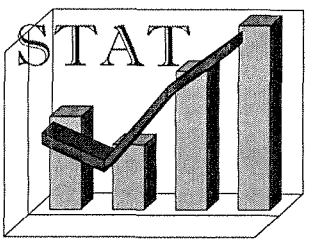
Collected Project Reports

Volume II

METHODOLOGY



Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

BIDE
Boston Institute for Developing Economies

Development Alternatives, Inc.

Volume I : Agriculture

#	Report #	Author	Title	Date	Page
1	4	Vijay Verma	Surveying Agricultural Households	June 00	I.1
2	23	Vijay Verma	Agricultural Census and Surveys	Mar 01	I.7
3	33	Suwandhi Sastrotaruno	Preparations for the 2003 Agricultural Census: Progress Report 1	May 01	I.21
4	39	Suwandhi Sastrotaruno	Preparations for the 2003 Agricultural Census: Progress Report 2	Aug 01	I.29
5	44	Suwandhi Sastrotaruno	Preparations for the 2003 Agricultural Census: Progress Report 3	Nov 01	I.72
6	48	Suwandhi Sastrotaruno/ Choiril Maksum	Aggregate Rice Data in Indonesia : A Brief Overview	Feb 02	I.95
7	53	Suwandhi Sastrotaruno	Preparations for the 2003 Agricultural Census: Progress Report 3A	June 02	I.110
8	61	Suwandhi Sastrotaruno	Preparations for the 2003 Agricultural Census: Progress Report 4	Sept 02	I.135
9	64	Yahya Jammal / Arizal Ahnaf	Is Susenas rice Consumption in Urban Areas of Indonesia Understated?	Nov 02	I.181
10	66	Suwandhi Sastrotaruno	Preparations for the 2003 Agricultural Census: Progress Report 5	Dec 02	I.216

Volume II : Methodology

#	Report #	Author	Title	Date	Page
1	1	Vijay Verma	Using Population Census 2000 to Improve BPS Integrated Household Survey Designs	June 00	II.1
2	2	Vijay Verma	Area Sampling for Economic and Agricultural Censuses Based on Population Census 2000	June 00	II.15
3	3	Vijay Verma	Notes on Sample Design and Re-Design for SUSI	June 00	II.31
4	7	Robert Torene	Evaluation of Establishment Directory	Aug 00	II.44
5	8	Robert Torene	Evaluation of SUSI Sample Design and Results	Aug 00	II.51
6	9	Robert Torene	Preparing a CHROMY Sample for Kabupaten Estimates	Aug 00	II.87
7	18	Sri Budianti Sukmadi	Use of Population Census 2000 to Improve Household Survey Frame	Dec 00	II.123
8	20	Vijay Verma	SUSI Estimation Procedures	Feb 01	II.161
19	21	Vijay Verma	Lectures Notes: Sampling Over Time	Feb 01	II.218
10	28	Sri Budianti Sukmadi	Area Sampling for Agriculture Census	Mar 01	II.231
11	32	Sri Budianti Sukmadi	Choice of Sampling Unit in the Agriculture Sample Census 2003	May 01	II.251
12	37	Sri Budianti Sukmadi	Sample Selection Method for Agricultural Sample Census 2003	July 01	II.267
13	43	Sri Budianti Sukmadi	Sample Design for the Second Pilot Planned under the Agriculture Census 2003	Nov 01	II.281
14	50	Sri Budianti Sukmadi	Geographical Integration of Household Surveys	Mar 02	II.298
15	52	Sri Budianti Sukmadi	Using Rolling Samples in Selected BPS Surveys	May 02	II.318
16	58	Sri Budianti Sukmadi	Linking Design and Organization in the Agriculture Census Sample	Aug 02	II.333
17	62	Vijay Verma	Notes on Estimation Procedures	Oct 02	II.351
18	63	Vijay Verma	Implementation of SUSI Estimation Procedures	Oct 02	II.372

Volume III : Labor Force					
#	Report #	Author	Title	Date	Page
1	5	Hananto Sigit	Employment Data in Indonesia : A Review of Existing Sources	June 00	III.1
2	10	Hananto Sigit	Earnings Data in Indonesia: A Review of Existing Sources	Aug 00	III.37
3	12	John Kuiper	Notes on Short-Term Labor Market Indicators	Aug 00	III.67
4	17	Hananto Sigit	*Survei Triwulanan Pemulihan Kesempatan Kerja*: A Proposed Survey to Monitor Short-Term Employment and Earnings Changes	Dec 00	III.77
5	25	Vijay Verma	Selected Issues in Labor Force Statistics	Mar 01	III.109
6	35	Uzair Suhaimi / Yahya Jammal	Measuring Open Unemployment in SAKERNAS	June 01	III.122
7	45	Hananto Sigit	Evaluation of the First Round of the Informal Sector Pilot Survey	Feb 02	III.140
8	65	Hananto Sigit	Evaluation of the Informal Sector Pilot Survey	Nov 02	III.162
9	68	Yahya Jammal/ Uzair Suhaimi	Quarterly Wage Survey of Livestock and Fishery	Apr 03	III.191

Volume IV : National Accounts					
#	Report #	Author	Title	Date	Page
1	6	Frank de Leeuw	Preliminary Evaluation of National Accounts and Prompt Indicators	July 00	IV.1
2	14	Yahya Jammal	Revision of Quarterly Production Accounts	Oct 00	IV.26
3	19	Frank de Leeuw	National Accounts and Prompt Indicators: Progress & Future Priorities	Feb 01	IV.61
4	24	Yahya Jammal	Quarterly National Accounts in Indonesia: Current Sources & Methods	Feb 01	IV.88
5	34	Yahya Jammal	Chain Linking National Accounts	June 01	IV.142
6	36	Frank de Leeuw	Unit Value Indexes for Exports and Imports	July 01	IV.160
7	42	Frank de Leeuw	Measuring Changes in Inventories in Indonesia	Sept 01	IV.180
8	51	Kusmadi Saleh	Indonesian Income Account: Progress Report 1	Apr 02	IV.200
9	54	Kusmadi Saleh	Indonesian Income Account: Progress Report 2	June 02	IV.206
10	55	Frank de Leeuw	Changes in Inventories Held by Manufacturers, 1998-99: Analysis of Annual Survey Data	Jul 02	IV.214
11	60	Kusmadi Saleh / Yahya Jammal	Toward Income Accounts for Indonesia	Sept 02	IV.232

Volume V : Industry

#	Report #	Author	Title	Date	Page
1	15	Yahya Jammal / Rosniaty Ismail	The Monthly Manufacturing Production survey	Nov 00	V.1
2	22	Sugiarto	Progress Report: Monthly Manufacturing Survey	Feb 01	V.68
3	30	Yahya Jammal / Vijay Verma	Monthly Manufacturing Production and Employment Indices	Apr 01	V.76
4	31	Sugiarto	Progress Report on Industrial Production	May 01	V.140
5	38	Sugiarto	Manufacturing Survey: Progress Report	July 01	V.149
6	40	Yahya Jammal / Rifa Rufiadi	Classifying Manufacturing Production by Commodity vs by Establishment Evidence Using 1998 Annual Survey Data	Aug 01	V.154
7	41	Sugiarto	Pilot for the Quarterly manufacturing Survey	Sep 01	V.196
8	47	Frank de Leeuw	An Expanded Industrial Production Index for Indonesia	Feb 02	V.207
9	49	Sugiarto	Planned Pilots for Manufacturing Surveys	Feb 02	V.234
10	56	Sugiarto	Progress on Pilots for Manufacturing Surveys	July 02	V.256
11	59	Sugiarto	Evaluation of Manufacturing Survey Pilots	Sept 02	V.291
12	67	Sugiarto/Suryamin	Capacity Utilization in Indonesian Manufacturing	Feb 03	V.308

Volume VI : Other Topics

#	Report #	Author	Title	Date	Page
1	11	John Kuiper	Seasonal Adjustments of Indonesia Time Series	Aug 00	VI.1
2	13	John Kuiper	Comparison of Implicit Price Indexes for Manufacturing	Aug 00	VI.50
3	16	Sugito Suwito	Enhancing the Role of Functional Staff and Improving Performance Evaluation	Nov 01	VI.81
4	26	John Kuiper	The Consumer Price Index : Selected Conceptual Issues	Mar 01	VI.98
5	27	John Kuiper	Converting the Wholesale Price Index into a Producer Price Index	Mar 01	VI.118
6	29	Bruce A. Johnston	An Interactive Survey Questionnaire for the Badan Pusat Statistik of Indonesia	Mar 01	VI.125
7	46	John Kuiper	Seasonal Adjustment for Lebaran	Feb 02	VI.140
8	57	John Kuiper	Operational Manual for Benchmarking Using the BENCH Program	July 02	VI.183
9	69	Yahya Jammal / Wendy Hartanto	Benchmarking Susenas Health and Education Data	May 03	VI.220
10	70	Yahya Jammal	STAT Project Summary Report	June 03	VI.

USING POPULATION CENSUS 2000 TO IMPROVE BPS INTEGRATED HOUSEHOLD SURVEY DESIGNS

Report # 1

by
Vijay Verma

June, 2000

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. CURRENT STATISTICS 3

II. PILOT STUDY USING THE HOUSEHOLD APPROACH 5
- **A. Comparison with Current Statistics** 5
- **B. Analysis of Variation** 7

III. PROPOSAL: A SURVEY OF FARMING HOUSEHOLDS 12

IV. AGRICULTURAL CENSUS 12

APPENDIX A BRIEF NOTE ON CURRENT AGRICULTURAL STATISTICS METHODOLOGY 14

I. CURRENT STATISTICS

This report is a follow-up to my earlier report *Surveying Agricultural Households*[1]. In this report, some more empirical evidence is examined. This reinforces the main conclusions presented in the earlier report. These are as follows:

1. Area under cultivation (and other current statistics) are collected routinely on a complete coverage basis. Complete coverage and high frequency of reporting on areas and other aspects are burdensome.[2]

2. In parallel, yields are obtained for paddy and other food crops using the crop-cutting method (with field work shared equally by extension agents and BPS statistical agents). While the sample size has been greatly reduced, from earlier 110,000 crop-cuts annually to around 30,000 in year 2000, this still remains a major and costly operation.

3. It appears certain that this system results in low quality of the data provided. The independence of the two systems for estimating areas and yields is prone to the danger of incompatibility of concepts and coverage, and hence to the danger of systematic biases in estimation of production computed as the product of two independent measures.

4. Some of the fluctuations from one year to the next appear very implausible. These suggest a rather uncontrolled data collection operation, and the presence of large and unstable biases in the results.

5. Comparisons with data on food consumption indicate that over-estimation of production has existed for a long time. It is believed that this results from over-estimation both in cultivated areas and in yields.

Table 1 presents some empirical information supporting the above statements. The table summarizes the reported change in cultivated area, yield and production between 1997 and 1998, as published by BPS. The figures are at the national level, based on complete coverage (in the case of areas) or very large samples (in the case of yield), and hence are practically unaffected by sampling error. At the same time, it appears very unlikely that such large changes in a single year could be real. Also, the pattern seems to lack consistency: whether we look across Java and Outer Islands, or across different crops.

---

1 STAT Project Report # 4, June, 2000.

2 See Appendix A below for an outline of the current methodology

For instance in Java, paddy and peanuts indicate a moderately large increase in the cultivated area between 1997-98, but an almost equally large reduction in yield, so that the recorded change in production is relatively small. Such a 'balance' is not expected.

By contrast, maize, mungbeans and sweet potatoes in Java all record very large (implausible?) increases in the area cultivated, nearly constant yields, and hence also very large increases in the total production. In Outer Islands, this pattern appears only for mungbeans.

Table 1
Percentage Change Recorded Between 1997 and 1998 in Area, Yield and Production

	Java			Outside Java			Indonesia		
	area	yield	production	area	yield	production	area	yield	production
paddy	**6.90**	**-6.99**	-0.58	**3.80**	**-3.56**	0.10	5.29	-5.30	-0.28
maize	**26.85**	-0.18	26.62	1.44	1.35	2.79	14.68	1.11	15.95
soybeans	-3.64	-1.66	-5.24	0.30	-1.41	-1.12	-2.15	-1.73	-3.78
peanuts	**5.70**	**-4.18**	1.26	0.07	-0.74	-0.61	3.65	-3.01	0.58
mungbeans	**12.40**	0.67	13.22	**18.52**	2.27	21.19	15.31	1.35	16.96
cassava	-5.28	0.00	-4.90	-0.31	0.92	0.10	-3.06	0.00	-2.89
sweet potatoes	**17.84**	-0.92	17.14	-4.55	1.16	-3.89	3.41	1.05	4.74

Recommendations:

1. It is highly desirable to examine the consistency and plausibility of the data from the survey. This should be done ***at the level of the greatest possible disaggregation***: by year, quarter, even by month where possible, for the long time series of the information which has been collected; and also by province, kabupaten, and even by individual kecamatan where possible. If changes as large as those in Table 1 can appear for the whole country, it is highly likely that some entirely implausible levels of change are present in the data when disaggregated to smaller units of time and space. 'Plausibility rules' should be established to edit and correct implausible data points.

2. With kecamatan if possible (at least kabupaten otherwise) as the unit, and by major crop, BPS should study the mean values and standard deviation of the basic variables: (i) reported yield; (ii) cultivated area/agricultural population in the unit; and (iii) production/agricultural population in the unit. (The denominators for (ii) and (iii) may be estimated from census and population projections, and from extrapolations of the information on the proportions in agriculture as required.) Such information is useful not only for evaluating the data, but also for sample design.

II. PILOT STUDY USING THE HOUSEHOLD APPROACH

A very large scale pilot study for the measurement of agricultural areas, production, yields and other variables was conducted in Java during 1996-97. The sample covered 6,000 wilcah, and as many as 262,000 households in Java. The results have been analysed and published by BPS in detail.

The system differs as follows from the current methodology:

Source	Measured	Computed
Current statistics	Area (A) 'eye estimates'; Yield (Y) 'crop cuts'	Production P = A*Y
Household approach	Area (A); Production (P) both from 'farmers reports'	Yield Y = P/A

A. Comparison with Current Statistics

Table 2 shows the level of difference from the current statistics approach, by 4-monthly periods and by province. Compared to the household approach, the current method 'over-estimates' cultivated areas by 20%, yields by over 25%, and hence production by over 50%. (Of course, alternatively this may be seen as 'under-estimates' from the household approach, compared to the current methodology, by 17% in area, 21% in yield and 35% in production. The truth may well lie in-between the two. Still, as to which approach is *more correct* needs to be established by a careful evaluation of the results.)

Note in particular the *above average* discrepancies for the January-April and the Jawa Timur area data, and for both the area and yield data for DI Yogyakarta. One source of the difference can be under-coverage of the sampling frame with the household-based approach. This would result in under-estimation of area (and production), but is unlikely to be an explanation for the big difference between the two sources in the yield values.

Table 2
Result of Pilot: Comparison of Existing System with Household Survey Approach (Paddy)

	(H) household survey			(M) MOA+cropcut			'over-estimate' (M-H)/H %		
	area	yield	production	area	yield	production	area	yield	production
jan-apr	1,984	50.1	99,458	2,690	63.3	170,274	35.6	26.3	71.2
may-aug	1,592	47.5	75,538	1,744	59.3	103,490	9.6	25.0	37.0
sep-dec	703	47.3	33,294	726	60.5	43,902	3.2	27.8	31.9
Jawa Barat	1,705	49.0	83,551	1,963	60.7	119,198	15.1	23.9	42.7
Jawa Tengah	1,290	47.9	61,738	1,551	61.3	95,029	20.2	28.0	53.9
DI Yogyakarta	79	46.5	3,666	102	64.5	6,597	29.7	38.7	79.9
Jawa Timur	1,205	49.3	59,342	1,543	62.9	97,017	28.1	27.6	63.5
total	4,279	48.7	208,315	5,160	61.6	317,947	**20.6**	**26.6**	**52.6**

Recommendations:

These are very large differences.

1. Firstly, it should be investigated whether the two systems are measuring the same things. For instance, are different types of land use – agricultural area, cultivated area, harvested area, etc. - being correctly distinguished? Also, is production by different end use – total production, own consumption, given as wages, animal feed, wastage, marketed, etc. - being correctly and consistently distinguished in the two systems?

2. In any case, in the collection of the data in either of the systems, separate reporting should be sought distinguishing area by land-use and production by end-use, so that the final variables can be constructed without confusion.

3. As for the current statistics noted in the previous section, it is highly desirable to examine the consistency and plausibility of the comparative data from the two sources ***at the level of the greatest possible disaggregation***: at least by 4-month period classified by province, and by individual kabupaten where possible. If differences as large as those in Table 2 can appear for the whole country, it is highly likely that some very extreme (hence highly implausible) differences are present in the data when disaggregated. *Comparing in detail these differences at the disaggregated level with the detailed time-series of comparisons from Recommendation 1 above may often provide a clear indication as to which source is likely to be more in error.*

B. Analysis of Variation

The study of mean values and standard deviation by kabupaten, both with the household and wilcah as the units of analysis, is useful for evaluating the data as well as for sample design. Such study should be done for all the basic variables by major crop.

With the household as the unit, this should include: (i) reported cultivated area; (ii) reported production; and (iii) computed yield. With wilcah as the unit of analysis, this should include: (i) reported cultivated area per agricultural household in the wilcah; (ii) reported production per agricultural household in the wilcah; and (iii) mean yield ('combined ratio' of the above two figures). Also would be useful for sample design purposes information on (iv) number of agricultural households in the wilcah, and (v) number of households in the sample from the wilcah.

The BPS have provided some information on mean values and variation of *yields* analysed in Tables 3 A-C. Table 3A shows the mean, standard deviation and their ratio (coefficient of variation) of paddy yield with the household as the unit. (The data have been provided by BPS by province and kabupaten, but in the tables here, breakdown by kabupaten has been shown only for one of the provinces.) The computed yields by household are quite variable (average cv=40%) – despite the claim sometimes made that there is little variation in these values. (In 60% of individual kabupaten in the province, estimated cv is in the range 30-45%.)

Table 3B shows the same computations performed with wilcah as the unit of analysis, i.e. it shows standard deviation and cv's for wilcah means[3]. The average cv among these values is surprisingly large, at over 25%.

Table 3C uses this information to estimate the design effects to which the sample for the Pilot is subject. Assuming that we are dealing with a simple random sample of wilcah, with a random sample of households within each sample wilcah, we have:

$$deft^2 = \left(\frac{S_w^2}{S_h^2 / \bar{b}} \right) = 1 + roh * (\bar{b} - 1)$$

Here $deft^2$ is the factor by which variance is inflated due to the use of a two-stage sampling design (wilcah, followed by households), and roh is the corresponding intra-cluster correlation coefficient, with b-bar as the average number of sample households per wilcah. S_h is the standard

[3] The mean values in Tables 3A and 3B differ somewhat because the latter presumably gives the simple mean on wilcah means in the kabupaten, while the former is the true per household mean.

deviation of household yields, and S_w that of wilcah means[4].

If wilcah were merely random groupings of households, we would expect:

$$S_w^2 = S_h^2 / \bar{b} ,$$

but S_w exceeds this value because of the positive correlation between households in the same wilcah.[5]

Table 3C shows extremely large values of design effect, with $deft^2 > 25$ (representing the factor by which the effective sample size has been reduced compared to a SRS of households), or deft>5 (representing the factor by which standard error has been inflated). This is the result of *very large cluster sizes* (65 sample households per wilcah), and also of the *very high degree of homogeneity* among households within the same wilcah (average roh=0.4).

Recommendations

These results have serious implications for the design of a household-based approach:

1. If such high values of roh are real, then this implies that ***much smaller clusters*** (i.e., numbers of sample households per wilcah) should be used. With roh=0.4, and for example 6 sample households per wilcah, we will have $deft^2=1+0.4*(6-1)=3.0$, or deft=1.7.

2. However, it is also possible (indeed likely) that these large roh values are caused in part by some problems with the data. How are the enumerators reporting area and yield values such that they result in so uniform yield values within wilcahs? It is necessary for BPS to investigate data at the micro-level, and also by individual enumerator.

3. The above points apply to variance of yield values. It is also necessary to analyse those for areas and production figures before final conclusions about the sample design can be drawn.

---

4 In the actual computations in Table 3C, I have used the cv's rather than the standard deviation (S) values, because the latter are affected by the slight differences in the mean values as a results of the manner in which these means have been computed. (See previous footnote.)

5 The BPS (unpublished) data provided to me also include an additional column in each of the tables A and B. This is labeled as 'standard error of the mean', and computed as the ratio [3]/[1] in Table A and as [7]/[5] in table B. While the latter is essentially correct as the actual standard error of mean yield, the former is not: it is what the standard error would have been with a simple random sample of households, i.e. with deft=1.0.

Table 3
Variation of Yield (pilot survey - household approach)

3A. Yield by household				
province	[1]	[2]	[3]	[4]
kabupaten	Households	mean	StDev	cv(%)
32 1	2,675	.472	0.209	44.2
2	3,048	0.359	0.180	50.2
3	3,987	0.436	0.171	39.2
4	6,357	0.458	0.159	34.8
5	5,431	0.480	0.135	28.2
6	4,915	0.445	0.133	29.9
7	5,353	0.455	0.193	42.5
8	6,157	0.421	0.152	36.2
9	6,900	0.438	0.152	34.8
10	2,704	0.445	0.143	32.2
11	1,837	0.495	0.145	29.3
12	3,571	0.481	0.177	36.8
13	3,397	0.445	0.123	27.6
14	3,882	0.458	0.159	34.7
15	3,830	0.475	0.156	32.8
16	1,428	0.400	0.168	42.0
17	2,901	0.549	0.128	23.4
18	1,863	0.500	0.226	45.1
19	2,013	0.414	0.221	53.4
20	2,814	0.484	0.227	46.9
71	5	0.452	0.268	59.2
72	38	0.561	0.224	40.0
73	154	0.491	0.180	36.6
74	1	0.350		
75	56	0.791	0.508	64.3
32 total	75,317	0.454	0.170	37.4
33 total	77,497	0.462	0.186	40.3
34 total	18,468	0.445	0.185	41.5
35 total	75,696	0.471	0.203	43.1
grand total	246,978	0.461	0.186	40.4

Table 3 (cont.)
Variation of yield (pilot survey - household approach)

3B. Yield averaged over wilcah

province		[5]	[6]	[7]	[8]
	kabupaten	no. of wilcah	mean	StDev	cv(%)
32	1	32	0.468	0.099	21.2
	2	45	0.358	0.124	34.7
	3	62	0.458	0.112	24.4
	4	101	0.474	0.083	17.5
	5	91	0.500	0.091	18.1
	6	75	0.439	0.082	18.6
	7	78	0.453	0.113	24.9
	8	89	0.434	0.078	17.9
	9	105	0.444	0.078	17.6
	10	40	0.443	0.074	16.8
	11	32	0.505	0.110	21.8
	12	51	0.482	0.065	13.4
	13	50	0.443	0.051	11.5
	14	63	0.459	0.092	20.1
	15	67	0.478	0.108	22.5
	16	23	0.413	0.054	13.2
	17	43	0.535	0.066	12.4
	18	33	0.496	0.087	17.5
	19	30	0.492	0.100	20.3
	20	47	0.457	0.125	27.4
	71	2	0.530	0.042	8.0
	72	3	0.603	0.060	10.0
	73	8	0.429	0.091	21.3
	74	1	0.350	0.000	
	75	3	0.593	0.527	
32 total		1,174	0.461	0.099	21.5
33 total		1,135	0.463	0.120	25.9
34 total		254	0.487	0.188	38.6
35 total		1,254	0.464	0.128	27.7
grand total		3,817	0.464	0.121	26,0

Table 3 (cont).
Variation of yield (pilot survey - household approach)

3C Design effects

province	kabupaten	hhs/wilcah [9]= [1]/[5]	deft2 [10]= [9]*{[8]/[4]}^2	roh [11]= {[10]-1}/{[9]-1}
32	1	84	19.2	0.22
	2	68	32.5	0.47
	3	64	24.9	0.38
	4	63	15.9	0.24
	5	60	24.6	0.40
	6	66	25.4	0.38
	7	69	23.6	0.33
	8	69	16.9	0.23
	9	66	16.8	0.24
	10	68	18.5	0.26
	11	57	31.6	0.54
	12	70	9.3	0.12
	13	68	11.8	0.16
	14	62	20.7	0.32
	15	57	27.0	0.46
	16	62	6.1	0.08
	17	67	18.9	0.27
	18	56	8.5	0.13
	19	67	9.7	0.13
	20	60	20.5	0.33
	71	3		
	72	13	0.8	-0.02
	73	19	6.5	0.30
	74	1		
	75	19		
32 total		64	21.3	0.32
33 total		68	28.1	0.40
34 total		73	62.9	0.86
35 total		60	24.9	0.40
grand total		65	26.9	0.41

III. PROPOSAL: A SURVEY OF FARMING HOUSEHOLDS

As recommended in a previous report[6], the proposal for consideration is to ***supplement the existing system by a regular agricultural survey at the household level.*** The survey will aim at overcoming the shortcomings of the current system. The already reduced crop-cut sample will perhaps release some resources for that purpose.

The pilot surveys already conducted should be analyzed further. It may be appropriate to begin with a limited geographical coverage before expanding the new approach to the national level.

One further advantage of a household based approach is that information on production and yields can be related to a host of other policy relevant variables, such as on farming practices, use of machinery, etc., and also to personal characteristics of individual farmers and agricultural households.

IV. AGRICULTURAL CENSUS

The planning and design of the next Census of Agriculture is a major and complex operation, and its many details are not considered in this report. Two main recommendations are made here.

Recommendations

1. The Census of Agriculture should be based on a large sample drawn from the last Population Census.

2. The household listing operation within the sample wilcah for the Census of Agriculture must be greatly simplified. This is because this is a large-scale operation which has to cover all households in sample wilcah. The collection of detailed information should be confined to the *sample* of 'potentially' agricultural households finally included in the Census of Agriculture.

The first recommendation has been briefly discussed in my earlier report *Area Sampling for Economic and Agricultural Censuses based on Population Census 2000.*[7] Of course, BPS needs to give much more detailed and careful consideration to the issues.

---

[6] *Surveying Agricultural Households, Ibid.*

[7] STAT Project Report # 2, June 2000

The second recommendation has the following background. Hitherto, the Census of Agriculture has involved three main enumerations:

(1) an operation to list households in selected wilcah;

(2) collection of quite detailed information for each household on whether it is eligible for the Census of Agriculture, and for which of the many questionnaires;

(3) application of the particular questionnaire to each eligible household, covering a separate sample for each application (sector).

This system is costly and complicated. The proposed alternative is to *eliminate* step (2), but to include in step (1) one or more simple questions to identify whether the household is 'potentially' an agricultural household; and to relegate the collection of detailed information to the selected sample in step (3). The elaborate FAO criteria for the inclusion of individual households into the Census of Agriculture can be based on the information collected during step (3), and refined at the *analysis stage*.

APPENDIX A
BRIEF NOTE ON CURRENT AGRICULTURAL STATISTICS METHODOLOGY

Cultivated Area

Current statistics on agriculture are collected as follows. From each Kecamatan agricultural extension agents (Mantri Tani) report (a) every month on area under paddy and other food crops, on vegetable production, and on area of damage by pests or calamity; (b) every quarter on fruit production; and (c) annually on area by land utilisation and use of agricultural machinery. This information is obtained on a complete census basis. The results are published annually by the BPS, separately by month, quarter, 4-monthly round or yearly as determined by the frequency of collection.

Yield

In parallel, yields are obtained for paddy and other food crops using the crop-cutting method. For paddy, sweet potatoes, peanuts and soybeans, crop-cuts over a random sample of 2.5 by 2.5 m subplots are used; for maize and cassava, the "row-subplot" variant (harvesting a specified length of plant row and measuring average distance between rows) is used. Field work is shared equally by extension agents and BPS statistical agents. The sample size has been large in the past - around 110,000 crop-cuts annually, each cut representing around 120 ha of cultivated land in the country – but reduced to 30,000 crop-cuts in year 2000 due to budgetary reasons. Field work is shared equally by extension agents and BPS statistical agents.

Production

Total production is estimated by multiplying yields and harvested areas estimated from the two independent systems.

Reporting Domains

For each category of production, the country is divided into two domains: (a) "potential areas" which cumulatively account for 90% or more of the total production in the category, and (b) the remainder, "non-potential areas". In (a) data are reported disaggregated to the Kabupaten level, and in (b) to the provincial level. It is clear that both in time and space the level of disaggregation presented is very detailed.

AREA SAMPLING FOR ECONOMIC AND AGRICULTURAL CENSUSES BASED ON POPULATION CENSUS 2000

Report # 2

by
Vijay Verma

June, 2000

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 17

II. ADDING INFORMATION FROM PAST ECONOMIC AND AGRICULTURAL CENSUSES 18
- **1. Some Procedures to Facilitate Matching** 19
- **2. Note On Estimation Procedure** 19

III. STRATIFICATION ON THE BASIS OF ECONOMIC INFORMATION 20

IV. SELECTION OF AREAS FOR ECONOMIC AND AGRICULTURAL CENSUSES 21
- **1. Sample Size for the Censuses** 21
- **2. Allocation Across Domains (provinces and kabupaten)** 22
- **3. Allocation by Sector (ISIC)** 23
- **4. Combined Allocation** 24
- **5. Technical Procedures for Implementing the Required Sampling Rates** 25
  - **1. Using CHROMY Algorithm** 25
  - **2. Directly using the strata allocations f_{ij}** 26

ANNEX I CONSTRUCTION OF "STRATA OF CONCENTRATION" 27
- **A. Objectives** 27
- **B. Data And Notation** 27
- **C. Definition and Refinement of the "Strata of Concentration"** 28

ANNEX II A NUMERICAL ILLUSTRATION CONSTRUCTED BY BPS STAFF 30

I. INTRODUCTION[1]

In a census of establishments, medium and large sized units are covered through a list frame, normally on a 100% (census) basis. By contrast, small and household-based establishments – whether agricultural or non-agricultural – are covered through a large sample of area units. This report is concerned with the latter.

In the last two rounds, Indonesia has conducted the Population, Economic and Agricultural Censuses as separate operations. Each of the three censuses independently involved two phases: pre-listing and construction of the frame, followed by actual enumeration for the census. The 1990 Population Census involved enumeration on a complete basis for basic items and on a sample basis for more detailed information. The other two censuses have been conducted, as is the common practice, on a large sample basis. Independent design and operations resulted in differences in the area frames on which the three censuses have been based.

The great reduction in the available resources necessitates a fundamental redesign and integration of these operations. Statistical considerations also require linking of the operations, at least to ensure that the various censuses are based on the same type – and indeed the same frame – of area units.

Population Census 2000 is providing a new frame of census blocks (CB's) covering the whole country. It will be efficient and convenient to use this frame directly for the Agricultural and Economic Censuses to be carried out, respectively, in 2003 and 2006 according to the present plans. The same frame can also be used as the basis for agricultural and economic surveys involving small-scale and household-based establishments, resulting in substantial improvements in the quality and cost-efficiency of those surveys.

A number of operations will have to be carried out on the Population Census 2000 CB Frame before it can be used for agricultural and economic censuses and surveys:

1. Adding to this frame information from the most recent Agricultural (1993) and Economic (1996) censuses.

2. Stratification of the CB's, making use of the information available in the frame, to suit the separate requirements of the two censuses.

3. The selection of a 'census sample' of CB's from Population Census 2000, again separately for each of the two censuses.

Preceding these steps is of course the construction of the Population Census 2000 CB frame itself.

---

[1] This report has been prepared promptly, largely during the Consultant's visit to the BPS. The various recommendations and details in the report should therefore be viewed as tentative, and may need to be revised and developed further at the time of implementation.

This has been discussed in the report 'Using Population Census 2000 to Improve BPS Integrated Household Survey Designs'.[2] The Population Census 2000 frame involves a complete list of CB's with the various population characteristics aggregated to the CB level. These include (i) geographic, administrative and other information on characteristics of the area; (ii) some demographic and social information on the population in the area; and (iii) various economic characteristics of the population.

Items (i) provide variables for the construction of 'primary strata', which define the reporting domains and are common to all types of household, economic and agricultural surveys and censuses. Items (iii) include information on distribution of the population according to broad sectors of agricultural and non-agricultural activity, which is useful for constructing frames and selecting samples for censuses and surveys on these topics. However, the information on sector of activity included in the Population Census 2000 is too limited for this purpose, especially for the Agricultural and Economic censuses which must provide very detailed information by ISIC. It is essential therefore to incorporate information from the most recent Economic and Agricultural censuses onto the Population Census 2000 frame.

II. ADDING INFORMATION FROM PAST ECONOMIC AND AGRICULTURAL CENSUSES

The operation to incorporate information from the last Economic Census is already underway at the BPS. Similar work in relation to the last Agricultural census should be started soon. This work can be carried out in parallel with constructing the basic Population Census 2000 CB frame, as the required information already exists.

The linking operations are entirely feasible, though time-consuming and somewhat complex. Difficulties arise because the economic and agricultural censuses have been based on different types of area units than CB's of Population Census 2000. Also, the former censuses are based on (albeit very large) samples, and hence do not cover all the CB's in the country. These difficulties are believed to be more serious in the case of the earlier agricultural census compared to the more recent economic census.

BPS work to-date shows that, at least in relation to the economic census, one-to-one matching with population census blocks can be successfully achieved in a vast majority of the cases. It is recommended that simple, even if crude, procedures be used to complete this process for the problematic cases. The objective is to enhance efficiency of stratification in the frame to the extent possible, rather than to produce actual estimates. In any case, only a small proportion of the areas appear problematic.

---

[2] STAT Project Report #1, June 2000.

1. Some Procedures to Facilitate Matching

(1) Ignore any differences in area boundaries.

(2) If an area in the earlier census corresponds to more than one areas in the population census, assign appropriately proportioned shares from former to all CB's in the latter. Quantities such as the number of establishments by sector may be distributed simply in proportion to the CB 'population' size. By 'population' here we mean the number of persons engaged in agricultural activities (for the Agriculture census frame) or in non-agricultural activities (for the Economic census frame), as recorded in the Population Census 2000. (Note. Such proportionate allocation can be done only after the processing of L2 information from Population Census 2000, at the beginning of the next year.)

(3) If information from the economic/agricultural census covers only a part of the CB area, or is missing altogether, go to a higher level – ideally to the smallest area for which the required economic information can be estimated, but even to the (urban/rural area of) the whole kabupaten if that is necessary. Then assign this to all CB's with missing information in this area, distributing it in proportion to the CB 'population' as above

(4) If the whole area was excluded from the earlier census (for instance, if it was considered 'non-potential'), then assign 'zero' to the number of establishments in the CB.

(5) Where the earlier census was based on a sample, care is needed in estimating the information for the total area represented by the areas enumerated in the census. In such a situation, the sample values should be used to produce ESTIMATES for the whole area or stratum they represented, and then that information assigned to all CB's lying in that stratum. See the note below.

2. Note On Estimation Procedure

If the boundaries of the area for which the estimate is to be produced coincide with a stratum used in selecting the sample in the earlier census, then the estimate is simply

$$= \Sigma^{n} (Y_i / p_i)$$

where Y_i is the value for unit (area) i selected in the census with probability p_i, and the sum is over all SAMPLE units in the stratum.

If the boundaries of the area for which the estimate is to be produced do not coincide with a stratum (or strata) used in selecting the sample in the earlier census, then the estimate is modified as follows:

$$= \Sigma^{n}\left(Y_i / p_i\right) * \frac{\Sigma^{N}\left(p_i\right)}{n}$$

where the first sum is over the n sample units (areas)in the corresponding stratum (or strata) of the earlier census as before, and the second sum is over N units in the POPULATION in the earlier census comprising the area for which the estimate is being produced.

III. STRATIFICATION ON THE BASIS OF ECONOMIC INFORMATION

In future economic and agricultural censuses, stratification is required simultaneously in two dimensions: (i) stratification to control geographic/administrative distribution of the sample; and (ii) stratification to control the distribution of establishments in the sample according to detailed sector of activity (ISIC).

Stratification (i) is required to meet the requirement for disaggregated reporting by province, kabupaten, urban-rural etc. The same, common system of geographic/administrative stratification can generally serve all types of censuses and surveys, whether household, agricultural or non-agricultural. Of course, the detail in which the stratification is applied may be varied depending on the size of the survey, and the surveys may also differ in the part of the population covered. In principle, such stratification is straightforward, simply involving a classification of the PSU's (areas) in terms of their location.

By contrast, stratification (ii) is more complex. It has to be developed separately for agricultural and non-agricultural sectors, reflecting their different types and distribution of establishments. In either case, the main complexity arises from the fact that, while the units to be stratified and sampled are AREAS, the characteristics used for defining the strata pertain to individual ESTABLISHMENT. The sample needs to control the coverage of establishments by detailed ISIC, yet the sampling units must be whole areas, each containing a mixture of ISIC's in general. Whatever the actual method of sample selection, it is obvious that control over and efficiency of the sample can be improved through stratification which divides the area units into strata on the basis of the type of establishments present in the area. A technique of stratification according to the predominating sector of economic activity is described in Annex I below. A

numerical illustration of the technique, constructed by BPS staff, appears in Annex II.

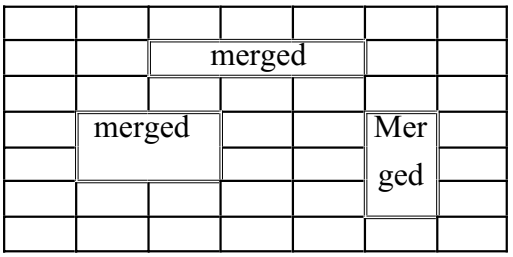
The overall system of stratification can therefore be seen as a cross-classification of two sets:

- 'strata of concentration', classifying areas (CB's) in terms of the predominating sector of economic activity (ISIC),

by

- classification of areas in accordance with the geographical domains (provinces, kabupaten, urban-rural) requiring separate reporting.

In practice, small cells in this classification may be merged to obtain strata of a 'reasonable' size, for example as illustrated in the following diagram:



IV. SELECTION OF AREAS FOR ECONOMIC AND AGRICULTURAL CENSUSES

The main issues requiring a decision are:

1. Sample size and allocation: What should be the overall sample size for the census, and how should the sample be distributed by geographical domain (provinces, kabupaten, urban and rural areas) and by sector (ISIC)?

2. Selection of units: How should the unit (CB) selection probabilities vary within the above strata? Should there be subsampling of establishments within sample areas?

1. Sample Size for the Censuses

The total sample size will be determined primarily by the size of the budget available, rather than by purely statistical considerations. Another point to note is that one should think in terms of the SAMPLE SIZE, and not the sampling rate. Rather than starting from some

assumption about the level of precision desired and working out the overall sample size needed for that purpose, practical constraints require one to begin with some feasible sample size and then work out the level of precision, and even more importantly, the degree of detailed breakdown of the results that it would provide. This information can then be fed back to adjust the sample size if needed, and this procedure repeated iteratively.

It is not possible to give a precise answer to the question of sample size at this stage. The following is a VERY TENTATIVE suggestion: Perhaps a sample of 30,000-50,000 census blocks can be considered. This would give an average of over 100 areas per kabupaten.

2. Allocation Across Domains (provinces and kabupaten)

The basic model is similar to that suggested for the master sample for household surveys.[3] The allocation has to be disproportionate across provinces, and increasingly among kabupaten as well in view of the need for decentralised reporting. As a compromise between the requirements of overall (national level) and domain (e.g. provincial level) estimates, a commonly used procedure is to allocate the sample in proportion to the square-root of province size. Note that 'size' here refers to the number of establishments as of the last economic/agricultural census and not to the size of the population or the number of households).

This basic allocation may be modified by two factors. Firstly, as in the case of household surveys, the allocation of smaller provinces may need to be further be increased in so far as smaller provinces also tend to have smaller kabupaten. Secondly, the economic/agricultural censuses would normally involve a single stage design: selection of areas followed by complete enumeration of the areas selected (i.e. without sub-sampling). This means that the allocation in terms of the number of establishments should be adjusted by average number of establishments per area to obtain the allocation in terms of the number of CB's to be selected.

The suggested model therefore is:

$$n_{i.} = n * \left(\frac{E_{i.}^{A-a} * K_i^a / B_i}{\Sigma \left(E_{i.}^{A-a} * K_i^a / B_i \right)} \right)$$

corresponding to the sampling rate (for the selection of areas)

---

[3] See my report 'Using Population Census 2000 to Improve BPS Integrated Household Survey Designs', STAT Project Report # 1, June, 2000.

$$f_{i.} = \frac{n_{i.}}{(E_{i.}/B_i)} = \frac{n_{i.}}{(N_{i.})}$$

where

i = a province

i. = the '.' notation is used to indicate that the quantity refers to the total for a province, summed over all sectors (ISIC)

n = overall sample size (number of CB's)

$n_{i.}$ = sample size (number of CB's) allocated to province i

$E_{i.}$ = size of the province (total number of establishments in all sectors, obtained through linking the last economic/agricultural census data to Population Census 2000)

K_i = number of kabupaten in the province

$N_{i.}$ = number of CB's in the province eligible for inclusion in the survey (for instance, it may exclude all areas definitely known to be without any establishments of interest)

B_i = average number of establishments per Population Census 2000 CB in the province = $E_{i.}/N_{i.}$

A = index for the basic allocation by province, = 0.5, say

a = a small number such as $0<a<0.25$, chosen on the basis of empirical information on the relationship between K_i and N_i in Indonesia.

Note also that such a model can provide no more than a starting point: the resulting allocations have to be individually checked and adjusted if required. Also, experience in Indonesia shows that in some of the smallest kabupaten, all the CB's may have to be included in the economic/agricultural census, i.e. without involving any sampling at this stage.

As to the allocation among domains such as urban-rural areas and kabupaten within a province, it will perhaps be simplest and most convenient to retain this allocation proportionate to the number of Population Census 2000 CB's in the domain.

3. Allocation by Sector (ISIC)

Again, the allocation should be disproportionate to ensure adequate sample sizes for small sectors. A suggested model is of the following form. The basic allocation has to be expressed in terms of the number of establishments (e.j), rather than the number of areas. In any case, the outcome is the set of required sampling rates by ISIC (f.j).

Allocation

$$e_{.j} = e_{..} * \left(\frac{E^S_{.j}}{\Sigma\left(E^S_{.j}\right)} \right)$$

with e.. as the estimate of total number of establishments in the sample, determined from quantities $n_{i.}$ and $B_{i.}$ defined above:

$$e_{..} = \Sigma_i \left(n_{i.} * B_{i.}\right)$$

so that the corresponding sampling rate is:

$$f_{.j} = \frac{e_{.j}}{E_{.j}}$$

where

j = a particular sector (ISIC)

.j = the '.' notation is used to indicate that the quantity refers to the total for a sector, summed over all provinces

$e_{.j}$ = sample size allocated to sector j (expected number of establishments, as of the last census)

$E_{.j}$ = size of the sector (total number of establishments in j, according to the last economic/agricultural census)

S = index for the basic allocation by ISIC

In comparison with the national versus provincial allocation discussed above, it seems appropriate to give higher importance to estimates for the total economy versus those for individual sectors. In other words, the allocation should perhaps be more nearly proportionate. This implies a larger value of S compared to A of Section IV.B, such as: S=2/3.

4. Combined Allocation

The combined effect of the allocations by province and by ISIC may be written as

$$f_{ij} = \frac{f_{i.} * f_{.j}}{f_{..}}$$

where f.. is a constant (like the overall sampling rate) determined to obtain the required total sample size when the f_{ij} are applied to individual Province-by-ISIC cells. That is:

$$f_{..} = \frac{n}{\Sigma_i \Sigma_j \left(f_{i.} * f_{.j} * N_{ij} \right)}$$

As noted earlier, any such model provides no more than a starting point: the resulting allocations should be individually checked and adjusted if required.

5. Technical Procedures for Implementing the Required Sampling Rates

Different procedures can be used for determining the appropriate sampling probabilities to be applied for the selection of individual CB's for the economic/agricultural census. Two such procedures are noted below.

1. Using CHROMY Algorithm

As used in the last SUSI, CHROMY algorithm may be used to determine selection probabilities (hence measures of size for PPS selection) of individual CB's. The input required is the number of establishments by ISIC x Area (E_{ij}), and a specification of target coefficients of variation (CV's) for the geographic-ISIC matrix (i,j) of the type defined above (Section IV.D.). The f_{ij} values there can first be used to determine the target sample sizes (in terms of the numbers of CB's to be selected for the economic/agricultural census) as:

$$n_{ij} = (f_{ij} * E_{ij}) / B_i$$

which can then be converted to corresponding CV's on the basis of information on population means, variances and assumed design effects. (Computations for such population parameters will need to be performed using data from past economic/agricultural censuses.)

2. Directly using the strata allocations f_{ij}

My report on SUSI[4] describes a procedure for determining sampling rates which can be applied directly to the matrix of geographic domain by 'stratum of concentration', so as to achieve the required rates f_{ij} by geographic domain and ISIC. Within these strata, the required number of CB's can be selected systematically with probability proportional to size (number of establishments in the area, as of the last economic census).

Notes:

(1) It is important that with either methodology of assigning selection probabilities to area units, the actual selection is done using a system of stratification of the type described above. This should be noted in particular if the CHROMY algorithm is used. The system of stratification is useful irrespective of the technique used for assigning the CB selection probabilities (or measures of size for PPS sampling).

(2) I have some doubt about the variance estimator used in CHROMY algorithm which, if correct, may influence the results through its effect on the presumed CV's. Please refer to the report 'Notes on Sample Design and Re-design for SUSI'.

(3) It is assumed that no subsampling of establishments within sample areas selected for the censuses will be involved. This will normally be the appropriate practice. In any case, the procedure can be easily modified to accommodate such subsampling.

---

[4] 'Notes on Sample Design and Re-design for SUSI', STAT Project Report # 3, June, 2000.

ANNEX I

CONSTRUCTION OF "STRATA OF CONCENTRATION"

A. Objectives

This Annex describes a technique for the stratification of area units according to their composition in terms of the type of economic activity (ISIC) present in the area. The description is in terms of economic censuses and surveys, but the approach applies equally to the agricultural sector. The basic idea is to classify areas according to the 'strata of concentration' as described below.

For each ISIC, the corresponding 'stratum of concentration' is defined to consist of a set of areas in which that ISIC predominates in the sense as defined below. One such stratum corresponds to each ISIC for which the survey results are to be reported separately.

The objective of construction of such strata is to separate out areas, to the extent possible, according to their composition in terms of the relative number of establishments of different types (ISIC). Each stratum is then sampled at a uniform rate to select areas (apart from any variation according to 'size' of the area, as in a usual PPS scheme). However, the sampling rate may be varied from one stratum to another with the aim to achieve the desired sample allocation by ISIC.

The main point of this stratification is to eliminate, or at least reduce, the need to apply different second-stage sampling rates for the selection of establishments belonging to different ISIC's within the same sample area. This is a very important practical consideration. Distinguishing establishment by ISIC at the second-stage requires the collection of additional information during listing, and is expensive, complex, difficult to implement, and prone to errors. The introduction of such stratification may also increase the efficiency and control in the design.

B. Data And Notation

The basic data consists of a list of areas, with the number of establishments of different types (ISIC's) in each area as of the last economic census. In addition, employment by ISIC may also be available.

The following notation will be used for the variables available or to be defined.

i subscript identifying an ISIC, i=1 to I

k subscript identifying an area

N_{ki} number of establishments of ISIC i in area k (frame)

$N_{k.}$ total number of establishments in area k (all ISIC's)

$N_{.i}$ total number of establishments in ISIC i (all areas)

A_i number of areas containing at least 1 establishment of ISIC i (Nki>0)

B_i average number of establishments of ISIC i per area (counting only areas containing at least 1 such establishment) = ($N_{.i}$ / A_i)

P_{ki} 'index of relative concentration' of ISIC i in area k = N_{ki} / B_i

R_{ki} rank of P_{ki} among all P_{kj} values (j=1 to I) for area k

R_{kj}=1 for the largest P_{kj} value

R_{kj}=2 for the next to largest P_{kj} value, etc

....

R_{kj}=0 for all j with N_{kj}=0 (i.e. for ISIC's with no establishment in the area)

$C1_k$ identification of the predominating ISIC ('stratum of concentration') in area k

$C1_k$=0, 1 to I:

=i corresponding to the ISIC (if any) for which R_{ki}=1 in the area

=0 if area k is empty (N_k.=0)

$C2_k$ identification of the 'second predominating ISIC' in area k

$C2_k$=0, 1 to I:

=i corresponding to the ISIC (if any) for which R_{ki}=2 in the area

=0 if area k is empty (N_k.=0) or contains establishments of only one ISIC

$N1_i$ total number of establishments of ISIC i in all areas with $C1_{k=i}$ (i.e. in areas with i as the 'stratum of concentration')

X_i the proportion of establishments of ISIC i 'captured' in its stratum of concentration = ($N1_i$ / N_i)

C. Definition and Refinement of the "Strata of Concentration"

To begin with, variable $C1_k$ as defined above identifies the 'stratum of concentration' of each area k. ISIC i=$C1_k$ is the one which predominates in this area.

However, the extent to which establishments of this ISIC are concentrated in such a stratum is indicated by variable Xi defined above. It is desirable and efficient that the Xi are as large as possible. How can the values of Xi be improved, particularly for ISIC's which happen to have the smallest Xi values?

It is for this objective that the variable $C2_k$, identifying the 'second predominating ISIC' in each area has been computed. Consider the following table classifying areas (k) in the frame by rows $i=C1_k$ and columns $j=C2_k$. Cells show the number of establishments, N1ij. The marginal columns show N1i. (see above), N_i., and their ratio $X_i=(N1_i / N_i)$. It is important to examine the X_i values.

The first and second strata of concentration:

	$C2_k$= 1	…	**$C2_k$= j**	…	$C2_k$= i	…	$C2_k$= I	in stratum	in ISIC	ratio
$C1_k$= 1										
…										
$C1_k$=j					N1ji to cell ij?					
…										
$C1_k$= i			N1ij					N1i.	Ni.	Xi= N1i/Ni
…										
$C1_k$= I										

Suppose that in the above table, ISIC i has a low X_i value, while X_j for ISIC j is high. Perhaps the X values in these two ISIC's can be made less unequal by shifting the areas and establishments from cell (j,i) to cell (i,j)? What this means is that areas in which j came first and i second, are treated as if i came first but j second. (Mechanically, these means interchanging the $C1_k$ and $C2_k$ values for areas for which $C1_{k=j}$ and $C2_{k=i}$, and then recomputing the above table.)

The above is merely an illustration. The point is that the distribution can often be improved by judgement based on careful examination of the pattern of the data.

Another critical choice is the form of P_{ki}, the 'index of relative concentration' of ISIC i in area k. It has been suggested above to define it purely in relative terms $P_{ki}=N_{ki}/B_i$, irrespective of the actual size of the ISIC. There is nothing to preclude empirical exploration of other forms in so far that improves (increases) the X_i values achieved.

ANNEX II

A NUMERICAL ILLUSTRATION CONSTRUCTED BY BPS STAFF

The attached tables, prepared by BPS staff, provide a detailed numerical illustration of the procedure described in Annex I. The data are taken from the 1996 Economic Census for one province, covering around 4,300 census areas containing a total of nearly 360,000 small or household-based establishments.

As shown in Table 1, the stratification is in terms of ISIC at the 2-digit level, consisting of 37 categories, 31 of which are represented in the data set. This also is the number into which the census areas have been divided. Each stratum corresponds to and is identified in terms of a particular ISIC: namely the ISIC which 'predominates' (in the sense defined in Annex I) in the areas comprising the stratum. A diagonal cell in the table shows the number of establishments of a particular ISIC which are 'captured' in their own 'stratum of concentration'. Often a majority of the establishment in the stratum in fact belong to some different ISIC (this can be seen from ratio of the last two rows in the table). However, for any ISIC, the important figure is along the row: it is the ratio of the diagonal cell to the row total, which gives the proportion ($X_{i.}$) of establishments in ISIC i captured in their own stratum of concentration.

Table 2 shows the classification of establishments according to the first and second dominating ISIC's. Such a table makes it possible to explore ways in which areas can be shifted across cells of the table so as to increase the proportions captured in diagonal cells of Table 1.

Table 3 shows the composition of each cell of Table 2 in more detail, namely the distribution of establishments in the cell by ISIC. Such exploration of the data is essential for gaining an appreciation of the underlying patterns, and for developing good, practical sampling procedures.

NOTES ON SAMPLE DESIGN AND RE-DESIGN FOR SUSI

Report # 3

by
Vijay Verma

June, 2000

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 33

II. CURRENT APPLICATIONS 33
- **A. The Frame** 33
- **B. Stratification** 34
- **C. Area Probabilities of Selection** 34
- **D. Selection Procedure** 34
- **E. Evaluation** 35

III. REDESIGN 35
- **A. The Frame** 35
- **B. Stratification** 35
- **C. Assigning Selection Probabilities to Areas** 35

IV. CHECKING THE RESULTS 35
- **A. Achieved Versus Planned Sample Size** 36
- **B. Sampling Efficiency** 36
- **C. Consistency of the Weights** 36

V. COMMENTS ON VARIANCE ESTIMATION 37
- **A. The CHROMY-GEN Model** 37
- **B. The Actual Selection Method Used** 38

ANNEX I DETERMINING SAMPLING RATES FOR "STRATA OF CONCENTRATION" 40
- **A. The Data** 40
- **B. Sampling Rate by Stratum** 41

I. INTRODUCTION[1]

SUSI ("Survei Usaha Terintegrasi") is a regular survey of small and household-based establishments covering all non-agricultural sectors in an integrated design. It is based on an area sample selected from Economic Census 1996. The survey began in 1998 with an annual sample of around 90,000 establishments. It was expanded to a quarterly survey in 1999, with a sample of around 30,000 establishments per quarter, or 120,000 for the year. However, for the year 2000, the sample size has been reduced to 15,000 establishments per quarter.

The survey may be redesigned every five years or so. From next year or soon thereafter, the survey will need to be redesigned to shift to the new frame provided by Population Census 2000. For this purpose the frame will be enhanced by incorporating into it information from Economic Census 1996.[2] It is appropriate therefore to begin discussion on the possible directions of this redesign. The next opportunity for redesign will come, perhaps, in year 2007 when the new frame from the Economic Census of 2006 becomes available.

A number of operations have to be considered in discussing SUSI design:

(1) Construction of the frame of area units
(2) Stratification of the areas to enhance control and precision of the sample
(3) The determination of first stage sampling probabilities
(4) The procedure for selecting area units
(5) Listing of establishments in the sample areas
(6) Sub-sampling within areas, if any
(7) Actual enumeration for the survey

II. CURRENT APPLICATIONS

A. The Frame

The frame of area units is provided by the Economic Census 1996. Among other things, the frame provides information by census area on its location (for geographic, administrative and urban-rural classification, etc.), and numbers of establishments and persons employed by ISIC in all non-agricultural sectors. The census did not cover all areas in the country, but is itself based on a large sample of areas. Presumably, 'non-potential' areas (areas with no or low level of economic activity) are excluded.

---

[1] This report has been prepared promptly, largely during the Consultant's visit to the BPS. The various recommendations and details in the report should therefore be viewed as tentative, and may need to be revised and developed further at the time of implementation.

[2] Please see my report 'Area Sampling for Economic and Agricultural Censuses based on Population Census 2000', STAT Project Report #2, June, 2000.

B. Stratification

The area frame is stratified and arranged by location variables. No stratification by the size and type of economic activity in the area has been introduced.

C. Area Probabilities of Selection

As in the case of Economic Census 1996, SUSI has used the CHROMY-GEN algorithm to determine the area selection probabilities. The basic input required is a list of 1986 census areas, giving information on the number of establishments and employment by ISIC. The procedure requires the specification of target CV's by 'item', taken in the present case as 3-digit ISIC at the national level, and 2-digit ISIC at the provincial level. BPS specified the CV's as a simple function of the size of employment in the 'item' concerned. The program outputs unit (area) selection probabilities which satisfy the CV constraints by item. The algorithm assumes that each unit is subject to selection independently with the specified probability, and that there is no subsampling at the second stage.

D. Selection Procedure

Though CHROMY-GEN assumes a particular method of sample selection, it does not by itself do the selection. Based on the area selection probabilities (or measures of size) output from the program, the SUSI sample has been selected by BPS using the standard PPS systematic procedure, with areas arranged according to the 'location' variables (see above). For practical reasons, BPS introduced sub-sampling within sample areas to limit the number of establishments which have to be enumerated. The sub-sampling rate varied according to the size category (number of establishments) of the area, with different rates applied to three different types of establishments within the same area.

The BPS has found the use of CHROMY-GEN program convenient and useful, though both in the method of selection at the first stage and in the fact of sub-sampling at the second stage, the actual procedures did not correspond exactly to the model underlying the program. In both cases, the choice made by BPS has the effect of improving the efficiency and practicality of the resulting sample.

Another departure, but this time not a desirable one, has been the practice of selecting a number of samples, and then take the one which appears the 'best' in some sense. This is not recommended in good sampling practice. The proper procedure is to adopt a sampling design with sufficient controls to rule out bad samples, but then accept the outcome of the sampling process. Subjectively selecting the 'best' sample can introduce bias – for instance under-estimation will result if the 'best' is always taken to be the 'smallest in size'!

E. Evaluation

Now that a working system has been established for SUSI, it is important to carefully evaluate the results of the survey, and identify whether any improvements can be made in the next redesign. See Section IV for some basic statistical measures for evaluation.

III. REDESIGN

A. The Frame

As soon as possible, SUSI should be moved to the new frame from Population Census 2000, augmented with information linked from the 1996 Economic Census. One convenient feature of the new frame is that it will cover the whole country (on a 100% basis), rather than only a sample of areas.

B. Stratification

It is important that with either methodology of assigning selection probabilities to area units, the actual selection is done using a system of stratification, not only geographic, but also taking into account the distribution of economic activity by type. For this purpose, a system of 'strata of concentration' may be constructed, as described in my report 'Area Sampling for Economic and Agricultural Censuses based on Population Census 2000'[3].

C. Assigning Selection Probabilities to Areas

Annex I below proposes an alternative procedure for determining selection probabilities. Given the target sample allocation by province and ISIC, the procedure determines the required sampling rate by stratum. Within strata, the areas sampling probabilities can be varied by size (number of establishments or employment in the area). The first stage sampling rates can be determined so as to eliminate, or at least to minimize, the need for sampling different types of establishments at different rates at the second stage. This procedure should be tested and evaluated in the same way as the output from CHROMY-GEN. Some useful statistical checks are noted below.

IV. CHECKING THE RESULTS

Three basic checks are described below:

---

[3] STAT Project Report #2, June, 2000.

A. Achieved Versus Planned Sample Size

Is the expected sample size and distribution by ISIC as planned? How do final n'_i compare with the target n_i? More strictly, in terms of the target sampling precision, the required comparison is between the target sample sizes, and the "effective" sample sizes achieved:

$$n_{i.} \quad \text{and} \quad \left(n'_{i.} / W_i^2\right),$$

where W_i, as described below, is the loss in precision (hence in the effective sample size) due to arbitrary weights introduced because of the non-uniform sampling rates within the same ISIC. (These sample sizes, of course, also need to be divided by the design effect due to clustering and stratification to obtain the true 'effective sample sizes'.)

B. Sampling Efficiency

How efficient is the sample? A main factor determining efficiency of the design for any particular ISIC is the variation in its selection rates across the strata. Such variation increases the sampling error. The factor by which variance is increased compared to a sample of the same size, but selected at a uniform rate is for the whole ISIC:

$$W_i^2 = \frac{\Sigma_j \left(N_{ij} * f_j\right) * \Sigma_j \left(N_{ij} / f_j\right)}{\left(\Sigma_j N_{ij}\right)^2}$$

The ratio

$$\frac{\Sigma_i \left(n'_i / W_i^2\right)}{\Sigma_i n'_i}$$

indicates overall efficiency of the allocation.

C. Consistency of the Weights

Does the weighted sample correctly estimates the population? Let Y_k be a measure available for all units (areas) in the frame, such as the number of establishments or employment in the area. Consistency of the weights is checked by the comparison

$$\sum_{k \in S} \left(W_k * Y_k\right) \quad \text{versus} \quad \sum_{k \in P} \left(Y_k\right)$$

where the first sum is over all units (k) in the sample, and the second sum all units in the population. The above check should be applied separately to each stratum from which a separate sample has been selected.

In relation to all these checks, it is important to look for extreme values which may need changes in the design or other adjustments. Also, if W_i is significantly above 1.0 (say >1.1), the target sample size n_i for the ISIC may have to be increased.

V. COMMENTS ON VARIANCE ESTIMATION

The following presents my understanding of the variance estimation in the CHROMY-GEN model. These views (and doubts) are presented for further discussion. The BPS has been requested to look into the matter further.

A. The CHROMY-GEN Model

The determination of unit (area) selection probabilities in CHROMY depends on the variance estimation formula used in the algorithm. The formula in turn assumes a particular method of selection, though CHROMY itself does not do the selection. Consistent application of the algorithm requires that the selection method actually used is the same as the one assumed. The assumed selection method and the resulting variance estimator is as follows.

It is assumed that each unit i is equivalent to a stratum and is independently subject to selection with probability p_i. The achieved sample size is a random variable, since there are two possible outcomes of this selection procedure independently for EACH unit in the population:

- With probability p_i: the unit happens to be selected, giving a sample estimate $=(Y_i/p_i)$, where Y_i is the value of the unit for some variable of interest.

- With probability $(1-p_i)$: the unit is not selected, thus giving a sample estimate of 0 (zero).

In either case, the true population value is Y_i, so that the contribution of the unit to variance is

$$Var(Y_i) = p_i * \left(\frac{Y_i}{p_i} - Y_i\right)^2 + (1-p_i)*(0-Y_i)^2 = \left(\frac{1-p_i}{p_i}\right)*Y_i^2 = (w_i - 1)*Y_i^2$$

where $w_i = 1/p_i$, the sample weight of the unit. Variance of the population aggregate Y= Y_i is

$$Var(Y) = \Sigma Var(Y_i) = \Sigma(w_i - 1)*Y_i^2 = \Sigma w_i * Y_i^2 - \Sigma Y_i^2 \qquad [1]$$

where the sum is over all units in the population. It may be noted that a sample estimate of the variance is given by the sum for units in the sample, each weighted by the unit weight w_i:

$$Var(Y) \approx \sum_{sample} w_i * \left[(w_i - 1) * Y_i^2 \right]$$

B. The Actual Selection Method Used

As in SUSI, and indeed as in most surveys of this type, the units are selected with PPS systematic sampling, which provides implicit stratification, and also ensures that the sample size is fixed and not a random variable. The selection procedure is essentially equivalent to the following. The population is divided into strata, each comprised of a set t of units such that summed over the set $p_i = 1$. From each set, exactly one unit is selected with its probability p_i, and (Y_i/p_i) is used to estimate total Y_t for the set. The contribution of the SET to the variance of the total therefore is

$$Var(Y_t) = \sum_{i \in t} p_i * \left(\frac{Y_i}{p_i} - Y_t \right)^2 = \left(\sum_{i \in t} w_i * Y_i^2 \right) - Y_t^2 = \left(\sum_{i \in t} w_i * Y_i^2 \right) - \left(\sum_{i \in t} Y_i \right)^2 \qquad [2]$$

while similar to [1], the CHROMY estimator of the same set is

$$Var(Y_t)_C = \left(\sum_{i \in t} w_i * Y_i^2 \right) - \sum_{i \in t} Y_i^2 \qquad [1']$$

The difference between the two is in the last term. To understand the difference, let p be the average selection probability of units in the set (w=1/p being the average weight, which is the same thing as the number of units in the set, from which exactly one is selected). Then Y_i is only of the order of p times Y_t, so that the last term in [1'] is also only of the order of p times the last term in [2]. Thus judged against the method of selection actually used in practice, the CHROMY estimator over-estimates the variance. The over-estimation can be large for strata consisting of many small units (small p's or large w's).

In so far as the algorithm determines the sampling rates to meet specified precision criteria, such an over-estimation results in over-allocation of the sample to smaller units, compared to what the allocation would have been on the basis of the actual variances.

In a separate exercise, a comparison has been made of the existing (single stage) sample of the Quarterly Survey of Manufacturing Establishments, with what the sample size and allocation would have been using CHROMY-GEN. The comparison was on the basis of fixed

CV's by ISIC in the existing design, as computed using the CHROMY-GEN formula. The main results of the comparison were that (i) CHROMY-GEN allocation gives more smaller establishments (i.e. fewer selected with certainty), and (ii) the required CHROMY-GEN total sample size for the same levels of precision appears to be 5% smaller than the existing design. It would appear that at least in part these results can be explained by the point made above concerning the CHROMY-GEN variance estimation formula. If it is correct that the variance contribution of smaller establishments is over-estimated, that would tend to shift the sample allocation in favor of such units in comparison with the true optimum allocation. Also, the absence of such a shift (as in the existing design) would tend to be penalized in terms of apparent increase in the sample size required for a given level of precision.

ANNEX I
DETERMINING SAMPLING RATES FOR "STRATA OF CONCENTRATION"

A. The Data

It is assumed that areas (PSU's) in the frame have been stratified according to the 'strata of concentration' as defined in Annex I of my report 'Area Sampling for Economic and Agricultural Censuses based on Population Census 2000'[4]; and that the required sample sizes (number of establishments) by ISIC (the 'reporting domains') have been determined. These domains may be defined for the whole country, or separately by province etc when the sample sizes are sufficient for more detailed reporting.

For generality, we assume that the sample of establishments is to be selected in two stages – selection of areas, followed by the selection of establishments within sample areas. The procedure of course also applies when there is no subsampling within areas. The objective is to determine the overall sampling rate to be applied within each 'stratum of concentration' so as to achieve the specified sample size and its distribution by ISIC.

As explained in the above mentioned Annex, there is a one-to-one correspondence between the ISIC's and the strata. For each ISIC, the corresponding 'stratum of concentration' is composed of the set of areas in which that ISIC predominates in some sense, as defined there. The basic information is represented by the following square table. Subscript i (rows 1 to I) refers to ISIC, and j (columns 1 to I) to stratum. Generally, any ISIC is distributed over various strata, and any stratum contains establishments from various ISIC's. However, the DIAGONAL elements (i=j) predominate to the extent that ISIC's are geographically segregated, and each tends to be concentrated within its 'own' stratum.

N_{ij} refers to the number of establishments (as recorded in the sampling frame) of ISIC i which lie in sample areas in stratum j. Summing across columns,

$N_{i.}$ is the total number of establishments in ISIC i. Summing across rows,

$N_{.j}$ is the number of establishments in stratum j,

$n_{i.}$ is the target sample size for ISIC i, and

f_i $= (n_{i.} / N_{i.})$ is the required overall sampling rate for ISIC i.

Empty areas

Note that account must also be taken of 'empty' areas which, at the time the frame was prepared, contained no establishments of any type. Since these areas may no more be empty at the time of the survey, a sample must be selected from them – albeit at a low sampling rate.

---

[4] STAT Project Report #2, June, 2000.

Classification of areas and establishments by ISIC and the 'strata of concentration'

Stratum	1	…	j	…	I	total
ISIC						
1	i=j					
…		i=j				
i			Nij			Ni., ni.
				i=j		
I					i=j	
total			N.j			N.., n..
						'empty areas'

Ordering of strata and ISIC's

For the application of the procedure described below, it is convenient to order the strata (columns) according to increasing value of the ratio

$$N_{ij} / N_{.j} \quad (i=j)$$

i.e. the extent to which the diagonal element (i=j) predominates in the stratum.

Then the rows (ISIC's) are arranged to follow exactly the same order, i.e. such that column j=i gives the 'strata of concentration' of ISIC i.

B. Sampling Rate by Stratum

Given the target sample sizes $(n_{i.})$ for ISIC i, the required sampling rate for the ISIC is: $f_i = n_{i.} / N_{i.}$

To start with, we apply these rates to the corresponding STRATA, i.e.

[1] $$f_j = n_{i.} / N_{i.} \quad (i=j)$$

Note that we will consistently use subscript i (rows 1 to I) for ISIC, and j (columns 1 to I) for stratum. Thus a quantity such as $N_{i.}$ stands for the sum over strata for ISIC i, and $N_{.j}$ for the sum over ISIC's for stratum j. The diagonal cell is indicated by (i=j), and all off-diagonal cells by (i j).

Application of [1] gives a total sample of size

$$n' = {}_j (f_j * N_{.j}).$$

Next the f_j defined above are modified by a constant overall factor so as to obtain the required overall target sample size n:

[2] $$\text{new } f_j = \min[1, (n / n')*f_j]$$

Rate [2] yields the required sample size overall. However, its distribution by ISIC may be distorted. The sample size obtained for ISIC i is

$$n'_{i.} = \Sigma_j (f_j * N_{ij})$$

consisting of

[3]
$$n1_{i.} = (f_j * N_{ij}) \quad |j = i$$

coming from the diagonal element, plus

[4]
$$n0_{i.} = \Sigma_j (f_j * N_{ij}) \quad |j \neq i$$

from the remaining strata.

Correcting the achieved sample size by ISIC

We adjust the sampling rate for the 'strata of concentration' of each ISIC i (i.e. for column j=i, passing through the diagonal cell), so as to obtain the required sample size n_i for the ISIC. This is done by starting with the first row. (Note that ISIC's are assumed ordered as indicated in Section A above.)

[5]
$$f'_j = \min\left(1, \quad \frac{n_i - n0_i}{n1_i} * f_j\right) \quad |j = i = 1,$$

with no change for the other strata (columns):

[6]
$$f'_j = f_j \quad |j \neq i$$

With the revised f'_j values, we compute n0 for the next ISIC (i=2) using [4]. (Note that n1 value is not affected as f_j, j=i, remains unchanged due to previous adjustments applied to j<i.) Then we use [5] to obtain revised f_j values.

This is then repeated for i=3, 4, …I in turn, using at each step the most recent values of f_j. This completes one iteration.

Iterative cycle

The process [2]-[6] is repeated iteratively, starting with the f_j values resulting form the previous iteration. The iterative process may be stopped when

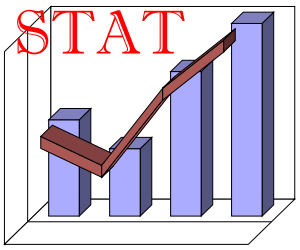
- all (n_i, n'_i) differ by <x (5?)%, and the total (n, n') by <y (1?)%,
- or the number of iterations exceeds z (100?).

EVALUATION
OF
ESTABLISHMENT DIRECTORY

Report # 7

by
Robert Torene

August, 2000



Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 46

II. PRELIMINARY REVIEW PROCEDURE 46

III. CONCLUSIONS & RECOMMENDATIONS 48

IV. DEVELOPMENT OF AN INTEGRATED DIRECTORY USING ADMINISTRATIVE DATA 49

I. INTRODUCTION

The BPS has constructed a Directory of Non-Manufacturing Establishments with Legal Entity. The task is to determine if a CHROMY sample can be drawn for such a Directory. The answer is twofold: **Yes and No**.

Yes, CHROMY can be used to determine probabilities for selecting samples from a Directory. As a strictly mechanical exercise it is totally possible to select a sample from the Directory. It is a relatively simple task.

No, a sample should not be selected from the Directory unless further review and verification indicates the Directory coverage is adequate. I suspect, and the preliminary review that we have conducted indicates, that the Directory procedure may be flawed and the coverage of the Directory is seriously lacking.

II. PRELIMINARY REVIEW PROCEDURE

To test the accuracy and completeness of the Directory we compared the 1999 SUSI listing for 12 wilkers in North Jakarta against the Directory. The results are shown in the table below.

Number of establishments in SUSI Listing 1999

Kecamatan	Kelurahan	Wilcah	NKS	Wilker	SUSI Establishments	Matched Directory	Out of Scope
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tanjung Priok	Kebon Bawang	31B0	210641	01	5	2	1
	Sunter Jaya	22B0	200613	01	1	-	-
Pademangan	Ancol	06B0	120613	02	2	1	-
		13B0	220619	01	8	3	-
	Pademangan Timur	15B1	210586	01	1	-	-
	Pademangan Barat	11B0	220586	01	1	-	-
Cilincing	Semper Timur	12B0	220893	01	2	-	-
	Cilincing	07B0	120886	01	5	2	-
Koja	Tugu Selatan	06B0	210684	01	2	-	-
	Lagoa	37B2	200714	01	9	3	-
Kelapa Gading	Pegangsaan Dua	14B1	220858	01	5	-	-
Penjaringan	Pejagalan	02B0	220509	01	4	1	
Total					**45**	**12**	**2**

The results were alarming. Of the 45 SUSI establishments with legal entity only 12 were matched to the Directory. It was decided to do some field checking in two of the wilkers to see if some of the reasons could be discovered. The assistance of the North Jakarta Statistics Office was requested and the original SUSI enumerators were assigned to assist in the checking.

For each unmatched establishment we attempted to locate the establishment and determine the reason that it was represented in the Directory. The following questions were asked for each establishment:

- Name and address
- Year started operation
- Type of business activity
- Legal entity

Of the 14 establishments with legal entity listed in the two wilkers, 9 were not in the Directory. The SUSI listing was conducted in June 1999. The latest Directory updating procedure took place in November 1998. All of the establishments interviewed were in operation before November 1998 when the Directory updating procedure took place, so they should have been listed in the Directory. The table below shows the result by establishment. Gray lines are matched establishments.

The Results of Visiting Two SUSI Wilkers

Propinsi : DKI Jakarta - Kota : Jakarta Utara - Kec. : Tanjung Priok - Kel. : Kebon Bawang
Wilcah : 31B0 - NKS: **210641** - Wilker : 01

VUSI99-L			The Results of Visiting		
Name	**Address**	**Legal Entity**	**Name**	**Address**	**Remarks**
PT. Jikian Triputra	RT. 0018/06 No. 1	2	PT. Jikian Triputra	Same with VUSI99-L	No name board
SPTSK	Jl. Bugis No. 1	6	Serikat Pekerja Tekstil, Sandang, Kulit	"	Out of scope of both SUSI and PD
PT. Yanita Mandiri	Jl. Bugis No. 1D	2	PT. Yadika Mandiri	"	No name board
PT. Tari Jaya Sentosa EMKL	Jl. Raya Bugis No. 146	2	PT. Talie Jaya Sentosa	"	No name board
PT. Akselatama Indoputra	Jl. Raya Bugis RT. 005/01	2	Akselatama Indoputra, PT	Bugis 122, l.	Matched to directory
PT. Nusa Cakrawala Sejati	Jl. Raya Bugis RT. 005/01	2	Nusa Cakrawala Sejati, PT	Bugis 22, Jl.	Matched to directory

Propinsi : DKI Jakarta - Kota : Jakarta Utara - Kec. : Pademangan - Kel. : Ancol
Wilcah : 13B0 - NKS : **220619** - Wilker : 01

VUSI99-L			The Results of Visiting		
Name	**Address**	**Legal Entity**	**Name**	**Address**	**Remarks**
PT. Laju Karya	J. Tongkol 6Q	2	Laju Karya Perkasa	Tongkol Blok 6 Jl.	Matched to directory
PT. Inti Manis	Jl. Tongkol No. 6Y	2	CV. Inti Manis	Same with VUSI99-L	No name board
PT. Tan Bun Siong	Jl. Tongkol No. 6Z	2	PT. Sinar Budi Packaging	"	No name board. Name different
PT. BMS	Jl. Tongkol No. 6 AD	2	PT. Bintang Mas	"	No name board
PT. Tata Cipta	Jl. Tongkol 6 AH	2	PT. Tata Cipta Pelangi	"	No name board
PT. Sinar Rahayu Abadi	Jl. Tongkol 6	2	Sinar Rahayu Abadi	Tongkol No. 6F PS. Ikan, Jl.	Matched to directory
PT. Rimba	JL. Krapu No. 10	2	PT. Rimba Kumala	Same with VUZI99-L	No name board
PT. Rukun Gaya Baru	Jl. Krapu No. 5	2	Rukun Gaya Baru PT	Tongkol. RT. 09/10 Jl.	Matched to directory

The results of our interviews show that the major reason these establishments were missed is that they ***do not have signboards*** in front of their establishments. These establishments provide sales and services to other business establishments, not the general public, and do not feel that they need a signboard for their customers to find them. They stated that their customers knew where to find them or that their customers found them through other means: yellow pages, internet, infomedia, trade shows, trade publications, etc.

III. CONCLUSIONS & RECOMMENDATIONS

- **Further checking needs to be done to review the accuracy and coverage of the Directory.** Although our review was very limited, it is a distinct possibility that there is a problem with the current Directory updating procedure. The SUSI listing process is a complete door to door canvass – a census of the enumeration area. The Directory updating process is not and is subject to substantial undercoverage.
- **To fully verify the accuracy of the Directory, select a representative sample of about 50 SUSI wilkers and perform the same type of coverage check.**
- **It would be wise to not conduct any surveys based on the Directory until its accuracy and coverage is determined more fully.**

• It is probable that the undercoverage is concentrated in establishments who service or sell to other establishments or institutions rather than the general public.

• Alternative directory building procedures should be seriously investigated. This is further discussed below.

IV. DEVELOPMENT OF AN INTEGRATED DIRECTORY USING ADMINISTRATIVE DATA

In October 1998, BPS staff and this consultant met with members of the Dirjen Pajak technical staff in charge of computer processing. The results of the meeting were encouraging from a technical perspective.

• Coverage is probably the most comprehensive and complete. Any large company must have an NPWP registration number to qualify for government contracting or belong to professional or business organizations. There were about 500,000 Badan registrations of which about 350,000 are mailed annual tax forms (Formulir 1771).

• The Master File of Registrations contains appropriate information for constructing a directory including the KLUI code and the NPWP of the Pusat if the establishment is a Cabang. This will allow BPS to develop company structures by linking all establishments with the same Pusat NPWP. Tracking of multilevel companies should be also be possible since there should always be a NPWP trail to the next highest level.

• Data on NPWP registrations is 100 percent computerized and resides on the computer in the Pajak headquarters. This information is updated from the local tax offices twice each month.

• The NPWP application contains the following information:

- Name and address
The address requested is the physical location, not simply a post office box. One of the requirements for an NPWP is the business location permit, which in turn requires a domicile permit. This is important for BPS data collection procedures.
- Status Usaha
Applicants report whether the business location is for the central office (pusat), branch (cabang), or a single-unit establishment (tunggal).
- Number of branches/Branch serial number
If the application is for a central office, they report the number of branches. If the application is for a branch, they report the serial number of the branch.
- NPWP of the central office
If the application is for a branch, they report the NPWP of the central office for the branch. This is extremely important for the directory since it will allow the BPS to develop company structures by linking all establishments with the same Pusat NPWP.
- Business activity code
This code is from the "Lapangan Usaha Wajib Pajak", an

industrial classification taken from the BPS KLUI coding system.
- Form of ownership code
This is a code assigned by the DirJen Pajak based on the form of ownership in the Firm Registration Act (Akte Pendirian/Akte Perubahan).

This should be a good time to approach the DirJen Pajak for this information since DirJen Pajak staff recently visited the BPS to request the BPS directory. In addition, USAID has an advisor, Geoff Walton, in the DirJen Pajak who recently met with Yahya Jammal and this consultant. Perhaps this is the most favorable time to obtain at least a test file to see if the coverage of the Pajak directory is as good as it seems.

EVALUATION
OF SUSI SAMPLE DESIGN
AND RESULTS

Report # 8

by
Robert Torene

August, 2000

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. SURVEY DESCRIPTION 52

II. DISCUSSION OF VIJAY VERMA'S REPORT 53
- **A. Current Applications** 53
  - **1. The Frame - "Non-Potential" Areas** 53
  - **2. Stratification** 53
  - **3. Selection Procedure** 54
  - **4. Clustering Effect** 54
- **B. Redesign** 55
  - **1. The Frame** 55
  - **2. Stratification** 56
  - **3. Assigning Selection Probabilities to Areas** 56
- **C. Comments on Variance Estimation** 57

III. PERFORMANCE OF SUSI SAMPLE COMPARED TO EXPECTED RESULTS 58

IV. QUESTIONNAIRE DESIGN CHANGES 61
- **A. Legal Status** 61
- **B. Type of Establishment** 62

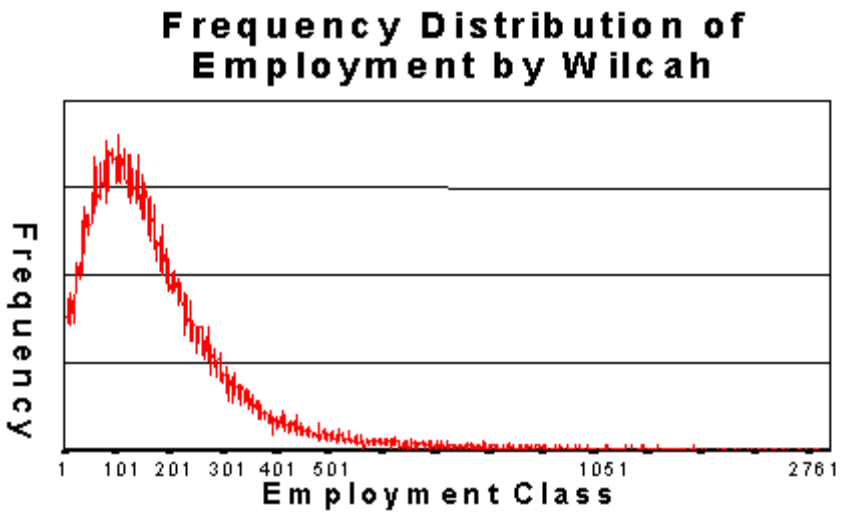
I. SURVEY DESCRIPTION

The SUSI Survey is an innovative area survey of small and micro enterprises with no legal entity. It uses an integrated, multi-dimensional approach to sampling of enumeration areas based on their total contribution of all targeted individual industries to national and provincial totals. This was made possible by the use of CHROMY-GEN, a program designed by the U.S. Census Bureau to provide minimum cost stratified samples. This survey is the subject of a paper given at the 2000 International Conference on Establishment Statistics in Buffalo, New York. It is attached (Attachment A) to provide further introduction to this survey.

Surveys of economic activity (especially area surveys) are among the most difficult sampling tasks.

- There are many characteristics that have to be considered and the number of first-stage primary sampling units is generally limited.
  - Industry classification (KKKP) and level of publication,
  - Type (establishment, enumeration area) and size of unit,
  - Levels of geography (provinces, kabupatens, etc.),
- Economic activities tend to cluster, i.e., businesses tend to locate in specific and concentrated areas and in totally different patterns than demographic data.

The size of economic activities varies greatly. This is especially true in list samples (where individual establishments can be extremely large or extremely small) but is also true in area samples as enumeration areas can vary greatly in size due to the tendency of business establishments to locate in specific and concentrated areas.



The chart above shows a frequency distribution of the enumeration areas (wilcahs) by employment size. Because of the clustering of economic units, the distribution is a typical Poisson distribution (L shaped with a long tail stretching to the right) where the larger wilcahs contribute disproportionately to the estimates.

Wilcahs with 301 or more pekerja account for only 14 percent of the wilcahs (11.127 of 77.717) and 30 percent of the total estimate for pekerja. The same is true for wilkers although to a slightly smaller extent: 15 percent of wilkers with more than 300 pekerja account for 29 percent of total pekerja.

II. DISCUSSION OF VIJAY VERMA'S REPORT

Vijay Verma's report, "Notes on Sample Design and Re-Design for SUSI"[1] has suggested that a stratified sample design would provide more reliable estimates than the current Poisson sample design using the CHROMY-GEN program. I disagree with his conclusions.

Before discussing his suggestions for an overall sample re-design I would like to review several statements that are in Section A ("Current Applications").

A. Current Applications

1. The Frame - "Non-Potential" Areas

On page 1 when discussing the Frame, Vijay Verma stated that "non-potential" areas (with no or low level of economic activity) are presumably excluded from the 1996 Census sample frame. This is incorrect. These areas were sampled at relatively high rates but are fully represented in the sample frame.

2. Stratification

On page 2, Vijay Verma states: "The area frame is stratified and arranged by location variables. No stratification by the size and type of economic activity in the area has been introduced." Although this statement is technically correct, it does not take into account the CHROMY-GEN process which inspects each enumeration area for its contributions to both size and type of economic activity. By its very nature and procedure, CHROMY implicitly stratifies for these variables at every level for which a CV constraint exists. In the SUSI sample this means that there is implicit stratification as low as Indonesia 3-digit and Province 2-digit.

In fact, this is one of the major benefits of CHROMY since (1) there are only minimal preparations of the sample frame before sampling can occur, and (2) stratification and targeting of the sample frame in the same way as CHROMY implicitly does would be an impossible task.

[1] STAT Project Report No. 3, June, 2000.

3. Selection Procedure

In this section Vijay Verma takes issue with the practice of selecting the "best sample". He states that "subjectively selecting the 'best' sample can introduce bias – for instance underestimation will result if the 'best' is always taken to be the 'smallest in size'!"

Perhaps some clarification will help. Many unbiased samples were selected, each starting with a new random number. Each sample was tabulated and compared to the Census totals for Indonesia at various KKKP levels. The "best" sample was the one with the smallest accumulated absolute differences from the census totals. This was done very "objectively" in the sense that a computer program was written to determine the absolute differences from each sample to the census totals and each sample was then ranked in order of increasing absolute differences.

At first, I was also somewhat concerned that this procedure was not "unbiased" and discussed it with Richard Sigman, the U.S. Census Bureau mathematical statistician who designed the CHROMY-GEN program. He stated that we are introducing efficiencies into the sampling process with this procedure and that if the chosen sample was selected using an initial random number and the CHROMY generated probabilities it is an unbiased sample.

I also raised the question and discussed our procedure at a session of the International Conference on Establishment Statistics titled "Estimation Strategies Using Variants of Poisson Sampling". The panelists, including Esbjorn Ohlsson, Stockholm University; Bengt Rosen, Statistics Sweden; and Pedro Saavedra, ORC Macro; all agreed that the resulting sample is unbiased.

Further, I would put these questions to Vijay Verma: "If the chosen sample was selected as the first and only sample, would it be unbiased? If so, how can it be a biased sample if selected with the procedure described above?" His email responses to Yahya's email raising these questions seem to assume that we are trying to determine the "best" sample using some rule other than the best replication of the universe, which is not the case. We have no other motive than to choose a sample that reproduces the census totals as closely as possible.

4. Clustering Effect

Although I found no reference to the clustering effect in his paper, I was told that Vijay Verma also suggested that the SUSI sample produces a clustering effect, which is not taken into account in the variance calculation. If I am misquoting, I apologize, but I would like to discuss this subject.

First, as stated above, by its very nature economic activity is highly clustered. Businesses tend to locate in specific and concentrated areas. Second, the issue of the clustering effect has more to do with a distribution of selected primary sampling units by geography and industry that is not representative of the actual distribution in the sample frame. By the nature of the sampling process used for SUSI, any clustering effect should be minimized, if not eliminated.

As Vijay Verma states on page 6, "… the units are selected with PPS systematic sampling, which provides implicit stratification, ….". The nature of the CHROMY program design further implicitly stratifies by the type of CV constraints that are used. As such the SUSI sample is implicitly stratified by Indonesia 1-, 2-, and 3-digit KKKP and by Province 1- and 2-digit KKKP. Both industry and geographic (Province) variables are implicitly stratified in this way. In addition, the sample was selected with the sample frame in geographic sequence, which implicitly stratifies by location variables.

B. Redesign

1. The Frame

Vijay Verma states that the SUSI sample should be moved to the new frame from the Population Census 2000. I disagree. The 1996 Economic Census sample frame cannot be improved upon by using Population Census data. There is also considerable doubt that converting enumeration areas to the Population Census census blocks can be achieved reliably. Further, there is little correlation between data on economic activities and population characteristics. Economic activities do not cluster in the same way as demographic data. I certainly agree that the sample frame can get very old and that it should be updated, if at all possible. However, using Population Census 2000 data is not the way to do that.

Recommendation 1: The Economic Census sample frame should remain as is with the same enumeration areas until the next economic census when the Population Census 2000 census blocks can be used.

Recommendation 2: For all future censuses, Population, Economic, and Agriculture, maintain "permanent" census blocks from census to census. I realize that Census Blocks are used to even out enumerators' work load but any changes should be designed to minimize BPS preparations and promote comparability and consistency. For example, target for a workload of 75-150 households/economic units per census block. If the census block rises above 150 units, cut the census block in half within the existing boundaries. If the census block falls below 75 units, merge it into a neighboring census block. In this way the census blocks will be much more comparable and consistent from census to census while maintaining an approximate workload.

Recommendation 3: For future censuses, I would suggest that the BPS should consider adding directory type information on Population Census L1 such as:

- **Is there any economic activity at this location?**
- **If so, describe the activity.**
- **How many persons are working in this activity?**

This information will provide the data necessary to update the economic sample frame every 5 years rather than 10. Coding of the activity description will be necessary but at least the raw material for a new sample frame will be available.

2. Stratification

Vijay Verma recommends that "… the actual selection is done using a system of stratification, not only geographic, but also taking into account the distribution of economic activity by type." The SUSI sampling scheme implicitly stratifies by area and industry. This topic is more fully discussed above.

He proposes that a "…. system of 'strata of concentration' may be constructed…….". **Using the CHROMY-GEN sampling system makes this type of stratified sample frame totally unnecessary**. CHROMY-GEN itself implicitly stratifies the SE-96 sampling frame in much greater detail than is possible with Vijay Verma's system.

Recommendation 4: Do nothing with respect to construction of a stratified sample frame for economic statistics. Use of CHROMY-GEN makes the construction of stratified sample frames unnecessary, especially for SUSI and other economic surveys. In general, CHROMY-GEN simulates the construction of its own stratified sample frames from the census data. It should be noted that the current SUSI sample design is a type of stratified sample. It is modified so that each first-stage sampling unit is a separate stratum and has its own individual probability of selection.

3. Assigning Selection Probabilities to Areas

Vijay Verma proposes an alternative procedure for assigning selection probabilities to wilkers that depends on the construction of "Strata of Concentration" proposed earlier. While, in theory, this stratification technique could produce a representative sample there are several drawbacks. In practice, the construction of "Strata of Concentration" is difficult and time consuming. It requires a lot of manual intervention and judgement.

It is very clear from the example tables in Verma's report, "Area Sampling for Economic and Agricultural Censuses Based on Population Census 2000" that the "Strata of Concentration" do not represent the wide range of ISICs (only 2-digit ISICs are used for stratification) that exist in the various strata. Verma recognizes this and further proposes that a second variable, the "second predominating ISIC" be computed to improve the stratification.

This second variable is then used to "make it possible to explore ways in which areas can be shifted across cells of the table so as to increase the proportions captured in diagonal cells of Table 1." I suggest that it is necessary to consider the work load and decisions required in actually preparing these tables and then making the decisions about "shifting" areas across cells. Is the amount of work worthwhile?

Contrast the above procedure with the SUSI Poisson sample design using CHROMY. By its very nature and method of operation where each wilker is a separate stratum, CHROMY inspects the contribution of each ISIC designated by a CV constraint from each wilker. Overall wilker probabilities are assigned not only on the major and secondary ISIC, but on every ISIC represented in the Census data. AND it is basically automatic. There is no need to go through the

time consuming process of constructing a sample frame based on "Strata of Concentration". From a practical viewpoint, the CHROMY-GEN program is much easier to use and produces minimum cost samples.

Recommendation 5: Continue using the CHROMY-GEN sampling program and expand its use to other surveys. CHROMY-GEN when used in a Poisson sample design is much more suited to economic surveys than is a stratified sample design. Currently, the BPS is preparing an interactive version of CHROMY that will make the selection of samples much easier and less time consuming and will prepare all relevant files for both stratified and Poisson sample designs

Verma does make two points that are important and need to be considered.

- **Empty Areas**

He states, "Note that account must also be taken of 'empty' areas which, at the time the frame was prepared, contained no establishments of any type. Since these areas may no more be empty at the time of the survey, a sample must be selected from them – albeit at a low sampling rate."

Recommendation 6: The SUSI sample should be reviewed to determine if empty areas were omitted from the sample. If so, a small additional sample should be selected to represent that portion of the universe. The CHROMY-GEN program can be configured to assign minimum probabilities to each EA to avoid the occurrence of this situation for future sample designs.

- **Differential sub-sampling by sector**

Verma states, "The main point to this stratification is to eliminate, or at least reduce, the need to apply different second-stage sampling rates for the selection of establishments belonging to different ISICs within the same sample area. This is a very important practical consideration. Distinguishing establishment by ISIC at the second-stage requires the collection of additional information during listing, and is expensive, complex, difficult to implement, and prone to errors."

I fully agree. However, I don't believe that Verma's stratification plan will eliminate this problem. Merely reducing it will still require considerable work in the field and still require the time consuming procedure of creating the "Strata of Concentration".

Recommendation 7: To eliminate or reduce differential sub-sampling rates by sector, experiment by reselecting a CHROMY sample, tighten up on the CV constraints for sectors where we want a higher sub-sample rate and loosen up the CV constraints for the sectors (notably sector 6) where we want fewer sub-sample establishments.

C. Comments on Variance Estimation

With respect to Vijay Verma's comments about overestimation of variance with the SUSI estimator, I am doing the following:

- Making available a report titled "The Annual Survey of Manufactures: A Report on Methodology". (This is a rather lengthy report but can be made available to

anyone interested. Please see Yahya Jammal for copies of this report.) The report contains a lengthy description of the Poisson sampling technique with many mathematical proofs. This survey was among the first (perhaps the first) to use Poisson sampling. The actual selection procedure is somewhat different than used for the SUSI. For the Annual Survey of Manufactures, random numbers (R_h) are drawn for each primary sampling unit (PSU). A PSU is selected whenever $R_h + P_h >= 1$. This, however, results in an indeterminate sample size. To ensure a fixed sample size, the PSUs in SUSI were selected with PPS systematic sampling.

• Sent a copy of Vijay Verma's report on SUSI to Richard Sigman, of the U.S. Census Bureau and the designer of the CHROMY program for his comments.

III. PERFORMANCE OF SUSI SAMPLE COMPARED TO EXPECTED RESULTS

The 1998 SUSI has performed very well in comparison with the results expected when the sample was first drawn. Table 1 below compares the distribution of number of units and pekerja by tabulated coefficient of variation (CV) in percent. Data are shown for the original sample selected from the 1996 Sensus Ekonomi with comparable data from the 1998 SUSI.

The 1998 coefficients of variation are somewhat higher than the expected values based on the 1996 tabulation of the sample. However, the 1996 CVs shown here were not calculated based on a sub-sample procedure. Although the sub-sample does not contribute greatly to the calculation of the variance there is some effect.

The sampling strategy for the SUSI sample was that the larger KKKP, those with a higher number of pekerja, would be assigned lower CVs. It was decided that these larger cells should be more reliable than smaller cells since they dominate the estimates. This strategy has worked quite well as shown in the table.

National Data:

At the 2-digit Indonesia level of tabulation 52 percent of the pekerja are in tabulated cells with CVs or 3 percent or less, and nearly all (98.1 percent) of pekerja are in cells with a CV of 15 percent or less.

As the sample begins to get thinner these rates decline but are still very respectable. At the 3-digit KKKP level for Indonesia 25 percent of the pekerja total are in KKKP with CVs of 3 percent or less, nearly 80 percent in KKKP with CVs of 5 percent or less, and over 96 percent in KKKP with CVs of 15 percent or less.

Even at the Indonesia 4-digit KKKP level, for which CV constraints were not supplied, almost 50 percent of the pekerja total can be published with CVs of 5 percent or less and 90 percent in KKKPs with CVs of 15 percent or less.

Province Data:

At the province level the sample begins to get very thin and the tabulated CVs at various levels of KKKP are correspondingly higher. This is reasonable since the major determinant of CV is the number of wilkers in the Province sample. No province total has a CV of 3 percent or less although Provinces with CVs of 5 percent or less account for 41 percent of total pekerja. All Province totals have CVs of 15 percent or less since that was a maximum constraint that was supplied to CHROMY-GEN for the Province totals.

The sample for KKKP data at the Province level gets even thinner and the CVs correspondingly higher but there is still a large body of data with very respectable CVs.

At the Province 1-digit KKKP level 24 percent of the pekerja totals are in cells with CVs of 5 percent or less; 71 percent in cells of 10 percent CV or less; and 91 percent in cells of 15 percent or less. Even at the Province 3-, and 4-digit KKKP levels for which CV constraints were not supplied substantial proportions of the data have CVs. at a 10 percent and 15 percent level. At the Province 4-digit KKKP level about 45 percent of the pekerja are in cells with a 15 percent CV or less while at the Province 3-digit KKKP about 61 percent of pekerja are in KKKP with 15 percent or smaller CVs.

Table 1 : Comparson of SUSI 98 with Original SE-96 90,000 Sample

KKKP Level	CV for Pekerja (%)	SUSI 98 Sample				SE-96 90,000 Sample			
		Units	% of total	Pekerja	% of total	Units	% of total	Pekerja	% of total
Indonesia 1-digit	3% or less	12.432.619	88,2	22.953.995	88,0	14.859.468	93,4	23.907.579	93.3
	5% or less	13.533.407	96,0	24.930.482	95,6	15.550.312	97,7	24.879.069	97.1
	10% or less	14.094.087	100,0	26.060.714	100,0	15.913.845	100,0	25.615.342	100,0
	15% or less	14.094.087	100,0	26.060.714	100,0	15.913.845	100,0	25.615.342	100,0
Indonesia 2-digit	3% or less	8.158.666	57,9	13.616.083	52,2	11.932.965	75,0	17.377.053	67,8
	5% or less	12.448.376	88,3	22.060.714	84,7				92,8
	10% or less	13.821.889	98,0	25.228.445	96,8				97,7
	15% or less	13.947.548	98,9	25.582.414	98,1				99,1
Indonesia 3-digit	3% or less	4.065.284	28,8	6.556.749	25,2				56,6
	5% or less	11.910.748	84,5	20.812.141	79,8				81,9
	10% or less	13.418.407	95,2	24.122.589	92,5				94,1
	15% or less	13.811.559	98,0	25.143.505	96,4				97,3
Indonesia 4-digit	3% or less	1.872.997	13,3	3.588.586	13,8				35,6
	5% or less	8.197.471	58,1	12.977.348	49,8				61,4
	10% or less	12.464.599	88,4	21.890.657	84,0				86,2
	15% or less	13.152.397	93,3	23.536.140	90,3				92,6
Province Totals	3% or less	0	0,0	0	0,0				0,0
	5% or less	5.992.488	42,5	10.771.621	41,3				60,4
	10% or less	11.679.976	82,8	21.359.773	81,9				86,3
	15% or less	14.099.025	100,0	26.070.199	100,0		100,0		100,0
Province 1-digit	3% or less	0	0,0	0	0,0				
	5% or less	3.585.663	25,4	2.164.611	23,8				
	10% or less	10.220.597	72,5	14.587.061	71,0				
	15% or less	12.935.057	91,7	19.959.959	91,3				
Province 2-digit	3% or less	0	0,0	0	0,0				
	5% or less	1.266.616	9,0	0	8,3				
	10% or less	8.520.524	60,4	10.472.062	56,0				
	15% or less	11.322.428	80,3	16.061.643	76,6				
Province 3-digit	3% or less	0	0,0	0	0,0				
	5% or less	0	0,0	0	0,0				
	10% or less	6.565.294	46,6	6.454.818	40,2				
	15% or less	9.392.571	66,6	11.887.851	61,6				
Province 4-digit	3% or less	0	0,0	0	0,0	0	0,0	0	0,0
	5% or less	0	0,0	0	0,0	111.510	0,7	189.969	0,7
	10% or less	4.124.757	29,3	6.454.818	24,8	6.598.462	41,4	8.639.665	33,7
	15% or less	7.376.872	52.3	11.887.851	45.6	9.645.670	60.6	13.580.868	53.0

In summary, for the national and province data the sample is providing a great deal of very high quality data, even for KKKP levels that were never considered to be candidates for CV constraints.

A proposed list showing the number of KKKP levels that could be published for both Indonesia and the Provinces using a 20 percent CV criteria is shown as Attachment B. A proposed table heading is shown below. This type of table would be extremely helpful to users as it would summarize all major publishable information in a single table for Indonesia and a similar table for the Provinces.

Proposed Table Heading for Integrated Table of Major Data Items

KKKP		Number of		Balas			Value	Relative Standard Error for:	
Code	Short Title	Units	Pekerja	Jasa	Omzet	Biaya	Added	Pekerja	Omzet
See proposed list of KKKP in Attachment B									

Note: One table for Indonesia data and one table for Province data.

IV. QUESTIONNAIRE DESIGN CHANGES

A. Legal Status

The question on Legal Status (Badan Hukum) is a mixture of Form of Ownership and Other Government Certificate or License. It is confusing and makes it more difficult to clearly identify the kind of establishments we are dealing with. I propose that this question be broken into two questions and should be on the SUSI questionnaire as well as the listing sheet. This is especially important on the listing sheet so that we can check any directories that the BPS will construct in the future. I realize that SUSI is a survey of establishments without legal status, however, this may be necessary if the BPS gets access to Pajak registers. It will be important to be able to record the actual form of ownership as well as any other government licenses and certificates held by the establishment.

1. Legal Form of Ownership

PT

CV

Partnership
Single proprietor
Etc.

2. Other Government Licenses/Certificates
SIPD
Etc.

B. Type of Establishment

There was a wonderfully informative question on the 1996 PND/URT questionnaire about the type of establishment that somehow has gotten lost on the SUSI questionnaire. While I can't remember the exact form of the question, it had to do with the physical status of the establishment. This question gave excellent information about the type of establishment, if it has a structure, whether the goods remained overnight, and the conditions under which it operates.

Physical status of the Establishment
Tempat tetap
Tempat tidak tetap
Etc.

ATTACHMENT A:
INTEGRATING NON-AGRICULTURAL SURVEYS INTO ONE: AN INDONESIAN CASE STUDY

Robert Torene, Weidemann Associates; Robert Aten, USAID, Jakarta;
Sri Budianti Sukmadi, Statistics Indonesia; Richard Sigman, U.S. Bureau of the Census;
Yahya Jammal, Boston Institute for Developing Economies
Contact Robert Torene, Weidemann Associates, 1490 Jordan Ave., Crofton, MD 21114 USA
rtorene@erols.com or rtorene@aol.com

Developing/emerging nations have a difficult problem collecting economic data. Their economies have two distinct and important parts: the traditional composed of many small establishments and the modern composed of large heterogeneous establishments. Statistics Indonesia has designed a two-part program. (1) For the traditional sector they are conducting an **innovative annual area sample survey of small enterprises covering all non-agricultural sectors.** For the difficult sampling this entails they are using a unique multivariate sample allocation program (2) For the modern sector they plan to develop a directory based on administrative sources from which they will select a list sample. The eventual goal is to combine the results of these surveys in a seamless way and publish comprehensive, timely annual data for all non-agricultural economic sectors by industry for Indonesia and its provinces. This approach can provide a useful model for other developing/emerging nations attempting to adequately cover their traditional as well as their modern economies.

Key Words: Economic Surveys, Integrated Sampling, CHROMY-GEN

I. Introduction

The international donor community is helping Indonesia improve the quality of its industrial production data and other key aggregates because such data are relevant to a number of key policy objectives. First, the International Monetary Fund has established key statistical standards for developing countries. Among these standards are economic performance data. Second, economic growth in Indonesia over the period from the early 1980s to 1997 has relied heavily on the concentration of exports of industrial commodities. The economic downturn of 1997 and 1998 directly occurred in these key sectors, particularly in manufacturing for the Indonesian domestic market. Thus, understanding the behavior of the industrial sector in Indonesia is key to understanding the likely behavior of the Indonesian macroeconomy. However, as Indonesian development continues, the role of the financial and other sectors in establishing and sustaining a rapid pace of economic growth becomes ever more critical. Third, small and microenterprise businesses, despite their limited impact on economic growth in the Indonesian macroeconomy, are nonetheless central to providing employment for the poor. Finally, in the absence of good data about the actual behavior of the Indonesian economy, bad economic decisions that have adverse impacts on economic growth and poverty alleviation are ever more likely. Good data helps to drive out bad decisions.

For many economic questions, time series data must be utilized. Rational economic analysis requires detailed statistical data about economic performance for industries over a period of years. The conclusion is that good economic analysis relies enormously on continuing availability of the kinds of detailed statistics describing the economic structure and annual performance of the economy, industry by industry. This Statistics Indonesia program is an opportunity to provide the kind of data most economic policy planners and analysts desperately need but rarely have: comprehensive, timely, annual data for all non-agricultural economic sectors by industry for Indonesia and its provinces.

II. The Problem

Most emerging/developing countries face difficult problems collecting data on their economic structure. From the viewpoint of statistical data collection, they have distinctly dual economies: one traditional and one modern. Different survey techniques are appropriate for these significantly different economies. Developed countries have predominantly modern economies and conduct their economic censuses and surveys in a relatively standard fashion. They develop mailing lists from administrative sources and generally collect the data by mail, for both large and small establishments. While such an approach, which requires a list sample drawn from a previously prepared directory, can be relied on to cover the modern economy, it cannot be used to cover the traditional economy. Only an area sampling

approach can be used for this purpose. However, given the relatively large size of the traditional economy in developing/emerging nations, such an approach may prove too costly. That is why it has so far been avoided to a large extent. This paper reports on a program started by Statistics Indonesia which, if successful, can provide a useful model for other developing/emerging nations attempting to adequately cover their traditional as well as their modern economies.

III. Description of the Statistics Indonesia Program

Indonesia may be unique among developing countries in having well-established and developed statistical systems and traditions with a decades long history. Among its regular activities is a decennial economic census which is supposed to cover all establishments in the country.[1] This is complemented by:

- an annual census of manufacturing establishments employing 20 or more workers. The main reason for such an activity was to help predict the country's annual GDP (since manufacturing was one of the most important sectors and these establishments dominated it in terms of value added),
- irregular surveys of small and "cottage industry" manufacturing establishments, which were conducted on average once every three years.[2]

Data for sectors other than manufacturing (e.g. mining, trade, services, construction, transportation) came from intermittent and infrequent surveys using different samples. Overall, therefore, significant gaps in the data existed. Statistics Indonesia recognized these shortcomings and limitations and requested USAID to assist in a review and improvement of their economic statistics program. As a result, Statistics Indonesia decided to design a new program for collecting economic statistics which would allow coverage, within one integrated system, of establishments of all sizes and in all sectors. The priority was, of course, on areas where deficiencies existed, namely in coverage of non-"medium and large" manufacturing establishments and establishments in all other non-agricultural sectors of the economy.

A. Establishment Classification: "Directory" vs "Non-Directory"

A statistician needs to deal with clearly defined and unambiguous boundaries. That is why he/she devises concepts like enumeration areas which can be mapped and have boundaries that can be clearly identified in the field. When designing a system of economic censuses and surveys, he/she tries to divide the non-agricultural economy into clear, distinct and unambiguously measurable units. This is generally more difficult than defining a physical unit like an enumeration area, because the boundaries, being conceptual rather than physical, cannot be seen or readily identified. The concept must reflect conditions that exist in the real world, thereby making it easier to identify the target units.

The eventual goal of this program is to divide the universe - the Indonesian non-agricultural economy - into two parts: **(1) Directory of large establishments (the list sample universe) and (2) everything else (the area sample universe).** Table 1 provides a graphical representation of these two universes.

Table 1: An Economic Map of Indonesia as Viewed by a Statistician

Sector 2, Mining and Quarrying	Sector 3, Manufac-turing	Sector 4, Electricity, Gas, and Water	Sector 5, Construc-tion	Sector 6, Wholesale and Retail Trade	Sector 7, Transporta-tion and Communi-cation	Sector 8, Real Estate, Financial, and Business Services	Sector 9, Personal and Other Services
		Directory Establishments					
		Area Survey Establishments					

[1] The latest census, conducted in 1996, estimated that the number of establishments in the country was ?? million, of which ?? million were engaged in "cottage industry" type activities.

[2] Statistics Indonesia follows the International Standard Industrial Classification (ISIC) definition of "establishment" but differentiates them according to employment size: those employing 20 or more workers are classified as "medium and large", those employing 5-19 workers are classified as "small" and those employing 1-4 workers are classified as "cottage industry".

Table 2: 1998 Publishable Data (20% Relative Standard Error Criterion)

Industry Level	Number of Industries			Publishable Data for Indonesia			Publishable Data for Provinces		
	Total	In sample	Pub-lished	No. of Units	No. of Workers	Receipts (mill. Rp.)	No. of Units	No. of Workers	Receipts (mill. Rp.)
Total				14,074.3	26,020.2	537,385.1	14,074.3	26,020.2	537,385.1
1-digit	8	8	6	14,064.4	25,994.2	537,008.9	12,438.9	22,770.3	494,716.7
% of total				99.9	99.9	99.9	88.4	87.5	92.1
2-digit	34	33	16	13,905.4	25,497.2	532,978.5	10,828.8	18,797.0	408,896.3
% of total				98.8	98.0	99.2	76.9	72.2	76.1
3-digit	95	91	27	13,547.9	24,656.7	523,789.3	8,557.2	14,324.1	284,169.2
% of total				96.3	94.8	97.5	60.8	55.1	52.9
4-digit	282	268	56	13,000.4	23,290.7	479,186.3	6,882.3	11,305.8	193,458.6
% of total				92.4	89.5	89.2	48.9	43.5	36.0

Table 3 below shows the reliability of the tabulated data for various industry levels for Indonesia and the Provinces for the 1998 sample. At the 2-digit industry level for Indonesia over 92% of total employment are in cells with a relative standard error of 5% or less. At the 3-digit industry level over 94% of total employment are in cells with a relative standard error of 10% or less. Even at the 4-digit industry level 35.6%, 61.4%, and 86.2% of total employment are in cells with a relative standard error of 3%, 5% and 10% or less respectively.

Table 3: Tabulated Employment Data By Relative Standard Error

Relative Standard Error	Percent of Total Employment at Specified Relative Standard Error Levels							
	Industry Level for Indonesia				Industry Level for Provinces			
	1-digit	2-digit	3-digit	4-digit	Total	1-digit	2-digit	3-digit
3% or less	93.3	67.8	56.6	35.6	0.0	0.0	0.0	0.0
5% or less	97.1	92.8	81.9	61.4	60.4	32.1	19.5	4.5
10% or less	100.0	97.7	94.1	86.2	86.3	73.6	61.3	45.4
15% or less	100.0	99.1	97.3	92.6	100.0	94.7	82.1	68.9

III. Statistics Indonesia Progress to Date

The SUSI survey has greatly increased the available data on small establishment-household activities in Indonesia. According to the government, in the next five year planning cycle (year 2000-2004) small establishments will be a priority concern. Emphasis will be given to improving management capabilities, creating supply linkages with larger firms, granting of credit, etc. To evaluate the condition of small and microenterprises Statistics Indonesia plans to continue the collection of data on small establishments-household activities through the SUSI survey.

The survey began in November 1998 with a sample of 90.000 establishments, using a time reference of one year. This was followed up in 1999 by quarterly integrated surveys with a sample of 30.000 establishments in each quarter (total of 120.000 establishments in the year 1999). Due to the economic crisis and limited government budget the sample for SUSI 2000 was reduced to only 60.000 establishments to be enumerated in quarterly samples of 15.000 establishments. We expect that this limited sample will produce reasonably consistent results with the 1998 and 1999 SUSI surveys.

Profile of Small Establishments

Reflecting the financial and monetary crisis that began in 1997 the SUSI data showed a steep decline from 1996 to 1998 of 16.1% in the number of establishments and 9.9% in workers. (See Table 4 below.) There was a slight recovery from 1998 to 1999 as the number of establishments increased 2.2% with workers increasing 1.3% for all of Indonesia. The average number of reported workers increased from 1.72 in 1996 to 1.85 and 1.83 in 1998 and 1999, indicating that the smaller enterprises had more difficulty staying in business during this economic crisis.

This division into two distinct and complementary universes may seem simplistic. However, one should point out, that it is true ***regardless of how the Directory is defined.*** In other words, in the context of an integrated survey system where the sum of the "Non-Directory" (area survey) and the "Directory" establishments represents the entire universe, any nonagricultural economic establishment defined as out-of-scope of the "Directory" will automatically fall into the "Non-Directory" area survey. Therefore the **entire universe will be represented no matter how the Directory is defined.** However, it is very important to use a definition that includes all of the large establishments, is unambiguous, is available in administrative records, and can be easily identified in the field. *(Further discussion of this topic is beyond the scope of this size-limited paper.)*

B. Survey of "Non-Directory" Establishments -- Survei Usaha Terintegrasi (SUSI)1. Sample Design: Major Considerations

In designing the sample for the area survey of the "Non-Directory" establishments, Statistics Indonesia needed to take into consideration several important factors.

- Emphasis on employment size was critical. Industries and geographic areas with high employment levels needed to be given more prominence than those with lower levels.
- Variances for all tabulations at the 1-, 2- and 3-digit industry level for Indonesia needed to be calculated, as well as that for the 1- and 2- digit industry levels for each province. Given such a large number of strata, it was necessary to find a sampling program that would provide a solution for this complicated sampling process.
- Given the large number of sectors and geographic areas to be surveyed, the cost of such an undertaking was expected to be high. One ideally wanted an automated program which allowed cost comparison of various potential samples. The "best" sample, i.e. that which provides the most appropriate CVs at the least overall cost, would then be chosen.
- To make the most effective use of the budget and reduce resulting variances a sub-sampling strategy within selected enumeration areas was adopted.

2. CHROMY-GEN: Why and How it was Used?

The CHROMY-GEN sampling program was developed by the U.S. Bureau of the Census[3] to help survey designers find minimum cost stratified samples that satisfy multiple variance constraints. It was then adapted for use in economic surveys that use Poisson sampling (in particular, the U.S. Annual Survey of Manufactures) where each sampling unit is treated as a separate stratum.

The CHROMY_GEN program was so named because it implements the multivariate-allocation algorithm described by Chromy (1987)[4], and it is a general-purpose program. No coding changes are needed to apply it to different surveys; one simply prepares different input files. Zayatz and Sigman (1995)[5] document the use of CHROMY_GEN for designing surveys that use stratified sampling.

Chromy's algorithm can also be applied to designing surveys that use Poisson sampling by treating each sampling unit as a separate stratum (Sigman and Monsour, 1995)[6]. For example, Zayatz and Sigman (1994)[7] and Bond, Struble, and Imel (1995)[8] describe the use of Chromy's algorithm to determine Poisson-sampling probabilities for the U.S. Annual Survey of Manufactures. Also, Saavedra and Weir (1999)[9] describe the use of Chromy's algorithm to determine the probabilities of selection of individual

---

[3] The description of CHROMY_GEN in this paper reports results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform parties and to encourage discussion.

[4] Chromy, J. (1987), "Design Optimization with Multiple Objectives," *Proceedings of the Survey Research Methods Section*, American Statistical Association, pp. 194-199.

[5] Zayatz, L. and Sigman, R. (1995), "CHROMY_GEN: General Purpose Program for Multivariate Allocation of Stratified Samples Using Chromy's Algorithm," ESMD Methods Report Series ESMD-9502, Washington DC: U.S. Census Bureau.

[6] Sigman, R. and Monsour, N. (1995), "On Selecting Samples from Frames Based on Lists of Establishments," in Cox, F.; Binder, D.; Chinnappa, B.; Christianson, A.; College, M.; and Kott, P. (eds), *Business Survey Methods*, New York: Wiley, pp. 133-152.

[7] Zayatz, L. and Sigman, R. (1994), "Feasibility Study of the Use of Chromy's Algorithm in Poisson-Sample Selection for the Annual Survey of Manufactures," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 641-646.

[8] Bond, D.; Struble, R.; and Imel, L. (1995), "Results of Using Chromy's Algorithm for the Annual Survey of Manufactures," *Proceedings of the Section on Survey Research Methods*, American Statistical Association

[9] Saavedra, P. and Weir, P. (1999), "Application of the Chromy Allocation Algorithm with Pareto Sampling," to appear in the *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

sampling units in a petroleum-products survey. When Chromy's algorithm is used in this manner, to determine the probabilities of selections of individual sampling units, it

- Minimizes:

  C = expected total variable survey cost

  $= \sum \pi_i c_i$

  where π_i is the probability of selecting sampling unit i, c_i is the variable survey cost associated with sampling unit i, and summation is over all sampling units in the survey population,
- Subject to multiple variance constraints (for the Horvitz-Thompson estimator):

  $$\sum_i y_{ij}^2 \left(\pi_i^{-1} - 1\right) \leq V_j^*, \; j = 1, 2, \ldots, J$$

  where y_{ij} is the data from sampling unit i for survey item j, V_j^* is the target variance for survey item j, J is the number of survey items, and summation is over all sampling units in the population..

Chromy's algorithm solves iteratively for the minimum-cost π_i values. Sigman and Monsour (1995) describe how a result by Causey (1983)[10] provides a stopping rule for Chromy's algorithm[11].

Statistics Indonesia used CHROMY-GEN to design an annual integrated survey of small establishments and household activities ("Survei Usaha Terintegrasi", SUSI). It selected enumeration areas based on their contribution to all national and provincial sectoral totals at once. This reduced the total budget since the total number of selected enumeration areas was smaller than if a separate sample of enumeration areas was selected for each industrial sector. It was also possible to set expected relative standard errors for individual industries at any desired level of detail. The output of the CHROMY program is a set of probabilities for each enumeration area. Using a random number start the enumeration areas were then selected systematically with probability proportional to size.

Additionally, use of a Poisson sampling system allows easy calculation of variances for all publication/tabulation levels. This allows the publication of the maximum data possible. Every tabulation cell that meets the coefficient of variation (CV) publication criteria can be published. This is true for any level of industry/geographic area combination.

C. Integrating "Directory" and "Non-Directory" Establishments

Although the Directory portion of this overall plan has not yet been implemented, it will be possible to combine both parts of the universe to achieve an overall total for Indonesia and its Provinces. Both tabulated totals and variances can be combined, tabulated, and published.

IV. Sample Results

Although the 90.000 sample for 1998 is very small (only one-half of one percent of total units) it is very efficient. The strategy for the sample selection was to provide reliable estimates for the most important industries, those with the largest number of workers. It does this very well. As shown in Table 2 below, at 20 percent relative standard error publication criteria, 98.0%, 94.8%, and 89.5% of total employment for Indonesia can be published by 2-. 3-. and 4-digit industry respectively. By Province, 87.5%, 72.2%, 55.1%, and 43.5% of total employment can be published for Provinces by 1-, 2-, 3-, and 4-digit industry respectively.

---

[10] Causey, B. (1983), "Computational Aspects of Optimal Allocation in Multivariate Sampling," *SIAM Journal of Scientific and Statistical Computing*, Vol. 4, pp. 322-329.

[11] The CHROMY_GEN program is written in SAS® and is freely available from the U.S. Census Bureau (rsigman@census.gov). Users must have access to a licensed version of SAS, however, to run CHROMY_GEN.

Table 4: Result of the SUSI Survey by Industry Sector and Province: 1996, 1998, and 1999

	Number of Establishments (000)					Number of Workers (000)				
	1996	1998	1999	% change		1996	1998	1999	% change	
	Census	SUSI	SUSI	98/96	99/98	Census	SUSI	SUSI	98/96	99/98
Indonesia, Total	16,780.6	14,074.3	14,462.1	-16.1	2.8	28,876.4	26,020.2	26,528.0	-9.9	2.0
By Industry Sector:										
Mining, quarrying	197.2	142.9	153.4	-27.5	7.3	369.0	279.5	290.3	-24.3	3.9
Manufacturing	2,859.3	2,189.9	2,438.8	-23.4	11.4	6,598.1	5,287.4	5,941.4	-19.9	12.4
Electricity, gas, water	8.9	4.9	3.8	-44.9	-22.4	14.3	9.5	5.5	-33.6	-42.1
Construction	151.3	93.7	90.1	-38.1	-3.8	598.1	366.3	291.5	-38.8	-20.4
Whole/retail trade, restaurants, hotels, etc	9,772.4	8,618.5	8,719.4	-11.8	1.2	15,821.8	15,626.3	15,541.2	-1.2	-0.5
Transport, storage, and communication	1,833.7	1,600.8	1,667.4	-12.7	4.2	2,307.3	1,993.3	2,088.4	-13.6	4.8
Financial institutions	16.6	4.8	3.1	-71.1	-35.4	43.9	16.8	11.9	-61.7	-29.2
Real estate, services	1,941.2	1,418.8	1,386.1	-26.9	-2.3	3,123.9	2,441.1	2,357.8	-21.9	-3.4
By Province:										
Dista Aceh	352.6	301.9	202.8	-14.4	-32.8	606.2	566.4	372.6	-6.6	-34.2
North Sumatera	662.2	598.0	589.5	-9.7	-1.4	1,184.1	1,145.1	1,124.9	-3.3	-1.8
West Sumatera	365.3	306.7	319.3	-16.0	4.1	660.2	660.3	639.0	0.0	-3.2
Riau	227.7	174.6	171.2	-23.3	-1.9	432.2	378.1	346.9	-12.5	-8.3
Jambi	143.1	117.4	130.1	-18.0	10.8	270.2	229.0	248.5	-15.2	8.5
South Sumatera	447.5	359.3	352.6	-19.7	-1.9	807.0	724.8	704.2	-10.2	-2.8
Bengkulu	90.0	69.4	68.3	-22.9	-1.6	171.6	137.6	128.2	-19.8	-6.8
Lampung	456.0	357.7	335.2	-21.6	-6.3	789.7	671.1	657.7	-15.0	-2.0
Jakarta	852.1	711.0	647.3	-16.6	-9.0	1,575.6	1,476.0	1,298.7	-6.3	-12.0
West Java	3,472.5	2,901.9	2,992.3	-16.4	3.1	5,620.0	5,114.1	5,191.2	-9.0	1.5
Central Java	2,899.3	2,515.0	2,798.6	-13.3	11.3	5,002.9	4,520.1	4,980.7	-9.7	10.2
Yogyakarta	364.4	304.6	327.3	-16.4	7.5	643.2	556.7	613.1	-13.4	10.1
East Java	3,470.7	2,879.4	3,043.8	-17.0	5.7	5,804.2	5,106.7	5,419.1	-12.0	6.1
Bali	307.9	273.8	255.2	-11.1	-6.8	553.1	532.9	509.0	-3.7	-4.5
West Nusa Tenggara	352.7	287.0	289.1	-18.6	0.7	616.5	514.5	498.3	-16.5	-3.1
East Nusa Tenggara	167.0	146.9	182.9	-12.0	24.5	314.9	304.1	359.7	-3.4	18.3
West Kalimantan	204.6	160.9	183.2	-21.4	13.9	398.2	351.1	416.3	-11.8	18.6
Central Kalimantan	139.7	126.7	122.9	-9.3	-3.0	286.4	263.7	252.1	-7.9	-4.4
South Kalimantan	337.1	271.0	275.9	-19.6	1.8	550.6	465.6	516.4	-15.4	10.9
East Kalimantan	156.7	142.1	153.4	-9.3	8.0	278.3	276.2	338.0	-0.8	22.4
North Sulawesi	268.1	224.1	217.6	-16.4	-2.9	487.3	431.1	377.6	-11.5	-12.4
Central Sulawesi	148.7	123.2	134.9	-17.1	9.5	267.7	232.2	252.3	-13.3	8.7
South Sulawesi	601.7	472.5	482.3	-21.5	2.1	979.4	852.2	934.1	-13.0	9.6
Sulawesi Tenggara	110.4	92.6	101.6	-16.1	9.7	215.2	181.6	180.1	-15.6	-0.8
Maluku	120.8	100.0	33.8	-17.2	-66.2	235.8	206.3	65.3	-12.5	-68.3
Irian Jaya	61.8	56.6	51.0	-8.4	-9.9	125.9	122.7	104.0	-2.5	-15.2

Attachment B

1998 SUSI: Proposed Data Table for Indonesia and Provinces

01/08/2000

	Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Relative Standard Error Pekerja
Indonesia	All	14.099.025	26.070.199	1,7	1,7
	2	142.898	279.462	8,7	9,3
	24	111.868	183.155	9,8	9,7
	240	111.868	183.155	9,8	9,7
	2401	44.709	78.075	14,6	16,4
	2408	57.313	86.471	13,7	12,3
	3	2.196.899	5.301.654	2,4	2,3
	31	772.192	1.889.892	3,5	3,5
	311	421.830	1.013.904	4,7	4,7
	3115	10.731	23.180	19,6	19,1
	3116	132.511	349.227	5,6	6,0
	3117	36.966	115.374	7,8	9,6
	3118	219.088	469.181	7,7	8,4
	312	324.865	803.077	4,8	4,9
	3123	28.398	58.664	8,7	9,1
	3124	119.841	307.188	6,8	7,6
	3125	92.386	248.040	10,8	9,4
	3127	72.265	153.246	10,0	10,0
	32	289.468	658.798	6,3	6,4
	321	177.315	295.348	8,2	8,3
	3211	144.724	238.707	9,3	9,6
	322	84.704	252.759	10,1	10,3
	3221	58.042	188.545	11,1	9,4
	33	783.787	1.645.289	3,9	4,0
	331	619.053	1.127.824	4,6	4,5
	3311	50.863	149.336	7,9	8,8
	3313	513.104	860.534	5,3	5,4
	3314	34.552	72.692	17,5	17,2
	3319	18.026	35.654	18,1	16,4
	332	164.734	517.465	5,5	7,4
	3321	136.321	446.645	5,9	8,1
	34	13.754	47.099	12,4	14,7
	342	10.733	33.745	13,5	15,5
	3420	10.733	33.745	13,5	15,5
	36	208.097	704.680	5,8	5,9
	363	15.751	71.469	12,8	15,6
	364	182.120	609.368	6,4	6,4
	3642	162.271	565.061	6,4	6,6
	38	60.519	192.234	10,4	10,2
	381	52.131	168.310	11,7	11,2
	3811	27.785	80.257	13,9	16,1

			Relative Standard Error	
Kkkp	Establishments	Pekerja	Establishments	Pekerja
39	51.386	114.636	10,0	11,1
390	51.386	114.636	10,0	11,1
3901	9.828	17.676	17,8	17,4
3909	19.142	45.785	16,3	19,5
5	93.870	366.892	5,5	6,0
52	82.017	324.729	5,9	6,3
521	73.590	295.339	6,3	6,8
5211	48.244	196.716	8,0	8,4
5212	25.057	97.513	9,4	11,1
522	8.427	29.390	14,6	15,5
54	9.001	32.062	15,2	18,3
542	7.269	22.627	17,6	20,0
5420	7.269	22.627	17,6	20,0
6	8.634.171	15.656.942	1,9	1,9
61	253.460	644.101	5,9	7,1
613	148.978	396.904	7,9	9,4
6131	122.943	315.854	8,7	10,6
6133	24.208	76.533	14,3	19,6
614	78.779	198.153	10,8	12,5
6141	61.321	142.231	12,9	15,4
62	6.603.335	11.737.796	2,0	2,0
621	1.097.185	2.103.138	3,1	3,3
6212	1.097.185	2.103.138	3,1	3,3
622	112.324	227.059	8,8	9,6
6222	112.324	227.059	8,8	9,6
623	2.514.955	4.686.722	2,6	2,7
6231	641.958	1.098.136	4,6	5,4
6232	1.872.997	3.588.586	2,6	2,8
624	805.666	1.913.960	4,0	4,4
6241	315.226	694.822	6,4	6,9
6242	102.153	252.489	7,9	9,1
6243	59.487	142.594	11,7	11,0
6244	103.793	312.773	7,7	9,5
6245	80.095	144.224	7,8	8,1
6246	34.571	94.875	11,8	14,5
6247	47.444	151.050	11,8	16,6
6248	32.101	55.079	11,2	13,0
6249	30.796	66.054	12,9	14,2
625	86.187	204.071	8,8	10,0
6253	16.233	41.269	14,2	16,7
6257	34.774	82.535	15,1	17,7
626	1.987.018	2.602.846	2,9	3,0
6261	1.237.498	1.627.393	3,4	3,7
6269	749.520	975.453	4,3	4,5

			Relative Standard Error	
Kkkp	Establishments	Pekerja	Establishments	Pekerja
63	1.773.710	3.268.775	2,8	3,0
631	1.773.710	3.268.775	2,8	3,0
6312	246.827	673.021	5,5	6,5
6314	716.484	1.448.861	3,4	3,7
6315	803.965	1.127.972	4,3	4,7
7	1.601.549	1.995.399	2,8	2,8
71	1.555.331	1.878.287	2,8	2,8
712	1.550.329	1.870.027	2,8	2,8
7122	656.392	671.252	4,0	4,0
7125	761.089	947.090	3,5	3,5
7126	131.820	250.473	5,5	5,9
72	33.977	74.204	11,3	19,0
722	29.227	49.337	11,7	12,8
7221	21.872	33.129	13,7	13,8
8	323.912	483.878	5,1	5,3
82	319.126	467.107	5,2	5,4
821	262.546	333.598	5,9	6,3
8211	118.571	138.394	6,2	6,3
8212	143.975	195.204	9,1	9,5
822	43.831	95.513	7,9	9,6
8221	16.872	26.361	11,6	14,7
8223	16.737	49.425	10,7	13,5
829	10.738	33.650	12,9	13,8
8299	7.846	24.426	14,5	16,1
9	1.100.788	1.976.487	2,3	3,0
92	17.716	87.061	13,3	15,6
922	17.716	87.061	13,3	15,6
93	109.477	153.956	5,0	5,9
932	109.477	153.956	5,0	5,9
9322	9.594	15.776	16,8	18,6
9323	28.027	47.907	8,7	10,1
9326	64.200	75.498	6,7	7,9
962	5.861	20.071	16,8	19,3
97	960.021	1.671.880	2,4	3,0
971	304.021	576.069	3,3	4,0
9711	117.874	298.710	4,9	5,8
9712	50.619	64.682	7,3	7,5
9713	28.129	35.466	8,9	9,5
9714	69.739	105.427	6,3	7,9
9719	37.660	71.784	8,2	9,7
972	19.141	37.276	11,5	18,2
9720	19.141	37.276	11,5	18,2
974	120.876	202.743	5,0	5,7
9741	50.061	64.994	7,4	7,7

			Relative Standard Error	
Kkkp	Establishments	Pekerja	Establishments	Pekerja
9742	70.815	137.749	6,4	7,5
979	512.116	845.799	3,0	4,3
9791	19.329	40.390	11,9	12,9
9792	302.341	487.603	3,5	4,1
9799	188.958	315.771	4,5	9,0

				Relative Standard Error	
	Kkkp	Establishments	Pekerja	Establishments	Pekerja
	626	60.362	82.640	7,1	7,7
	6261	46.555	63.319	7,9	8,4
	6269	13.807	19.321	13,0	14,0
	63	94.045	186.125	5,7	6,0
	631	94.045	186.125	5,7	6,0
	6312	6.231	15.543	18,2	19,9
	6314	63.633	130.802	6,3	6,6
	6315	21.389	34.046	9,9	11,0
	7	60.746	67.870	7,0	6,8
	71	60.270	66.871	7,1	6,9
	712	60.219	66.820	7,1	6,9
	7122	22.389	22.512	10,4	10,4
	7125	35.671	39.868	8,0	7,9
	7126	2.159	4.440	17,6	19,8
	8	8.985	13.449	13,3	12,7
	82	8.822	13.055	13,5	13,0
	821	7.391	9.002	15,5	16,3
	8211	6.688	7.724	16,8	17,9
	9	65.163	124.829	5,3	6,5
	93	7.408	12.416	10,5	12,2
	932	7.408	12.416	10,5	12,2
	9323	3.035	5.563	13,9	14,6
	97	55.584	102.631	5,4	6,2
	971	22.674	48.028	7,0	8,3
	9711	10.500	29.833	9,2	10,9
	9712	4.499	5.618	15,3	17,5
	9714	3.220	4.827	16,0	17,9
	974	7.038	11.266	11,7	17,2
	9741	3.282	3.883	15,7	15,9
	979	23.068	39.554	7,1	8,0
	9792	16.986	27.760	7,9	9,1
	9799	5.558	10.506	12,1	14,7
13 Sumatra Barat	All	306.736	660.289	8,9	9,2
	3	47.458	133.422	12,6	16,7
	31	16.924	48.274	17,6	19,3
	6	202.695	418.297	9,5	10,6
	62	155.189	315.157	10,1	11,5
	623	74.533	154.223	12,5	13,2
	6232	57.753	120.423	11,8	12,3
	626	37.466	58.046	14,3	15,8
	6261	26.722	41.671	16,1	18,2
	63	35.813	76.772	14,7	16,6
	631	35.813	76.772	14,7	16,6
	6314	24.918	51.927	16,2	18,0

	Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Relative Standard Error Pekerja
	7	18.344	28.897	13,1	14,0
	71	17.572	27.638	13,2	14,3
	712	17.572	27.638	13,2	14,3
	7125	10.854	17.206	15,7	16,5
	9	23.492	39.839	12,1	14,4
	97	20.240	32.633	12,6	14,4
	971	7.373	12.942	15,9	16,4
	979	9.887	15.719	16,0	18,9
	9792	7.008	10.154	17,6	18,4
14 Riau	All	174.648	378.135	10,9	10,6
	3	14.928	38.892	18,9	18,3
	6	111.641	266.031	10,9	12,2
	62	85.704	203.533	11,3	13,4
	621	14.038	33.685	17,2	18,7
	6212	14.038	33.685	17,2	18,7
	623	33.854	76.016	16,4	17,8
	6232	30.026	66.711	17,8	19,5
	63	24.229	58.794	15,4	16,0
	631	24.229	58.794	15,4	16,0
	6314	16.241	38.429	17,8	18,2
	9	14.725	29.406	14,0	17,0
	97	12.243	22.010	14,5	15,7
	979	6.018	9.749	16,8	18,5
15 Jambi	All	117.367	229.019	11,8	11,9
	6	72.513	138.125	12,2	12,1
	62	62.248	118.426	13,1	12,8
	623	34.490	67.605	14,9	14,9
	6232	27.177	55.662	15,8	16,0
	626	8.704	10.888	19,9	20,0
	7	14.351	19.105	15,3	15,1
	71	12.311	15.925	16,3	15,9
	712	12.311	15.925	16,3	15,9
	7125	9.875	11.838	17,2	16,9
	9	11.371	20.423	14,4	16,9
	97	9.185	17.509	15,7	18,6
	979	4.433	7.734	18,4	19,1
16 Sumatra Selatan	All	359.273	724.672	9,2	9,4
	3	36.501	98.809	13,1	14,0
	6	239.205	497.679	9,5	9,9
	62	189.738	388.921	9,8	10,3
	623	92.941	204.938	10,7	10,9
	6232	79.839	179.058	11,0	11,4
	626	49.764	80.223	15,1	17,8
	63	41.015	90.067	16,4	15,9

		Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Relative Standard Error Pekerja
		631	41.015	90.067	16,4	15,9
		7	39.803	50.152	15,0	14,7
		71	36.621	45.691	15,5	15,3
		712	36.621	45.691	15,5	15,3
		7125	13.313	16.175	16,8	16,2
		9	27.065	42.491	13,2	14,7
		97	21.816	35.982	14,2	16,1
17	Bengkulu	All	69.402	137.628	10,6	11,2
		3	6.813	16.594	14,3	16,1
		31	2.652	6.001	17,7	19,1
		6	47.602	92.925	10,9	11,3
		62	37.247	72.175	11,4	12,0
		621	17.617	38.233	12,9	14,8
		6212	17.617	38.233	12,9	14,8
		63	7.275	15.816	17,1	18,9
		631	7.275	15.816	17,1	18,9
		7	6.377	8.598	16,6	17,8
		71	6.377	8.598	16,6	17,8
		712	6.377	8.598	16,6	17,8
		7125	4.907	5.760	17,2	17,7
18	Lampung	All	357.699	671.072	9,1	9,0
		3	55.737	152.413	11,5	11,6
		31	24.627	59.927	15,1	15,3
		311	18.639	44.722	18,0	18,2
		36	14.457	57.493	19,3	19,6
		6	230.469	410.048	9,6	9,8
		62	193.409	331.482	10,1	10,2
		621	44.834	87.199	14,0	14,7
		6212	44.834	87.199	14,0	14,7
		623	77.532	132.128	13,4	13,3
		6232	55.347	96.330	13,8	14,3
		626	50.825	66.900	13,6	13,7
		6261	25.407	35.244	16,6	17,1
		6269	25.418	31.656	17,9	18,6
		63	33.166	64.992	15,9	19,1
		631	33.166	64.992	15,9	19,1
		7	34.320	45.986	17,2	15,4
		71	33.664	44.730	17,5	15,7
		712	33.664	44.730	17,5	15,7
		9	30.545	44.123	13,0	14,0
		97	24.755	37.155	13,9	15,2
		979	10.867	17.383	17,7	18,2
31	D.K.I. Jakarta	All	711.036	1.476.220	5,6	6,3
		3	20.314	105.878	11,0	13,5

		Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Relative Standard Error Pekerja
		32	9.036	53.179	14,2	15,2
		322	7.650	45.091	15,4	15,5
		3221	7.058	43.275	15,6	15,9
		6	515.293	1.063.176	5,9	6,7
		62	334.030	710.769	6,4	7,3
		621	55.055	109.228	11,6	12,1
		6212	55.055	109.228	11,6	12,1
		623	128.981	260.342	8,2	9,2
		6231	40.789	72.864	14,3	16,4
		6232	88.192	187.478	8,9	10,3
		624	78.295	227.882	11,8	13,7
		6241	38.184	110.641	18,8	18,9
		626	50.240	57.956	10,8	11,4
		6261	28.199	33.876	14,7	15,7
		6269	22.041	24.080	14,2	14,2
		63	173.014	324.614	7,8	9,2
		631	173.014	324.614	7,8	9,2
		6312	29.399	105.965	11,0	18,9
		6314	45.539	99.822	12,7	12,9
		6315	97.719	116.621	9,8	10,9
		7	56.791	67.033	9,3	10,1
		71	55.429	58.774	9,4	9,3
		712	55.429	58.774	9,4	9,3
		7125	46.018	48.748	10,0	10,0
		8	34.299	42.831	7,8	8,0
		82	34.190	42.490	7,8	8,0
		821	30.831	35.020	8,2	8,5
		8211	26.939	30.276	9,0	9,3
		8212	3.892	4.744	18,9	18,7
		822	1.999	3.154	19,0	17,0
		9	82.986	192.542	6,7	7,7
		97	73.243	162.719	6,9	7,3
		971	21.711	54.835	9,3	9,9
		9711	10.538	32.537	11,6	12,4
		9714	5.429	10.049	14,7	17,0
		974	14.427	29.118	11,7	13,0
		9741	4.170	7.168	16,4	18,6
		9742	10.257	21.950	14,1	15,7
		979	32.613	68.466	9,2	10,4
		9792	19.457	44.000	10,0	12,9
		9799	10.528	18.314	16,6	16,9
32	Jawa Barat	All	2.901.883	5.113.960	5,1	5,2
		3	307.891	833.773	7,1	6,8
		31	126.841	307.226	11,3	10,3

			Relative Standard Error	
Kkkp	Establishments	Pekerja	Establishments	Pekerja
311	79.182	177.902	13,9	13,0
3116	21.198	51.292	18,6	18,5
3118	48.082	100.453	19,8	19,6
312	44.100	120.735	16,4	14,9
33	100.130	236.954	12,0	11,3
331	74.236	145.311	14,9	13,9
3313	65.118	118.489	16,6	15,7
332	25.894	91.643	15,8	17,2
3321	18.849	71.464	17,2	18,9
36	29.545	111.061	14,2	16,7
364	27.654	105.938	14,6	17,3
3642	26.254	100.203	14,9	17,9
6	1.810.591	3.213.058	5,6	6,0
61	48.683	112.921	19,0	19,5
62	1.364.675	2.425.330	5,8	6,2
621	222.099	424.491	9,4	9,8
6212	222.099	424.491	9,4	9,8
623	478.690	984.390	7,7	8,0
6231	91.746	188.116	15,3	16,8
6232	386.944	796.274	8,1	8,3
624	138.131	328.395	11,8	14,0
626	483.482	606.041	8,1	8,7
6261	275.299	344.396	10,0	11,3
6269	208.183	261.645	11,1	11,6
63	397.233	674.807	8,5	9,4
631	397.233	674.807	8,5	9,4
6314	104.703	213.209	12,5	13,1
6315	249.991	340.885	10,2	11,8
7	507.521	589.702	6,3	6,4
71	504.202	579.926	6,3	6,5
712	504.012	579.736	6,3	6,5
7122	155.901	157.696	10,3	10,3
7125	311.414	357.266	7,0	7,1
7126	35.886	63.963	14,3	15,6
8	78.663	110.858	12,5	13,2
82	77.751	106.455	12,6	13,4
821	64.381	76.239	14,0	15,1
8211	23.764	25.170	19,2	19,0
9	179.018	317.636	7,6	8,8
97	159.792	280.716	8,0	9,2
971	53.244	98.836	10,8	13,2
9711	22.251	52.457	16,8	19,8
974	25.138	41.796	15,7	17,7
979	80.065	133.240	10,1	11,7

		Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Relative Standard Error Pekerja
		9792	54.136	93.135	10,8	13,4
		9799	22.230	33.002	19,0	19,5
33	Jawa Tengah	All	2.515.013	4.519.803	4,1	4,2
		2	40.793	56.741	17,8	16,9
		24	40.168	55.584	18,0	17,2
		240	40.168	55.584	18,0	17,2
		3	553.102	1.357.447	5,2	5,2
		31	225.841	549.985	6,8	7,3
		311	123.469	290.227	9,0	10,3
		3116	16.141	42.679	14,1	14,2
		3118	97.973	218.116	10,8	12,9
		312	100.058	238.669	9,3	8,6
		3123	8.641	16.575	16,9	17,0
		3124	43.025	112.144	11,4	12,1
		3125	32.438	76.299	20,0	16,0
		32	49.184	113.798	15,8	14,4
		33	202.801	449.221	8,9	9,4
		331	159.885	277.034	10,7	10,7
		3313	147.218	241.130	11,4	12,0
		332	42.916	172.187	12,1	16,3
		3321	36.699	160.748	13,1	17,3
		36	53.202	171.165	11,6	11,9
		364	46.356	146.948	12,7	13,1
		3642	40.301	135.644	12,9	13,6
		6	1.477.930	2.449.684	4,6	4,8
		61	48.277	130.613	13,8	17,6
		613	34.101	91.937	17,3	18,5
		62	1.127.134	1.799.186	4,9	5,1
		621	121.481	214.330	8,3	8,9
		6212	121.481	214.330	8,3	8,9
		623	413.267	699.680	7,0	7,1
		6231	127.052	212.739	10,8	13,2
		6232	286.215	486.941	7,1	7,1
		624	118.936	260.692	11,6	11,4
		6242	16.398	31.379	18,4	20,0
		6244	17.512	54.534	17,1	18,0
		625	14.640	31.741	19,1	19,5
		626	442.816	560.736	5,7	6,2
		6261	296.387	383.959	6,6	7,3
		6269	146.429	176.777	8,6	9,1
		63	301.986	519.352	6,5	6,7
		631	301.986	519.352	6,5	6,7
		6312	49.958	119.455	12,3	13,1
		6314	112.111	207.625	7,8	8,2

		Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Pekerja
		6315	138.524	188.241	9,5	10,3
		7	220.203	277.730	6,4	6,3
		71	217.858	274.481	6,4	6,4
		712	216.228	272.294	6,5	6,4
		7122	127.390	127.599	8,4	8,4
		7125	69.061	105.755	8,6	9,0
		7126	19.777	38.940	12,5	13,1
		822	9.487	17.363	17,9	16,7
		9	181.963	299.724	5,7	7,8
		93	18.454	22.684	11,9	13,3
		932	18.454	22.684	11,9	13,3
		9326	12.230	12.698	14,2	14,1
		97	156.556	254.074	6,0	8,2
		971	42.879	80.494	8,2	11,9
		9712	11.401	14.741	15,6	16,2
		9714	9.221	13.830	14,4	19,8
		974	17.865	25.791	11,6	13,3
		9741	8.011	9.487	16,8	17,4
		9742	9.854	16.304	14,9	17,5
		979	93.669	141.385	7,1	11,2
		9792	56.850	80.492	8,4	8,4
34	Jogjakarta	All	304.583	556.683	7,8	8,1
		3	74.790	162.366	9,9	9,3
		31	24.783	56.494	12,2	12,4
		312	19.876	41.399	13,0	12,2
		3124	11.554	26.364	14,2	14,9
		33	26.465	44.884	14,7	15,1
		331	18.774	29.085	17,8	18,7
		6	161.508	284.776	9,0	9,0
		62	117.027	190.565	9,5	9,4
		621	15.362	29.526	14,4	16,0
		6212	15.362	29.526	14,4	16,0
		623	39.537	65.391	14,6	13,7
		6232	23.993	40.690	13,6	14,1
		624	17.185	35.660	16,5	16,9
		626	40.825	53.003	12,0	12,8
		6261	32.092	40.247	12,9	13,4
		63	36.557	75.246	13,2	13,9
		631	36.557	75.246	13,2	13,9
		6314	13.714	29.844	16,8	18,6
		7	16.698	24.455	14,1	12,7
		71	16.432	23.961	14,2	12,8
		712	15.293	21.907	14,3	13,1
		7125	4.514	8.112	16,3	17,4

					Relative Standard Error	
		Kkkp	Establishments	Pekerja	Establishments	Pekerja
		8	17.449	27.620	16,7	17,3
		82	17.183	27.007	16,9	17,5
		821	15.445	21.137	17,9	18,9
		9	24.850	40.699	9,3	10,2
		97	21.490	35.013	9,8	10,7
		971	7.656	12.406	13,2	14,6
		979	11.226	18.426	11,6	13,5
		9792	7.493	12.630	13,7	15,7
35	Jawa Timur	All	2.879.439	5.106.603	3,8	4,0
		3	445.562	1.002.750	5,7	5,4
		31	127.142	323.640	9,4	9,2
		311	56.767	139.156	14,5	13,2
		3116	13.073	37.429	13,2	15,1
		312	65.775	178.468	11,6	12,5
		3124	37.502	87.057	14,8	17,2
		32	53.207	105.820	17,7	15,4
		33	184.800	340.049	8,4	8,1
		331	151.558	252.838	9,8	9,7
		3313	139.157	224.812	10,4	10,5
		332	33.242	87.211	12,0	13,2
		3321	25.917	72.574	13,0	14,2
		36	51.239	157.591	13,4	13,0
		364	43.728	121.159	15,1	15,2
		3642	40.464	116.060	15,7	15,6
		5	8.731	40.573	17,5	17,8
		52	6.374	33.545	18,4	19,6
		6	1.699.838	2.965.977	4,2	4,4
		61	41.751	121.612	12,5	17,6
		62	1.266.616	2.164.611	4,5	4,8
		621	150.361	270.598	8,8	9,1
		6212	150.361	270.598	8,8	9,1
		623	526.586	915.504	5,8	6,2
		6231	147.608	236.784	10,7	13,5
		6232	378.978	678.720	5,6	5,9
		624	150.234	349.496	9,3	10,3
		6241	51.242	115.804	14,7	18,1
		6244	20.433	52.598	15,8	18,0
		6245	16.511	28.212	20,0	19,8
		626	404.947	556.541	6,0	6,4
		6261	251.967	347.212	7,1	7,7
		6269	152.980	209.329	8,8	9,6
		63	390.234	678.430	5,6	5,8
		631	390.234	678.430	5,6	5,8
		6312	64.053	141.891	11,0	11,6

				Relative Standard Error	
	Kkkp	Establishments	Pekerja	Establishments	Pekerja
	6314	176.035	316.990	7,0	7,3
	6315	149.115	214.220	8,3	8,7
	7	348.184	431.738	5,5	5,6
	71	341.240	403.445	5,6	5,6
	712	339.248	399.667	5,6	5,6
	7122	201.673	203.563	7,2	7,2
	7125	109.208	143.307	7,7	7,7
	7126	28.367	52.797	10,8	11,4
	8	76.013	121.479	11,5	11,8
	82	75.333	118.666	11,5	11,9
	821	64.950	89.027	12,6	12,8
	8212	49.306	70.215	14,5	14,5
	9	270.896	491.018	4,7	7,7
	93	28.828	34.089	10,6	11,3
	932	28.828	34.089	10,6	11,3
	9326	20.674	22.342	12,9	13,4
	97	233.821	401.687	5,0	7,5
	971	63.403	113.248	6,9	8,5
	9711	18.067	47.733	11,1	14,0
	9712	13.770	16.246	14,3	15,0
	9714	17.317	23.860	12,3	15,2
	974	22.376	41.017	12,0	14,4
	9741	7.053	8.845	16,4	18,7
	9742	15.323	32.172	15,7	17,6
	979	144.096	241.717	5,8	10,7
	9792	72.572	114.446	7,4	9,1
	9799	68.101	121.898	7,7	18,8
51 Bali	All	273.775	532.978	9,5	9,5
	3	66.770	146.394	13,3	13,1
	33	39.417	76.288	17,7	18,1
	331	35.933	67.799	18,9	19,3
	6	164.895	313.106	11,2	12,0
	62	132.849	233.731	12,2	11,7
	623	50.781	79.190	15,5	14,6
	6232	33.241	54.150	16,2	14,9
	626	29.979	44.180	16,1	17,9
	7	13.681	16.042	19,1	18,1
	71	13.520	15.441	19,2	18,6
	712	13.520	15.441	19,2	18,6
	9	16.888	34.148	13,0	14,1
	97	14.696	28.703	13,3	14,5
	971	5.904	12.286	15,8	17,7
	979	6.555	10.600	18,2	19,6
52 N.T.B.	All	286.951	514.538	9,8	10,1

				Relative Standard Error	
	Kkkp	Establishments	Pekerja	Establishments	Pekerja
	3	75.804	159.631	12,9	12,1
	33	33.963	60.589	17,2	16,6
	331	29.310	52.092	18,4	17,8
	6	151.045	239.523	10,8	11,0
	62	126.600	195.189	11,3	11,4
	623	45.550	73.729	15,3	15,5
	6232	36.245	61.911	15,0	15,8
	626	48.660	60.943	15,4	15,5
	6261	38.646	49.317	18,0	17,8
	63	19.633	34.817	19,4	19,4
	631	19.633	34.817	19,4	19,4
	7	31.946	50.850	13,4	17,5
	71	31.824	50.241	13,4	17,6
	712	31.824	50.241	13,4	17,6
	7122	20.867	28.252	16,4	18,2
	9	12.647	20.064	14,7	16,4
	97	11.435	17.612	15,4	17,2
53 N.T.T.	All	146.947	303.998	10,8	10,9
	3	69.443	143.943	13,6	13,5
	32	45.751	91.344	16,8	17,0
	321	45.602	91.195	16,8	17,0
	3211	44.283	89.518	17,1	17,2
	6	62.173	119.524	13,6	14,2
	62	58.403	109.581	13,5	13,8
	621	18.022	33.281	18,6	18,2
	6212	18.022	33.281	18,6	18,2
54 Timor Timur	All	24.731	50.589	12,7	13,1
	6	15.586	30.430	14,6	14,6
	62	14.394	26.689	15,1	15,1
	621	8.125	15.593	15,7	15,6
	6212	8.125	15.593	15,7	15,6
61 Kalimantan Barat	All	160.871	351.055	12,4	12,1
	3	26.759	61.716	18,5	18,5
	6	102.956	227.652	12,8	13,6
	62	85.263	180.068	12,8	13,3
	621	28.958	64.439	17,7	17,9
	6212	28.958	64.439	17,7	17,9
	623	30.324	64.058	19,2	19,4
	6232	20.691	44.842	18,5	18,6
62 Kalimantan Tengah	All	126.696	263.607	11,2	11,4
	3	11.296	28.437	16,8	17,0
	6	77.770	152.905	12,4	12,6
	62	61.390	118.099	13,6	13,8
	621	12.739	27.949	18,6	18,8

				Relative Standard Error	
	Kkkp	Establishments	Pekerja	Establishments	Pekerja
	6212	12.739	27.949	18,6	18,8
	623	21.718	43.721	17,4	17,9
	6232	16.834	34.964	16,9	17,1
	63	12.123	24.252	16,4	17,8
	631	12.123	24.252	16,4	17,8
	7	15.511	19.533	15,0	14,9
	71	11.696	13.065	17,5	17,5
	712	11.696	13.065	17,5	17,5
	7125	8.453	9.608	18,4	18,7
	9	8.599	13.461	15,6	17,2
	97	8.105	12.797	16,0	17,7
	979	3.576	5.042	19,0	19,5
63 Kalimantan Selatan	All	270.998	465.530	10,3	10,3
	3	49.743	97.245	13,1	12,6
	33	26.458	52.069	16,1	15,9
	331	24.343	46.957	16,9	16,5
	3313	18.871	35.241	18,9	18,7
	6	143.715	244.628	11,1	11,1
	62	107.471	176.657	11,6	12,2
	621	26.328	49.083	16,9	16,7
	6212	26.328	49.083	16,9	16,7
	623	23.419	38.315	17,4	19,2
	6232	16.527	27.384	16,3	18,0
	626	40.235	55.073	14,4	14,5
	6261	23.311	31.614	17,1	18,2
	6269	16.924	23.459	18,2	18,5
	63	34.094	61.870	15,1	15,2
	631	34.094	61.870	15,1	15,2
	6314	20.217	38.150	15,1	15,6
	7	36.748	43.583	15,0	17,0
	71	32.651	38.546	15,5	17,9
	712	32.651	38.546	15,5	17,9
	7125	20.636	21.956	17,9	18,5
	9	21.871	34.537	14,3	15,3
	97	20.284	31.493	14,5	15,8
64 Kalimantan Timur	All	142.082	276.173	10,8	10,8
	3	9.439	19.900	16,5	15,7
	31	3.248	7.211	18,9	19,3
	6	89.926	186.653	11,2	11,7
	62	68.796	137.387	11,8	12,3
	623	42.481	86.824	14,2	14,7
	6232	33.151	71.953	14,2	14,6
	63	19.245	43.986	17,8	18,8
	631	19.245	43.986	17,8	18,8

				Relative Standard Error	
	Kkkp	Establishments	Pekerja	Establishments	Pekerja
	7	16.092	21.312	14,9	15,7
	71	12.246	14.831	16,3	16,8
	712	12.246	14.831	16,3	16,8
	7125	10.175	11.152	16,1	16,1
	8	9.326	11.513	17,9	17,7
	82	9.265	11.391	18,0	17,9
	821	8.307	9.312	19,3	19,1
	8211	8.307	9.312	19,3	19,1
	9	14.093	25.364	12,7	15,0
	97	12.734	21.718	13,3	14,9
	971	4.945	9.580	17,5	20,0
71 Sulawesi Utara	All	224.082	431.078	10,4	10,8
	3	52.094	128.531	14,8	15,5
	31	32.661	81.848	19,8	19,5
	6	121.777	213.701	11,7	12,0
	62	107.338	184.081	11,9	11,9
	621	30.260	55.668	16,1	15,6
	6212	30.260	55.668	16,1	15,6
	623	34.262	63.562	17,2	17,7
	6232	24.382	47.493	17,7	18,8
	626	30.892	41.133	18,4	17,7
	7	22.924	29.115	15,0	15,7
	71	22.705	28.831	15,1	15,8
	712	22.705	28.831	15,1	15,8
	9	17.836	29.670	18,0	18,5
	97	16.728	27.034	18,1	18,9
72 Sulawesi Tengah	All	123.161	232.105	10,5	11,3
	3	19.195	43.335	12,3	13,1
	31	6.066	12.264	16,9	17,0
	311	4.888	9.750	19,5	19,7
	33	9.820	19.967	15,6	15,7
	332	3.800	7.865	18,5	20,0
	6	81.983	143.917	11,1	10,9
	62	68.297	119.184	11,5	11,2
	621	22.021	40.971	14,4	14,5
	6212	22.021	40.971	14,4	14,5
	623	30.976	55.085	16,0	15,8
	6232	21.229	41.638	15,4	16,5
	63	12.383	21.082	19,1	19,7
	631	12.383	21.082	19,1	19,7
	7	7.111	10.380	16,1	15,2
	71	6.749	9.655	16,7	15,9
	712	6.749	9.655	16,7	15,9
	7125	3.052	4.315	18,6	18,2

		Kkkp	Establishments	Pekerja	Relative Standard Error Establishments	Pekerja
		9	8.754	15.786	14,4	16,6
		97	7.939	14.064	14,3	16,0
		971	2.198	4.388	18,3	19,7
		979	4.518	7.738	16,6	19,4
73	Sulawesi Selatan	All	472.503	852.278	8,8	9,3
		3	83.120	171.870	11,0	11,9
		31	27.094	59.500	13,9	14,7
		311	23.589	52.146	14,6	15,5
		3116	16.505	37.440	15,0	16,8
		6	294.930	520.561	9,5	9,7
		62	249.255	420.246	9,9	10,1
		621	89.738	177.059	13,1	14,5
		6212	89.738	177.059	13,1	14,5
		623	63.151	108.478	14,3	15,0
		6232	50.166	88.196	15,0	15,7
		626	70.500	83.549	14,1	13,5
		6261	52.185	62.685	14,6	14,4
		7	60.315	86.250	14,9	18,2
		71	56.952	68.139	15,6	14,9
		712	56.952	68.139	15,6	14,9
		7122	36.179	36.489	18,7	18,6
		9	24.820	44.521	16,1	18,1
		97	23.229	38.101	16,4	18,9
74	Sulawesi Tenggara	All	92.625	181.629	11,5	11,5
		3	18.536	39.065	16,4	14,8
		6	58.676	112.904	12,5	13,1
		62	55.253	105.788	12,4	12,9
		621	15.902	32.415	15,8	16,6
		6212	15.902	32.415	15,8	16,6
		623	19.503	39.543	15,4	15,9
		6232	16.179	34.123	15,8	16,6
		71	4.482	5.945	19,7	19,2
		712	4.482	5.945	19,7	19,2
81	Maluku	All	100.016	206.323	12,6	13,6
		3	10.812	25.558	17,0	18,6
		6	71.433	144.608	13,6	13,9
		62	65.172	131.074	13,5	13,6
		621	23.899	56.150	15,0	15,9
		6212	23.899	56.150	15,0	15,9
82	Irian Jaya	All	56.575	122.585	13,0	13,3
		6	37.599	80.072	14,6	14,2
		62	30.635	61.276	15,4	14,8

PREPARING A CHROMY SAMPLE FOR KABUPATEN ESTIMATES

Report # 9

by
Robert Torene

August, 2000

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 89

II. SAMPLE CONSIDERATIONS 89
- **A. Sample Size** 89
- **B. Sub-Use of Existing Sample in Kabupaten Sampling and Estimation:** 89
- **C. Use of Supplementary Kabupaten Sample in Estimation for National/ Provincial Level:** 90

III. STEPS FOR PREPARING CHROMY SAMPLE FOR KABUPATEN ESTIMATES 90
- **A. Step 1. Convert 1996 L2 Data to New Kabupaten Definitions** 90
- **B. Step 2. Prepare Actual Enumeration Cost Data for Each Wilker** 91
- **C. Step 3. Prepare Required CHROMY Files Including FREEZE** 91
  - **1. FREEZE File:** 91
  - **2. ITEM File:** 91
  - **3. STRATA File:** 91
  - **4. POP File:** 91
- **D. Step 4. Run CHROMY - GEN** 92

IV. RESULTS 92

V. OBSERVATIONS ON THE CHROMY RUNS 92
- **A. Stopping Rules** 92
- **B. FREEZE File and Stopping Rules** 93

VI. CONCLUSIONS AND RECOMMENDATIONS 93
- **A. Recommendation 1.** 93
- **B. Recommendation 2.** 93
- **C. Recommendation 3.** 94

I. INTRODUCTION

The original SUSI sample was selected to provide data for the larger KKKPs (in terms of employment) at National and Provincial levels. The sample was not designed to provide reliable data for Kabupatens, although, obviously, all selected wilkers are already located in Kabupatens/Kotamadya.

Under the new system of greater autonomy and budget for Provinces and Kabupatens, it may be possible that some Provinces/Kabupatens will approach the BPS requesting additional data from the SUSI for planning purposes. Therefore, the BPS wanted to determine the best method to supplement the existing National/Provincial sample, if requested to do so by the Kabupatens, and to determine the approximate cost of any additional sample. The sample problem is further complicated due to the:

- budget cut resulting in only 4.000 wilkers for the year 2000 SUSI, the
- decision to use a panel of 50% of the first quarter sample in each succeeding quarter, and the
- creation of new Kabupatens (only 304 Kabupatens existed in the 1996 SE; there are now 341).

It was decided to design a CHROMY sample for Kabupaten totals and Kabupaten 1-digit KKKP totals. It should be noted that this is not a final solution; it is basically an exercise to determine the appropriate techniques to use to select additional sample for Kabupatens on request.

II. SAMPLE CONSIDERATIONS

A. Sample Size

Although the original sample of approximately 8.000 wilkers was not designed to provide Kabupaten estimates, the SUSI wilkers and their weights should represent the economic activity of establishments with no legal status for each Kabupaten. This exercise is intended to help determine an appropriate sample size for each Kabupaten and serve as a starting point for discussions with Kabupaten and Provincial administrative staff.

B. Sub-Use of Existing Sample in Kabupaten Sampling and Estimation:

Since the existing wilkers are already weighted to estimate for the Kabupaten totals, one unbiased solution is to create a certainty stratum composed of these existing wilkers and then select another sample from the remaining wilkers. In this way, the existing wilkers are used in estimating for the Kabupaten total and 1-digit estimates.

CHROMY has the facility of using a FREEZE file. In stratified sampling the FREEZE file is used to indicate the minimum number of PSUs to be selected from a particular stratum. In the Poisson sample design used for SUSI, the FREEZE file indicates whether the PSU should be included with certainty resulting in a weight of 1,0. The PSUs listed in the FREEZE file do not contribute to the variance (since they have a certainty weight of 1,0) but do contribute to the final Kabupaten estimate. The CHROMY program will then work around the FREEZE file and produce a minimum cost sample that includes the FREEZE file wilkers and adheres to the CV constraints supplied to the program.

C. Use of Supplementary Kabupaten Sample in Estimation for National/ Provincial Level:

Since only a relatively small number of Kabupatens are expected to request and pay for additional sample, it was decided that the supplementary sample could only be used for making estimates at the Kabupaten level. The characteristics of establishments in a particular Kabupaten do not necessarily represent the characteristics of establishments throughout the rest of Indonesia. It would, therefore, be a biased procedure to include any supplemental sample in the national and provincial estimation procedure.

If a Province decides to supplement the sample for all of its Kabupatens, it will be possible to change the estimation procedure to take advantage of the supplemental sample and improve the Province estimates. The National estimation procedure can also be changed to include the improved Province estimates in with the National totals. The improvements at the National level will be only marginal (unless many Provinces decide to supplement the sample) and this should be reviewed to determine if the complexity is worth the additional improvement in the resulting data and relative standard errors.

III. STEPS FOR PREPARING CHROMY SAMPLE FOR KABUPATEN ESTIMATES

A. Step 1. Convert 1996 L2 Data to New Kabupaten Definitions

There are now 341 Kabupatens compared to 304 at the time of the SE-96. The new and revised Kabupatens are now smaller than before. This makes the selection of a reliable sample that much more difficult. In addition, to be responsive to the new and revised Kabupatens, it is necessary to convert the SE-96 data to the new Kabupaten definitions. Pak Kadarmanto supplied the conversions necessary to recast the SE-96 data from the old definitions to the new. What, at first, seemed like a relatively simple job turned out to be somewhat more complicated due to the use of the NKS identification for each wilcah. Although unique to each Kabupaten in the SE-96,

the NKS does not represent the hierarchical nature of the geopraphic code system. The kecamatan and kelurahan codes cannot be directly identified from the NKS code.

B. Step 2. Prepare Actual Enumeration Cost Data for Each Wilker

In previous CHROMY samples it was assumed that enumeration costs for all wilkers were the same. In this sample selection, we assigned estimated enumeration costs based on (1) the actual listing cost plus (2) the actual enumeration cost per unit. The listing cost was calculated at Rp. 152.000 per wilker including supervisory cost. The enumeration cost was calculated at Rp. 9.500 per establishment including supervisory cost. The total estimated enumeration cost was arrived at by multiplying the enumeration cost per unit by the estimated number of enumerated establishments (the SE-96 data on number of establishments by KKKP divided by the SUSI sub-sampling rate).

Total estimated enumeration cost per wilker ranges from a low of Rp. 161.500 to a high of Rp. 940.500. For this exercise, the CHROMY-GEN program will determine probabilities for each wilker while minimizing the total enumeration cost.

It should be noted that this is not the total cost. BPS staff will have to provide cost estimates for processing, transportation, etc. to arrive at the final total cost. However, it is a useful first step in developing a total cost by wilker for CHROMY sampling for SUSI and for other surveys.

C. Step 3. Prepare Required CHROMY Files Including FREEZE

1. FREEZE File:

This is the first sample using the FREEZE function in the CHROMY-GEN program. There were 2.219 Wilkers already included in the SUSI 2000 National/Provincial sample. These Wilkers were designated as certainty wilkers and were given weights of 1,0.

2. ITEM File:

There were 3.054 items designated. Items 1-341 were the Kabupaten totals with the remainder being Kabupaten 1-digit KKKP.

The designated CVs were assigned on a sliding scale: larger cells were assigned smaller CVs and vice versa.

3. STRATA File:

This file was prepared in similar fashion to the original SUSI samples – a record for each wilker in the SE-96.

4. POP File:

The POP file is a cross-classification of the ITEMS and STRATA files.

D. Step 4. Run CHROMY - GEN

CHROMY was run five separate times. Before each additional run the CVs were adjusted to loosen up the constraints so that the first run is the most costly and the 5th run the least costly option.

Constraint loosening was done using the "Shadow Price" information in the INFO output file. The "Shadow Price" is the approximate cost reduction for an Item if the CV is increased by 10%. Where the "Shadow Price" was relatively large (generally over Rp. 200.000 the CV was increased by 10%; from Rp.100.000 to 200.000 the CV was increased by 20%.

The large majority of constraining CVs were for the Kabupaten totals since this was the level where we wanted better estimates with resulting lower CVs.

IV. RESULTS

The results of the CHROMY runs are shown in the Attachments. Attachment 1 shows summary results for all 341 Kabupaten totals for all five CHROMY runs. Attachment 2 shows more detailed data for Samples 1, 3, and 5. The column, "Extra Cost" is reported in thousands of Rp.

For this exercise, there was no attempt to select the "best sample" for any of the runs. The first random number drawn was used for the sample shown in the attachments.

It is a most difficult exercise to achieve reasonably reliable results for smaller areas. In Sample 1 (the largest sample) CVs for the largest Kabupatens, (200.000 pekerja and more), range from 6 to 8 percent. For the smallest Kabupatens (1.000 to 2.000 pekerja) generally have CVs of 50% and more. In fact, some of the Kabupatens are so small that only a census will achieve reliable data.

As expected, calculated CVs are generally identical to the target CVs in each sample, except for the 1-digit Kabupaten totals which generally had a lot of "slack", the calculated CV being lower than the target. The constraining levels are almost always the Kabupaten totals.

V. OBSERVATIONS ON THE CHROMY RUNS

A. Stopping Rules

These CHROMY runs reacted very differently from the previous National/Provincial CHROMY runs. One of the "Stopping Rules" used in the CHROMY-GEN program is where the "Estimate of cost minus minimum cost" (EC) falls below a certain value entered by the user. For the previous SUSI samples, the EC for each iteration would always be lower than the preceding

iteration allowing the stopping rule to take effect. With these samples the EC always got exponentially larger for each iteration; therefore, the EC stopping rule never took effect.

Another "Stopping Rule" is the "Number of Iterations" (NI). Since this was a test using a very fast, powerful computer and I wanted to see how CHROMY reacted, I set the NI at 500 for four of the sample runs. On one run we let the computer work overnight to see what result could be obtained. In the morning the computer had performed 364 iterations and the exponent for the EC was e^{364}. Since the "Cost" was only going up by fractions of a rupiah, the computer was halted. Further iterations would not be beneficial.

Although the EC kept getting exponentially larger with each iteration, the sample probabilities resulted in samples that met the target CVs in all cases.

Perhaps, we can revise the CHROMY-GEN program to change the stopping rule when the EC gets larger rather than smaller. This will be discussed with Richard Sigman when I return to Washington.

B. FREEZE File and Stopping Rules

This was the first time a FREEZE file was used. It is very important to note that the FREEZE file PSUs are only entered in the SIZE file when CHROMY stops in accordance with one of the stopping rules. It was very puzzling to me that four of the samples seemed too low – only one seemed to estimate the Kabupaten totals reasonably well. After some research, I found that only the CHROMY run that had stopped normally included the FREEZE file PSUs in the resulting SIZES file. The SIZES file for the other four, which were stopped manually, had no FREEZE file PSUs. When these were put back in manually, the estimates were greatly improved.

VI. CONCLUSIONS AND RECOMMENDATIONS

A. Recommendation 1.

Prepare separate CHROMY runs for each Kabupaten or Province that requests additional data.

It seems clear that the results obtained can be improved on an individual Kabupaten or Provincial basis. This exercise shows examples of samples that can be selected, however, tailoring the additional sample to the Kabupatens' budget and requirements for data precision makes more sense for this type of application.

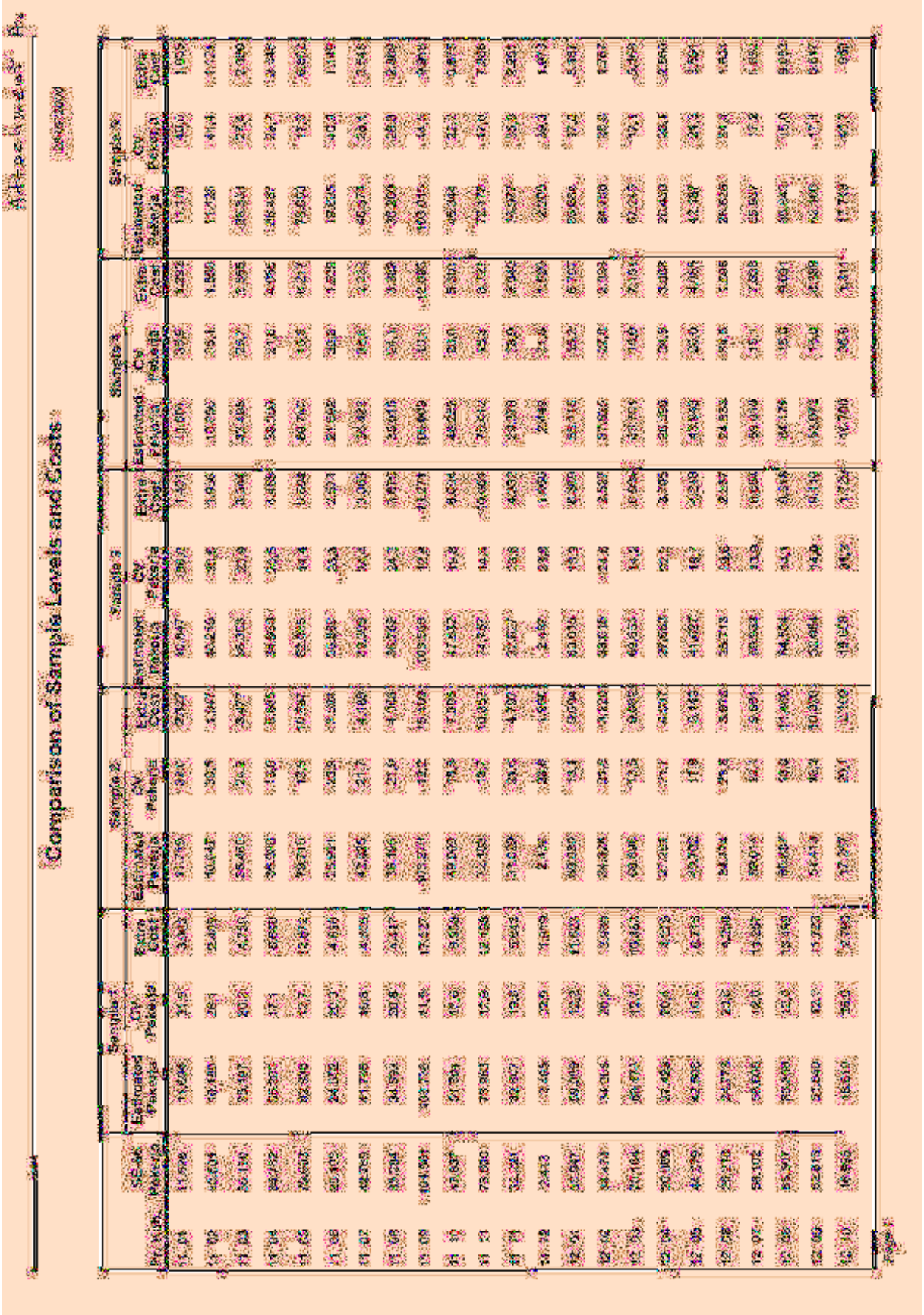
B. Recommendation 2.

Use the "Baby" CHROMY Interactive Program to prepare separate runs as suggested in Recommendation 1 above.

This will be a perfect test of the "Baby" CHROMY system that Farid has been preparing.

C. Recommendation 3.

Be sure to include a stopping rule that works when working with FREEZE files. If CHROMY runs are stopped manually, the resulting SIZE file does not include the FREEZE PSUs.



		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
12 11	21.316	21.227	20,4	4.161	22.294	23,5	3.600	21.194	25,3	2.992	21.368	24,7	2.774	22.398	26,0	2.859
12 12	174.119	177.291	8,0	25.229	176.930	9,0	20.852	173.012	10,0	17.109	175.186	11,0	14.183	176.124	12,2	12.198
12 13	83.988	85.021	12,1	11.495	82.895	13,1	9.747	83.243	14,0	8.588	84.617	15,1	7.047	85.478	17,2	6.118
12 71	12.332	11.647	25,6	2.992	13.227	28,9	2.585	13.962	31,5	2.536	12.792	36,1	1.966	11.482	46,2	1.158
12 72	14.361	14.656	24,7	3.125	14.745	29,7	2.356	14.986	32,7	2.109	13.960	37,3	1.662	16.056	35,4	1.814
12 73	26.202	26.723	19,8	5.025	24.682	21,1	4.683	27.415	21,4	4.322	28.843	25,4	3.163	27.404	30,9	2.479
12 74	15.708	16.432	24,2	3.144	15.200	25,8	2.916	14.673	31,5	2.014	15.647	31,7	2.137	14.455	37,2	1.548
12 75	229.074	226.974	8,0	31.131	227.087	9,0	25.450	229.715	10,0	22.087	226.705	11,1	17.812	236.309	12,0	16.169
12 76	27.768	27.978	20,2	4.987	27.554	23,9	3.648	28.032	29,7	2.736	31.678	33,3	2.185	32.033	37,9	1.598
13 01	1.281	1.256	58,6	779	1.250	56,7	760	1.250	56,7	760	1.125	62,2	598	1.176	60,9	579
13 02	38.934	39.628	16,1	8.512	38.369	17,3	6.868	38.325	19,2	5.738	38.801	19,9	5.158	39.966	21,1	4.873
13 03	40.384	40.233	18,5	7.248	41.356	17,8	5.928	39.715	19,8	4.769	41.064	20,1	4.512	43.657	21,3	4.237
13 04	38.256	39.091	20,1	5.225	37.979	22,0	4.303	40.196	23,5	4.246	39.481	23,8	3.648	39.855	26,3	3.287
13 05	56.283	56.679	12,2	12.473	56.108	13,3	10.963	57.139	14,1	10.003	57.032	14,9	9.224	57.075	16,9	7.248
13 06	75.253	76.124	12,1	13.794	75.649	13,1	11.837	74.506	14,1	10.165	74.973	15,1	9.329	72.819	17,4	6.963
13 07	69.461	65.801	12,5	14.079	67.736	13,2	13.129	67.396	14,1	11.675	66.949	15,0	10.148	65.274	17,1	8.369
13 08	52.576	50.791	12,7	13.689	51.178	13,3	12.188	50.972	14,2	10.925	50.385	15,3	9.699	51.455	16,1	8.493
13 09	61.508	61.041	12,7	11.485	59.637	13,5	9.766	60.644	14,3	9.006	62.550	15,0	8.483	59.765	17,2	6.393
13 71	93.201	90.478	12,1	14.240	92.573	12,9	12.768	94.687	14,0	11.476	90.123	15,1	10.013	95.301	17,0	8.635
13 72	9.830	9.843	25,3	3.040	9.515	30,4	2.289	9.513	31,4	2.270	10.502	34,4	1.900	8.571	35,7	1.767
13 73	6.913	7.050	24,3	2.916	7.366	29,4	2.318	6.336	38,3	1.434	6.037	48,5	1.064	6.340	45,9	1.054
13 74	7.852	8.114	24,1	2.812	7.137	30,5	1.900	7.826	36,3	1.548	9.121	40,9	1.425	6.105	52,8	817
13 75	23.003	24.802	25,7	3.885	23.617	24,6	3.695	22.563	28,0	3.002	23.690	30,0	2.878	24.427	32,9	2.346
13 76	18.539	18.034	25,0	3.496	18.120	27,9	2.650	21.495	31,6	2.717	16.739	35,2	1.788	17.229	36,2	1.833
14 01	23.413	23.336	21,7	4.607	23.334	23,4	3.733	22.195	25,0	3.135	23.852	24,9	3.211	25.224	27,1	2.697

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
14 02	30.761	30.969	20,4	5.111	30.255	22,4	4.227	29.162	24,2	3.498	33.284	28,2	2.878	27.318	35,8	1.624
14 03	35.101	35.158	20,0	4.797	36.904	22,4	3.960	32.052	24,4	3.230	33.749	24,5	3.097	36.379	25,7	3.011
14 04	6.624	6.944	26,5	2.356	6.571	31,3	1.653	6.840	31,0	1.736	6.855	34,8	1.311	7.079	35,0	1.406
14 05	16.069	16.798	25,5	2.565	15.261	26,5	2.213	16.267	29,2	1.966	16.224	29,7	1.919	16.921	36,1	1.282
14 06	27.704	27.875	20,3	4.911	27.625	24,4	3.600	25.461	29,5	2.413	27.684	33,8	1.847	25.634	35,8	1.586
14 07	14.628	14.331	25,3	2.945	14.925	27,7	2.546	13.864	29,4	2.118	13.578	35,3	1.634	14.675	35,9	1.672
14 08	37.246	36.515	22,3	5.006	42.445	24,8	4.417	36.955	25,1	3.895	37.406	26,3	3.306	34.587	30,7	2.945
14 09	24.963	25.929	20,8	4.769	24.136	22,2	4.313	24.726	22,8	3.847	23.948	26,3	2.783	25.618	30,3	2.289
14 10	26.283	26.924	19,8	5.234	25.961	22,1	4.265	26.949	24,0	3.914	26.148	29,1	2.726	26.676	34,8	2.014
14 11	15.339	15.271	25,3	3.154	16.032	27,7	2.793	14.764	33,1	2.185	14.594	34,9	1.914	13.709	37,5	1.539
14 12	4.132	3.887	25,3	2.432	3.901	31,5	1.833	4.126	35,8	1.558	4.461	41,9	1.206	3.481	56,3	636
14 71	72.710	73.152	11,8	14.250	73.344	18,1	12.587	72.439	14,0	11.210	73.190	14,8	10.488	73.199	16,9	8.312
14 72	34.358	33.622	20,0	5.890	29.850	21,7	5.234	34.868	22,3	5.909	35.931	23,4	5.415	34.893	26,6	4.683
14 73	22.220	21.660	25,4	3.059	23.952	30,3	2.356	23.561	29,4	2.270	22.697	35,1	1.833	22.254	36,0	1.453
15 01	32.233	31.649	20,2	4.465	32.336	19,9	4.541	30.688	21,6	4.018	32.536	22,2	3.971	33.123	23,7	3.610
15 02	26.695	26.638	20,2	4.436	25.490	24,1	2.907	27.364	27,9	1.947	26.831	33,5	1.463	25.445	35,6	1.244
15 03	26.935	26.181	20,0	3.619	27.364	21,6	3.182	25.316	24,5	2.422	26.360	28,4	2.375	27.398	28,5	1.852
15 04	16.649	15.324	25,0	2.327	17.374	26,4	2.527	17.287	30,1	1.814	16.200	36,6	1.254	17.883	35,8	1.349
15 05	18.232	18.511	24,6	2.954	17.795	25,1	2.622	15.909	28,7	2.042	17.524	30,2	1.672	17.135	34,0	1.444
15 06	19.048	20.704	25,1	2.612	22.393	29,3	2.023	23.061	29,5	2.109	21.765	37,3	1.235	21.036	36,1	1.282
15 07	16.528	17.186	28,3	2.556	18.184	29,8	1.691	15.650	30,6	1.577	16.798	34,0	1.263	19.156	34,2	1.273
15 08	14.528	14.867	25,6	2.261	14.941	28,2	2.042	15.442	29,3	1.786	13.517	31,8	1.415	15.228	35,2	1.178
15 09	26.693	26.525	20,2	3.173	25.406	20,5	3.049	23.464	21,9	2.346	26.426	24,0	2.052	23.715	28,4	1.387
15 71	48.069	45.726	15,6	7.704	50.144	18,0	6.403	49.290	19,9	5.405	47.374	20,2	4.674	48.611	21,7	4.493
16 01	92.013	93.400	12,1	13.651	94.842	13,0	12.407	93.971	13,9	10.829	101.077	15,4	9.604	95.215	17,1	7.600

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
16 02	93.935	98.136	11,7	15.494	95.598	12,3	13.423	95.076	12,9	12.663	97.547	13,2	11.970	97.814	14,1	10.725
16 03	59.345	59.914	12,3	12.635	59.904	13,2	11.143	59.610	14,1	9.756	60.309	15,0	8.797	62.304	17,0	7.153
16 04	61.256	62.001	13,6	12.141	63.272	14,2	11.058	60.728	15,3	9.120	63.226	16,1	8.398	62.644	17,6	7.115
16 05	75.199	76.272	12,1	14.202	77.222	13,1	12.853	76.933	14,0	11.229	77.652	15,1	10.174	73.385	17,2	7.638
16 06	89.517	97.183	12,1	13.338	94.098	13,2	11.096	87.881	14,1	10.070	89.500	15,7	8.455	88.675	16,9	7.134
16 07	56.303	55.380	12,6	13.214	55.585	13,4	11.789	54.314	14,5	10.250	56.062	15,2	9.728	55.964	17,0	7.980
16 08	18.768	18.421	24,9	3.258	19.134	30,4	2.441	18.765	30,0	2.375	17.945	34,0	1.995	19.000	35,8	1.748
16 71	162.911	162.882	8,0	31.825	162.558	9,0	26.258	163.978	10,0	22.391	161.023	11,0	18.325	159.044	12,0	16.311
16 72	18.005	18.394	24,0	3.458	17.984	27,6	3.239	18.522	29,4	2.564	16.916	38,8	1.615	17.445	41,1	1.577
17 01	27.774	26.762	20,0	4.398	27.159	23,1	3.315	26.271	24,0	3.125	29.654	25,8	2.897	25.698	29,8	2.014
17 02	37.896	36.520	20,3	4.474	37.921	24,4	3.088	34.515	26,6	2.261	34.924	35,6	1.520	41.286	33,9	1.662
17 03	34.383	35.169	19,8	4.257	36.750	23,3	3.078	32.851	24,8	2.688	31.678	26,2	2.470	31.678	29,8	1.758
17 71	32.804	31.736	19,6	5.348	32.181	22,3	4.455	33.222	23,8	4.132	30.684	29,3	2.669	37.314	33,5	2.251
18 01	32.864	34.192	21,3	4.638	32.817	24,2	3.600	35.565	28,3	2.593	29.966	35,2	1.567	31.753	35,3	1.615
18 02	80.629	82.680	12,3	12.730	81.298	13,3	10.963	81.987	14,2	9.813	81.695	15,1	8.559	79.932	16,3	7.942
18 03	113.828	116.472	10,7	17.831	118.211	11,4	15.996	117.767	12,2	14.003	117.538	13,1	12.321	113.811	14,2	10.583
18 04	122.498	121.573	10,2	19.960	122.026	11,1	16.720	121.875	11,9	15.114	122.840	12,1	14.497	123.643	13,0	12.766
18 05	134.938	135.928	10,2	20.662	137.511	11,0	18.107	137.120	11,8	16.577	136.046	12,0	15.409	135.500	13,0	13.357
18 06	45.830	45.552	16,1	7.799	47.729	18,0	6.593	49.417	19,9	5.396	41.368	22,1	4.360	45.949	24,1	3.847
18 07	25.565	25.504	20,7	4.246	24.456	24,7	3.108	25.303	29,4	2.270	23.392	34,8	1.539	31.610	38,8	1.377
18 08	43.020	46.318	16,4	7.847	44.299	17,8	6.004	43.780	19,5	4.987	43.679	20,0	4.607	42.834	21,9	4.037
18 71	99.753	98.221	12,1	16.292	97.569	12,9	14.468	100.170	14,0	12.302	101.415	14,9	11.827	98.211	17,1	8.493
18 72	19.318	20.010	24,2	3.090	18.215	29,2	3.258	18.113	30,4	3.021	20.514	29,2	3.030	16.826	33,8	2.185
31 71	286.701	284.005	6,0	54.368	285.754	7,0	43.008	285.094	7,9	35.084	290.762	9,0	29.222	296.587	10,0	23.721
31 72	335.118	337.912	6,0	60.249	333.061	7,0	45.742	337.068	8,0	37.639	334.944	9,0	30.495	335.864	10,1	25.374

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
31 73	219.879	220.591	8,0	30.485	219.582	9,0	24.709	215.881	10,0	20.140	225.054	11,0	17.609	218.714	12,0	15.162
31 74	349.321	348.156	6,0	55.394	349.389	7,0	43.586	353.157	8,0	35.482	345.867	9,0	29.383	340.832	10,1	23.047
31 75	232.751	236.310	8,0	37.496	232.135	9,0	30.096	232.067	10,0	26.505	236.372	11,0	21.973	229.992	11,9	18.288
32 01	135.284	135.933	16,6	15.038	137.339	17,1	13.176	137.533	17,6	11.172	134.925	18,5	9.908	135.514	19,2	8.322
32 02	89.571	93.139	12,3	13.433	92.389	13,1	11.932	92.709	14,1	10.374	95.708	15,0	9.082	92.036	17,0	7.296
32 03	336.405	340.597	6,1	58.282	343.947	7,0	45.647	341.325	8,0	35.375	345.288	9,0	29.888	338.050	10,0	24.225
32 04	249.714	255.145	8,4	50.359	257.636	7,2	40.641	254.081	8,1	31.920	254.754	9,1	25.973	261.697	9,9	23.436
32 05	200.002	196.316	6,1	33.734	196.062	9,1	27.512	198.363	10,0	23.484	203.010	10,9	20.254	197.302	12,0	16.957
32 06	428.092	427.969	6,1	60.676	430.727	7,0	46.825	428.122	8,0	38.541	426.305	9,0	31.445	439.835	9,8	26.647
32 07	207.445	208.328	8,1	33.820	205.463	9,1	26.837	205.018	9,8	24.139	208.358	10,0	22.344	201.590	11,0	19.389
32 08	267.665	267.501	6,1	56.810	269.636	7,0	45.334	292.547	8,0	36.584	290.993	9,0	30.096	284.634	9,9	25.213
32 09	238.495	236.755	8,1	34.741	234.690	9,0	27.749	238.250	9,9	24.348	236.421	10,0	23.132	239.943	10,9	20.909
32 10	100.463	97.504	12,0	14.734	100.519	13,0	13.100	98.099	14,0	11.010	99.095	15,1	9.823	101.796	17,0	7.790
32 11	251.275	254.496	6,5	52.250	255.748	7,3	41.971	256.882	8,1	34.323	255.854	9,0	27.607	253.454	9,8	23.693
32 12	161.695	160.581	8,2	31.758	163.185	9,0	27.350	161.435	10,0	22.543	163.286	11,0	19.332	164.534	12,0	16.891
32 13	98.533	96.551	11,7	16.026	100.550	12,0	15.019	101.408	12,9	13.642	97.855	13,1	12.445	103.630	14,0	11.751
32 14	210.575	212.186	8,1	33.582	212.076	9,1	27.398	210.967	10,0	23.104	213.000	11,0	19.209	220.037	12,0	17.337
32 15	189.458	196.620	8,2	30.875	196.887	9,0	25.194	197.386	10,0	21.717	194.604	11,0	18.008	198.304	11,9	15.589
32 16	78.795	80.115	12,1	14.867	79.892	13,0	12.815	80.748	13,9	10.915	80.594	15,0	9.443	79.936	17,1	7.267
32 17	203.389	208.311	8,1	31.141	208.278	9,1	25.536	209.770	10,0	21.660	208.042	11,0	17.650	208.800	12,0	15.323
32 18	253.096	253.665	6,7	51.879	253.232	7,4	42.398	254.314	8,3	35.549	252.567	9,1	29.345	250.334	10,1	23.769
32 19	240.652	240.204	8,0	35.995	243.244	9,0	30.058	247.168	10,0	25.289	245.542	11,0	20.681	248.909	11,8	18.278
32 20	179.488	179.450	8,3	32.358	180.690	9,1	27.569	180.282	10,1	22.971	179.334	11,1	19.114	179.440	12,1	16.492
32 71	42.380	41.958	16,6	7.894	44.320	18,6	6.726	44.182	20,2	5.833	45.816	23,9	4.655	44.388	24,0	4.541
32 72	19.402	19.635	24,2	3.334	18.484	26,3	3.144	16.110	29,9	2.593	19.938	29,6	2.441	19.619	34,1	2.004

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
32 73	302.705	309.544	6,0	57.684	300.105	7,0	43.681	308.536	8,0	36.185	299.024	9,0	28.014	303.674	10,0	23.626
32 74	40.212	39.878	16,2	8.160	39.464	18,5	6.593	42.859	20,5	5.510	39.629	22,9	4.816	42.302	23,2	4.408
32 75	121.179	121.617	9,9	21.707	121.104	10,9	18.620	121.555	12,1	15.779	122.679	12,9	14.202	123.339	14,1	12.464
32 76	92.996	93.908	11,9	16.473	93.003	13,0	14.354	95.033	14,0	13.146	94.087	15,1	11.039	94.756	16,9	9.272
32 77	98.490	98.085	12,0	16.378	100.118	13,0	14.654	101.736	14,1	12.616	98.573	15,2	10.754	99.008	17,0	8.930
32 78	29.315	30.301	19,9	5.719	28.860	24,3	3.762	30.127	30,1	3.030	27.987	29,7	2.631	27.684	33,1	2.422
33 01	215.034	216.422	8,0	33.905	214.671	9,0	27.350	215.185	10,0	22.534	215.154	11,0	18.868	210.605	11,8	16.549
33 02	226.013	223.497	8,1	33.715	223.458	9,1	27.350	215.935	9,8	24.234	225.214	10,0	22.914	222.005	10,8	20.284
33 03	137.297	139.253	10,4	21.023	140.307	11,2	17.945	138.528	11,6	16.767	140.074	12,2	15.161	140.077	13,0	13.176
33 04	124.960	123.684	10,1	23.180	122.721	10,6	21.793	123.589	11,0	19.674	124.502	11,9	17.290	127.075	12,9	15.352
33 05	194.736	197.695	8,6	31.868	199.020	9,3	26.951	195.955	10,2	22.344	197.552	11,2	19.038	191.724	12,2	16.093
33 06	108.166	107.425	10,5	19.218	107.591	11,3	16.936	109.563	12,0	15.038	108.652	13,1	13.186	108.787	14,0	11.647
33 07	138.276	139.035	12,1	18.734	140.974	12,6	16.264	138.150	13,7	14.174	139.938	14,1	13.262	135.597	15,2	11.381
33 08	211.444	210.822	9,6	30.996	211.831	10,2	26.391	213.364	10,9	22.363	216.115	11,6	20.216	211.797	12,4	17.233
33 09	111.372	112.355	10,4	19.152	112.805	11,2	16.511	110.793	11,8	15.475	112.330	12,1	14.240	112.659	13,0	12.264
33 10	189.803	191.934	6,1	31.682	190.801	9,1	26.030	189.710	10,1	21.451	189.109	10,9	18.924	188.030	11,8	16.425
33 11	104.688	103.964	10,1	20.691	103.206	11,0	17.328	105.861	12,0	15.817	107.437	12,9	14.098	107.080	13,9	12.226
33 12	135.178	138.632	10,1	20.681	135.976	11,0	17.670	138.101	11,8	16.406	139.017	12,0	15.352	132.947	13,1	12.473
33 13	133.599	136.852	10,1	21.441	138.266	11,0	18.810	134.208	12,0	16.834	135.550	13,0	13.461	135.510	14,2	11.884
33 14	133.006	134.971	10,3	21.546	135.054	11,1	18.496	135.731	12,0	16.007	132.959	13,0	13.414	132.460	14,1	11.647
33 15	144.877	146.493	10,4	19.466	147.067	11,3	16.292	150.618	12,1	14.658	148.835	13,5	12.245	152.512	13,9	11.191
33 16	74.617	74.269	12,2	13.395	74.700	13,1	11.694	73.998	14,0	10.136	73.885	14,9	9.281	71.303	17,3	6.973
33 17	69.793	70.043	12,8	13.110	69.411	13,6	11.466	71.236	14,3	10.535	73.244	15,1	9.937	75.362	17,7	7.752
33 18	164.330	163.997	8,6	29.250	163.744	9,3	24.633	161.636	10,2	20.567	162.406	11,1	17.575	162.952	12,1	15.190
33 19	103.654	102.655	10,4	18.857	105.323	11,2	16.938	103.416	12,2	14.250	105.397	13,1	12.634	105.043	14,0	11.048

			Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr.	Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
33	20	191.695	163.076	8,7	30.799	162.119	9,6	25.460	161.662	10,4	21.527	162.724	10,9	19.940	162.283	11,2	18.857
33	21	107.610	107.528	10,1	20.396	106.964	11,0	17.280	109.703	12,0	15.670	108.930	13,1	13.157	109.367	13,8	12.169
33	22	116.497	116.274	10,5	20.643	116.190	11,1	16.895	115.239	12,1	15.732	113.646	13,1	13.679	114.043	14,0	11.732
33	23	250.537	251.325	10,7	20.193	254.974	10,9	25.611	252.975	11,4	22.420	254.156	11,3	22.372	251.655	11,9	18.601
33	24	102.050	101.680	10,2	20.162	101.845	11,1	16.976	100.503	12,1	14.611	101.712	13,0	13.024	103.381	14,0	11.426
33	25	104.504	106.196	10,6	20.225	106.002	11,5	17.451	107.052	12,3	15.171	107.744	13,3	13.642	107.143	14,1	12.207
33	26	123.190	124.267	10,5	20.054	121.031	11,0	18.000	124.646	11,3	17.793	125.157	12,1	15.703	125.254	13,1	13.699
33	27	120.874	118.653	10,0	21.051	117.483	11,1	17.974	116.533	12,0	15.779	118.868	12,9	13.946	118.212	14,0	12.036
33	28	173.042	173.181	8,1	31.616	174.903	9,0	26.875	173.029	9,9	22.040	174.259	11,0	18.715	169.563	12,0	15.426
33	29	165.550	168.631	8,2	30.818	168.549	9,1	25.222	167.643	10,1	21.042	167.400	11,1	17.556	168.524	11,9	15.513
33	71	23.728	21.562	20,9	4.707	22.977	22,5	4.208	22.027	24,7	3.629	24.568	29,9	2.707	20.396	36,0	1.719
33	72	103.468	103.481	10,0	20.166	105.100	10,9	17.812	104.122	12,0	15.209	103.036	13,0	12.891	102.828	14,0	11.485
33	73	25.869	26.809	19,9	5.388	26.401	21,7	4.626	26.771	24,1	3.687	25.686	28,5	2.850	28.358	34,0	2.168
33	74	180.161	179.671	8,0	32.423	179.690	9,0	26.467	180.740	10,0	22.790	180.315	11,0	19.076	180.609	12,0	16.216
33	75	47.259	47.833	19,4	6.118	44.593	20,4	5.358	46.385	22,0	5.101	46.670	24,1	4.294	47.873	25,5	3.847
33	76	38.681	38.346	19,9	5.462	39.003	21,8	4.586	38.125	24,3	3.705	37.241	26,7	2.831	35.513	36,5	1.805
34	01	61.915	62.645	12,6	12.739	62.842	13,3	11.324	62.067	14,2	10.440	63.262	14,2	10.165	63.190	15,1	8.835
34	02	161.699	160.220	8,3	32.993	160.157	9,1	27.369	161.031	10,1	23.275	156.911	11,2	19.475	158.134	12,1	17.195
34	03	100.028	100.670	10,5	17.489	99.809	11,3	14.658	100.604	12,2	12.730	100.064	13,2	11.172	99.657	14,1	9.595
34	04	145.753	144.305	10,1	20.577	145.110	11,0	17.822	147.206	12,0	15.380	147.840	13,0	13.338	146.741	14,0	11.495
34	71	106.817	106.712	10,0	19.950	107.945	10,9	17.394	109.524	11,9	15.323	107.909	13,1	13.110	110.281	13,9	11.832
35	01	69.600	70.464	13,1	12.659	71.402	13,8	11.400	68.796	14,5	10.678	70.614	14,6	10.013	70.096	15,5	9.006
35	02	120.776	122.709	10,3	20.197	120.486	11,2	17.005	125.587	12,1	16.653	122.822	12,1	14.839	122.651	13,1	12.939
35	03	104.915	108.258	10,4	19.665	109.363	11,3	17.204	104.903	12,4	13.851	108.192	13,2	12.426	108.517	14,1	11.106
35	04	183.437	183.809	8,4	32.347	187.171	9,3	26.647	188.951	10,2	22.477	188.727	11,1	19.199	187.183	12,0	17.168

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
35 05	184.565	191.365	8,5	31.834	192.915	9,3	26.571	189.672	10,2	21.992	191.324	11,2	18.948	186.617	12,3	15.656
35 06	195.233	200.774	8,4	30.124	198.815	9,3	24.348	196.067	10,2	20.736	197.561	11,1	17.309	204.733	12,0	15.380
35 07	288.491	296.756	6,1	54.900	297.661	7,0	42.284	292.148	8,0	33.715	297.943	9,0	26.847	302.655	10,0	22.923
35 08	126.328	129.229	10,0	20.985	127.589	11,1	17.358	126.301	12,0	15.171	129.574	13,0	13.176	132.512	14,0	11.552
35 09	259.098	267.321	8,2	53.561	267.852	7,1	41.524	269.021	8,0	34.371	261.000	9,0	26.835	257.307	10,0	22.021
35 10	242.696	247.886	8,2	33.345	247.049	9,1	26.742	250.683	9,8	25.146	246.640	10,1	22.265	245.309	11,0	18.373
35 11	92.892	97.360	11,9	14.696	95.156	13,2	12.112	95.723	14,5	10.374	92.749	15,5	9.148	94.310	16,1	8.037
35 12	87.014	89.081	12,0	14.839	88.442	12,9	12.730	88.460	13,9	11.039	91.203	14,9	10.421	85.574	17,0	7.999
35 13	126.353	126.999	11,4	18.116	128.029	12,0	15.817	127.582	13,0	13.642	129.687	13,9	12.046	125.258	14,7	10.621
35 14	185.946	190.483	8,1	32.727	188.956	9,0	25.039	190.847	10,0	22.249	186.750	11,0	18.715	190.395	11,9	15.637
35 15	207.992	210.252	8,1	34.732	209.128	9,0	26.082	210.439	10,1	23.769	202.004	11,1	19.380	214.415	12,0	17.717
35 16	131.662	132.705	10,0	20.491	130.632	11,1	17.185	133.336	12,0	14.829	129.190	13,0	12.245	130.249	14,1	10.858
35 17	166.316	173.207	8,3	30.789	175.738	9,1	25.507	172.302	10,1	21.185	167.751	11,1	17.727	169.209	12,0	15.238
35 18	166.907	161.290	18,0	24.225	161.572	18,4	20.320	162.452	18,7	17.727	161.315	19,3	14.677	161.561	19,7	13.243
35 19	76.141	78.605	12,3	13.974	76.908	13,4	11.704	77.188	14,2	10.298	79.520	14,9	9.614	77.406	16,9	7.361
35 20	104.181	104.381	10,3	20.909	106.416	11,1	18.449	105.793	12,0	15.922	104.500	13,1	13.547	104.760	14,0	11.969
35 21	116.794	119.578	12,0	16.074	120.455	12,8	14.107	117.599	13,8	12.074	124.260	14,6	10.716	118.536	14,6	9.737
35 22	139.249	139.067	11,0	19.907	142.875	11,5	17.938	140.047	12,1	16.378	139.439	12,4	15.096	141.398	13,3	13.290
35 23	128.627	129.520	10,3	19.760	128.670	11,3	16.998	125.730	12,3	14.506	131.745	13,2	12.435	129.572	14,1	10.815
35 24	154.207	153.886	8,9	34.722	157.455	9,5	27.055	158.351	10,2	22.220	155.460	11,1	18.031	158.807	12,0	14.307
35 25	149.861	153.779	9,1	31.131	153.986	9,8	28.362	153.277	10,6	22.097	154.904	11,5	19.133	151.456	12,5	15.751
35 26	101.964	103.584	11,3	24.434	103.742	11,9	19.864	100.981	13,0	15.903	101.645	13,6	13.660	104.834	14,1	11.818
35 27	76.323	76.365	13,3	11.590	76.650	13,8	10.858	76.921	14,7	9.747	75.716	15,2	8.741	76.229	17,1	7.305
35 28	88.118	87.387	12,6	12.958	86.393	13,6	11.276	85.458	14,5	10.003	85.478	15,8	7.875	86.955	16,9	7.571
35 29	160.981	168.151	8,2	29.696	166.867	9,0	24.709	167.477	10,0	20.216	165.633	10,7	18.838	165.219	11,1	16.948

			Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr.	Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
35	71	46.722	45.867	15,5	9.471	43.011	17,5	7.985	49.077	21,9	6.715	47.207	20,0	6.032	45.931	21,4	5.491
35	72	26.030	25.579	20,0	5.415	26.568	21,7	4.788	27.202	23,4	4.248	27.327	26,3	3.657	26.049	32,2	2.498
35	73	134.899	134.889	10,0	22.572	134.951	11,0	19.437	130.859	12,1	16.321	133.621	13,0	14.316	133.320	14,1	12.625
35	74	31.352	31.999	19,7	5.424	32.065	21,6	4.655	33.195	23,6	4.113	30.071	27,6	3.239	31.392	26,3	2.774
35	75	36.126	38.028	20,5	6.013	37.905	20,2	5.975	38.851	21,9	5.377	31.554	24,9	4.360	37.616	25,6	4.275
35	76	23.910	24.639	24,1	3.752	23.541	27,3	3.391	24.417	31,3	2.707	22.161	36,6	2.014	24.969	35,6	2.023
35	77	34.104	33.692	20,3	5.111	35.696	21,4	5.149	34.666	21,6	4.531	33.597	24,1	3.714	33.197	26,5	3.173
35	78	450.963	454.578	6,0	54.163	443.313	7,0	47.832	449.131	8,0	38.788	460.373	9,0	31.967	464.927	10,0	26.571
51	01	36.748	38.696	20,9	5.348	37.382	20,1	5.130	37.207	21,5	4.607	37.662	21,8	4.360	35.069	23,6	3.752
51	02	48.827	52.423	19,3	6.327	46.311	20,2	5.082	44.994	23,9	4.170	48.744	23,6	3.999	47.509	26,1	3.116
51	03	76.346	74.098	12,5	13.787	76.982	13,0	13.186	75.409	14,1	11.496	75.536	15,3	10.269	76.562	17,5	8.474
51	04	105.321	109.103	11,3	19.275	106.846	12,1	16.549	109.909	12,6	15.494	108.338	13,4	13.604	108.867	14,2	11.941
51	05	22.329	22.917	24,1	3.904	24.258	24,2	3.914	23.926	28,7	2.954	21.880	26,3	2.717	23.312	31,7	2.213
51	06	31.741	33.166	20,3	6.099	30.972	22,1	5.063	30.936	22,3	4.626	31.585	23,3	4.379	30.837	27,2	3.011
51	07	47.966	49.650	16,6	9.823	47.446	17,3	8.122	48.267	18,4	6.732	45.840	20,4	5.158	51.862	22,3	4.864
51	08	69.839	69.127	12,5	13.385	69.791	13,2	11.884	69.395	14,2	10.241	71.227	14,6	9.832	68.352	15,2	8.625
51	71	78.491	77.568	12,1	13.841	79.476	12,9	12.568	77.741	14,1	10.587	78.964	15,0	9.927	79.912	16,9	8.046
52	01	83.746	82.398	12,1	14.953	79.231	13,2	12.606	82.282	13,9	11.143	81.252	15,1	10.317	84.507	17,4	8.238
52	02	103.812	101.913	11,2	19.065	103.699	11,9	16.691	101.853	12,8	13.898	100.876	13,7	12.321	102.294	14,4	11.124
52	03	146.619	145.469	10,7	23.712	143.017	11,8	18.990	146.640	12,3	16.330	149.378	13,2	14.753	142.024	14,3	12.188
52	04	46.355	48.405	19,4	6.878	46.950	20,7	5.833	43.897	22,7	4.607	46.348	22,1	4.522	46.704	24,0	4.037
52	05	16.089	15.485	24,9	3.420	15.302	28,3	2.783	17.085	32,6	2.745	15.712	35,1	1.909	17.554	40,6	1.586
52	06	60.269	57.598	16,5	12.426	59.109	16,8	11.476	58.886	17,5	10.402	59.503	18,0	9.006	57.845	19,7	7.324
52	71	52.529	52.081	12,0	12.997	52.657	12,9	11.561	52.807	14,1	10.222	50.612	15,3	8.818	52.538	17,0	7.657
53	01	31.958	33.659	18,5	5.966	30.971	19,9	5.006	33.470	22,5	4.787	33.928	21,5	4.484	32.584	25,5	3.211

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
53 02	13.892	13.312	23,8	2.812	14.058	24,5	2.983	13.210	27,4	2.209	12.420	30,9	1.757	13.616	32,3	1.643
53 03	20.644	21.236	23,0	3.885	21.173	24,9	3.068	20.436	28,1	2.774	22.027	27,5	2.698	19.682	32,8	2.080
53 04	14.549	13.418	25,4	3.011	13.124	31,2	2.014	14.902	29,5	2.365	13.676	36,5	1.662	16.260	41,3	1.434
53 05	9.958	10.056	25,2	3.067	9.809	29,9	2.213	10.195	35,0	1.729	9.384	44,4	1.406	11.093	43,2	1.330
53 06	18.399	17.665	25,0	3.230	18.197	27,3	2.690	15.233	32,5	2.261	18.014	34,1	1.909	18.490	39,2	1.862
53 07	11.034	9.799	26,0	2.546	11.216	32,9	2.308	9.884	35,8	1.890	9.547	37,4	1.615	11.376	34,5	1.624
53 08	4.062	4.457	28,3	1.871	4.052	32,0	1.262	4.183	38,5	1.007	4.007	40,7	902	2.667	57,3	418
53 09	22.425	22.305	24,2	3.439	22.554	25,2	2.973	25.413	26,1	2.898	22.412	30,9	2.023	22.891	32,5	2.299
53 10	28.974	29.853	22,9	4.835	29.806	23,9	4.227	28.936	26,6	3.636	28.102	26,5	3.059	31.023	25,7	3.383
53 11	36.531	39.594	15,9	8.445	41.164	18,3	7.334	50.208	28,3	6.118	41.142	19,8	5.366	40.006	21,3	4.664
53 12	17.193	16.284	25,7	2.869	18.427	27,1	2.869	16.943	30,1	2.090	15.009	31,5	2.090	17.627	36,7	1.556
53 13	31.923	36.048	18,0	6.365	33.909	19,5	5.187	34.553	23,0	4.218	35.446	23,5	3.572	35.256	26,1	3.344
53 71	31.592	31.326	19,5	5.214	29.710	22,5	4.037	31.226	23,7	3.657	31.639	30,1	2.574	26.251	34,6	2.204
61 01	53.948	53.517	13,9	10.868	53.807	14,5	9.823	52.350	15,2	8.654	53.952	15,8	7.913	56.332	18,0	6.792
61 02	27.792	28.269	19,4	4.364	27.634	20,7	4.645	28.643	21,8	4.075	27.618	28,3	2.331	28.697	30,0	2.261
61 03	16.657	17.629	25,0	2.755	16.823	25,3	2.603	15.842	31,4	1.805	17.531	29,4	1.919	16.051	36,3	1.263
61 04	80.188	81.290	12,8	12.815	81.058	13,6	11.352	81.475	14,3	10.117	81.362	15,0	9.538	80.925	15,2	8.977
61 05	42.702	39.598	16,9	7.191	42.075	18,8	5.916	44.650	20,8	4.826	42.456	20,3	4.579	44.973	22,0	4.341
61 06	31.543	30.743	20,1	4.635	30.510	23,2	3.591	31.508	23,5	3.382	30.420	25,8	2.935	32.083	28,0	2.898
61 07	60.705	61.394	15,0	9.481	60.989	15,3	8.692	60.463	15,8	7.799	60.738	16,3	6.992	60.024	17,7	5.947
61 08	16.479	16.595	26,2	2.669	17.347	29,9	1.395	16.015	29,2	1.881	15.743	36,7	1.282	16.584	36,4	1.244
61 71	52.414	51.007	12,1	11.668	52.700	12,8	10.839	51.720	14,2	9.433	52.277	15,0	8.749	51.748	16,6	7.847
62 01	29.173	29.344	15,9	20.681	29.344	15,9	20.681	29.344	15,9	20.681	29.344	15,9	20.681	29.344	15,9	20.681
62 02	79.241	76.615	13,3	12.483	76.679	14,0	11.286	75.885	14,9	10.241	76.373	15,6	9.205	72.606	17,2	7.552
62 03	67.142	66.582	12,4	11.675	67.980	13,2	10.687	68.407	14,0	9.495	68.183	14,9	8.521	68.491	17,2	6.526

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
62 04	20.103	20.001	25,0	3.097	21.753	24,2	3.277	21.626	29,4	2.185	20.197	29,8	2.004	24.716	32,7	1.957
62 05	31.117	31.517	19,9	4.807	26.561	24,0	3.163	28.014	29,3	1.965	31.729	28,3	2.128	25.234	33,8	1.434
62 71	22.634	22.406	25,2	3.277	22.150	26,6	3.135	23.781	28,9	2.793	21.011	31,4	2.289	22.873	35,6	1.833
63 01	28.849	29.936	20,7	5.082	30.694	24,3	3.648	28.275	25,3	3.192	31.108	28,7	3.154	29.406	28,3	2.698
63 02	65.838	67.755	13,8	12.844	67.017	14,8	11.343	66.795	15,3	10.667	68.082	16,1	9.424	70.805	17,5	8.198
63 03	57.298	57.308	12,4	13.946	56.325	13,3	12.141	57.353	14,0	11.172	56.834	15,0	10.106	56.911	16,9	8.198
63 04	29.287	27.527	20,7	4.603	29.521	20,2	4.864	30.750	22,0	4.408	29.005	22,1	4.047	30.324	24,2	3.695
63 05	21.871	24.316	23,9	3.477	23.282	24,4	3.011	19.706	26,4	2.622	21.642	30,0	2.071	23.949	33,1	1.862
63 06	42.356	40.495	20,0	5.709	40.130	20,3	5.396	43.114	20,5	5.367	42.611	21,7	4.778	39.136	24,0	4.170
63 07	34.545	35.543	20,1	5.092	33.809	23,0	3.857	33.375	24,1	3.467	34.381	25,9	2.954	32.233	28,2	2.490
63 08	70.469	69.751	12,0	15.087	68.988	13,1	13.072	69.003	13,5	12.635	68.953	14,1	11.523	69.683	15,1	10.098
63 09	30.197	30.814	20,3	5.158	32.002	22,5	4.598	30.178	22,2	4.389	30.503	23,9	3.885	31.002	25,7	3.439
63 71	92.402	92.484	12,1	16.473	91.127	13,0	13.841	92.504	14,1	12.730	93.539	14,9	11.210	94.692	16,7	9.234
63 72	19.212	19.697	24,6	3.647	19.122	28,8	2.859	18.160	29,5	2.302	18.111	35,1	2.213	18.396	36,3	2.071
64 01	28.981	28.781	20,6	5.035	30.400	21,7	4.579	30.410	23,4	3.980	31.400	23,0	2.764	28.926	34,1	2.014
64 02	11.715	12.155	31,3	2.441	11.506	34,3	1.928	11.319	34,4	1.938	9.685	39,9	1.406	11.564	43,0	1.244
64 03	42.009	39.909	16,3	7.163	45.487	18,2	6.346	47.307	20,2	5.158	42.975	19,8	4.778	38.494	21,0	3.933
64 04	7.838	8.099	25,9	2.232	7.276	31,7	1.377	7.844	29,8	1.558	8.392	31,0	1.588	7.937	34,6	1.140
64 05	10.836	10.603	24,5	2.983	10.720	31,8	2.194	11.610	35,9	2.194	9.898	34,6	1.824	10.400	47,3	1.491
64 06	2.864	2.595	71,3	190	4.885	81,8	190	6.816	86,9	160	2.595	71,3	190	4.894	81,8	190
64 07	8.249	8.296	41,9	2.023	7.958	44,8	1.539	7.958	44,8	1.539	8.339	47,2	1.149	7.569	52,5	874
64 08	2.824	2.578	36,3	1.491	2.460	44,2	997	2.795	40,8	1.159	2.984	40,0	1.092	3.136	42,8	998
64 71	44.018	46.390	16,6	6.673	44.995	17,8	7.400	43.600	20,4	5.605	40.623	23,0	4.626	43.265	24,0	4.189
64 72	67.576	67.590	12,0	13.357	69.179	12,9	12.150	66.956	14,2	10.338	65.631	14,9	9.158	66.699	16,9	7.239
64 73	14.582	14.356	29,1	2.579	14.287	32,7	1.995	15.170	33,7	2.052	14.195	37,2	1.548	15.564	39,1	1.434

			Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr.	Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
64	74	9.279	9.284	25,7	2.878	8.300	30,7	2.109	10.544	28,4	2.517	9.906	34,3	1.881	7.510	47,4	997
71	01	16.741	16.236	25,6	3.135	18.504	28,2	2.650	14.426	32,3	2.175	14.494	36,2	1.634	13.571	36,5	1.577
71	02	49.781	53.770	16,3	8.863	50.708	18,1	7.372	63.350	19,5	5.958	49.636	19,9	5.424	50.215	21,8	5.044
71	03	40.770	37.123	20,5	4.940	40.598	22,0	4.617	40.730	24,7	3.781	37.696	28,4	2.935	41.221	28,7	2.736
71	04	214.473	214.190	10,2	28.823	213.147	10,7	23.873	214.453	11,3	20.254	217.719	11,9	17.907	216.745	12,9	15.304
71	05	38.830	37.781	20,2	5.092	39.655	21,2	4.512	40.426	21,3	4.474	37.727	23,2	3.638	37.297	24,8	3.116
71	71	18.257	18.691	24,7	3.553	19.765	28,2	3.182	24.136	30,1	3.125	18.853	28,6	3.249	18.233	29,6	2.460
71	72	46.714	44.167	17,0	8.368	49.492	19,7	7.400	45.505	20,4	5.909	48.729	23,1	5.082	43.246	24,3	4.142
71	73	16.145	16.041	24,6	3.382	16.650	28,6	3.154	16.073	29,0	2.631	16.901	35,8	1.767	17.072	42,7	1.218
72	01	17.670	17.583	28,8	2.688	17.336	29,5	2.555	17.358	31,3	1.936	17.184	34,3	1.548	17.852	34,1	1.415
72	02	32.734	33.385	20,0	4.864	33.764	23,7	3.572	31.724	24,4	3.239	31.714	26,2	3.002	35.003	26,2	2.964
72	03	13.024	13.542	29,7	2.375	12.172	34,0	1.710	13.514	34,4	1.767	14.400	35,7	1.491	12.271	38,4	1.301
72	04	34.538	34.261	20,2	4.037	34.845	24,1	2.935	34.508	27,4	2.242	34.566	29,1	1.938	35.086	33,6	1.415
72	05	73.730	74.280	12,3	12.369	74.352	13,1	10.725	75.113	13,9	9.509	74.642	15,0	8.512	78.378	17,0	6.859
72	06	25.713	25.253	21,0	4.638	25.497	22,1	4.132	24.969	24,2	3.534	22.165	28,4	2.546	23.670	29,7	2.299
72	07	12.553	12.221	25,3	2.945	12.508	29,2	2.546	12.990	38,6	2.194	11.708	33,8	1.881	12.073	40,5	1.358
72	71	34.442	33.138	20,0	5.149	33.781	22,2	4.398	33.734	24,0	3.942	36.234	26,9	3.097	36.028	33,7	2.251
73	01	11.523	11.854	25,2	2.973	11.747	25,2	2.916	11.637	25,5	2.850	11.912	24,7	2.964	11.872	24,7	2.945
73	02	40.563	42.017	15,9	7.847	40.975	17,2	6.784	42.920	18,6	5.886	40.309	20,3	4.828	39.862	21,0	4.769
73	03	16.011	16.778	24,7	3.372	17.562	26,0	2.926	19.187	32,3	2.565	17.243	33,2	1.947	18.628	36,4	1.662
73	04	43.904	43.304	21,4	6.023	45.275	25,8	4.569	44.149	24,4	4.569	44.130	26,2	4.037	39.987	30,9	3.429
73	05	32.878	35.657	20,0	5.044	35.593	23,6	3.628	34.521	24,2	3.562	33.392	26,9	2.964	36.910	26,4	2.916
73	06	61.203	65.797	12,2	12.093	63.768	13,3	10.260	64.839	14,2	9.300	62.980	14,4	8.625	64.738	15,0	7.904
73	07	21.126	22.115	24,6	3.496	22.264	24,7	3.182	21.242	29,2	2.527	20.315	30,1	2.156	21.152	34,2	2.042
73	08	27.691	29.835	20,1	4.987	28.731	24,5	3.515	28.281	29,5	2.337	26.032	35,9	1.757	34.383	32,3	2.052

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
73 09	29.750	30.884	20,2	4.702	29.636	22,5	3.771	31.838	24,2	3.410	30.402	28,8	2.394	28.911	34,7	1.615
73 10	21.276	22.588	24,9	3.534	23.857	28,9	2.774	20.052	38,0	1.738	22.644	39,9	1.311	20.298	42,8	1.206
73 11	65.044	66.310	12,2	13.927	65.332	13,2	11.903	65.966	14,1	10.630	66.006	15,0	9.414	64.456	17,0	7.866
73 12	24.266	24.942	21,0	4.218	24.898	22,3	3.866	24.253	23,8	3.543	24.623	24,2	3.135	25.679	25,9	2.850
73 13	61.078	63.630	12,5	13.764	63.010	13,4	12.055	63.637	14,1	10.801	63.437	15,0	9.775	61.477	16,9	7.913
73 14	26.414	27.874	20,2	4.978	27.012	22,2	4.484	27.685	23,7	3.971	27.334	24,1	3.733	29.646	25,1	3.467
73 15	31.799	30.046	20,2	5.358	32.436	21,7	4.769	28.413	24,8	3.733	30.590	27,9	2.793	33.337	32,9	2.270
73 16	12.321	12.175	25,7	3.068	11.735	30,4	2.318	14.117	33,5	2.137	11.927	35,4	1.919	12.498	44,5	1.292
73 17	39.225	41.050	15,7	8.132	39.752	17,9	6.545	40.573	19,7	5.576	38.617	20,0	5.082	39.370	21,0	4.856
73 18	30.517	30.507	20,4	4.911	31.319	22,1	4.190	28.764	23,9	3.952	29.924	27,4	3.135	31.266	28,5	2.527
73 19	52.473	56.663	11,9	13.166	55.564	13,1	11.428	54.811	14,0	10.127	54.708	15,2	8.550	56.560	16,5	7.089
73 20	14.386	13.581	24,9	3.125	15.456	30,9	2.783	13.561	30,8	2.346	11.965	38,9	1.805	14.870	35,0	1.938
73 21	21.070	20.974	24,3	3.391	21.947	29,5	2.498	22.406	29,2	2.498	21.043	38,2	1.634	23.234	38,7	1.425
73 22	39.729	38.828	20,5	5.206	38.447	22,4	4.484	35.621	24,3	3.695	40.132	23,9	4.056	39.209	25,9	3.553
73 71	135.571	135.123	10,0	21.945	137.404	11,0	19.066	136.763	12,0	16.207	137.292	12,9	14.354	135.526	14,1	12.084
73 72	12.834	13.416	24,5	3.249	12.417	30,5	2.232	12.042	31,2	2.204	12.887	35,7	1.681	13.059	42,7	1.254
74 01	46.129	46.735	16,1	8.113	45.957	17,4	6.764	43.636	19,4	5.358	45.918	19,6	4.921	43.206	21,3	4.322
74 02	33.618	36.091	20,1	4.778	34.504	20,0	4.218	31.645	23,3	3.391	34.465	24,1	2.992	34.905	25,6	2.774
74 03	34.181	34.906	19,6	4.493	35.431	19,8	4.550	32.552	21,2	3.847	32.472	22,6	3.306	34.059	24,2	3.021
74 04	31.146	31.479	20,2	4.693	31.152	21,8	3.990	32.011	23,9	3.363	33.533	28,3	2.774	30.066	28,4	2.223
74 71	18.031	17.751	25,2	3.059	19.131	30,3	2.878	18.794	28,5	2.498	18.736	34,9	1.719	17.808	47,4	1.007
81 01	10.210	8.970	44,3	693	8.970	44,3	693	8.963	47,1	627	11.296	55,3	418	11.726	45,0	665
81 02	22.089	23.045	46,6	1.862	21.122	48,6	1.814	22.417	46,3	1.966	22.393	46,9	1.615	19.269	53,4	1.187
81 03	52.447	53.134	13,5	10.488	52.239	14,2	9.272	50.885	15,2	7.799	51.562	15,4	7.657	53.258	17,0	6.308
81 04	8.966	8.453	28,8	1.985	9.451	30,2	1.852	10.099	37,4	1.539	8.390	36,3	1.254	10.095	42,8	1.368

		Sample 1			Sample 2			Sample 3			Sample 4			Sample 5		
Pr. Kab.	SE-96 Pekerja	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost	Estimated Pekerja	CV Pekerja	Extra Cost
81 71	35.671	36.665	19,6	5.709	36.552	21,9	4.683	36.876	23,9	4.132	32.529	28,3	3.059	33.679	29,9	2.983
82 01	48.218	47.897	18,0	11.656	47.897	18,0	11.656	47.897	18,0	11.656	47.897	18,0	11.656	47.897	18,0	11.656
82 02	18.423	17.695	25,1	2.479	18.618	26,3	2.327	17.306	30,0	1.852	20.748	37,4	1.366	15.102	38,0	1.149
82 71	12.482	12.537	28,0	2.612	12.774	29,2	1.881	12.360	31,9	1.643	14.307	33,9	1.567	11.265	36,2	1.368
91 01	9.408	9.649	30,3	2.071	9.274	33,9	1.539	9.803	32,9	1.577	8.421	37,2	1.140	8.585	38,6	1.216
91 02	9.045	9.156	24,9	2.622	8.975	30,5	2.042	8.705	30,9	1.957	8.724	36,9	1.520	10.141	41,5	1.263
91 03	20.147	20.728	25,2	2.527	20.133	26,2	2.593	19.316	28,0	2.080	21.223	29,4	1.968	24.264	35,6	1.539
91 71	12.223	12.066	30,8	2.413	11.346	34,6	1.814	11.195	33,3	1.919	12.128	37,1	1.577	12.399	44,4	1.149
92 01	1.679	1.657	39,2	1.178	1.645	38,6	1.176	1.655	39,2	1.187	1.617	40,0	1.007	1.799	40,6	988
92 02	6.431	6.624	25,5	1.938	6.557	29,1	1.510	6.460	29,7	1.482	6.057	35,7	1.035	5.879	45,9	685
92 03	969	937	34,5	1.130	1.012	33,8	1.130	992	37,4	931	1.078	40,8	760	814	57,1	380
92 04	4.609	4.715	23,5	1.947	5.085	23,9	1.928	4.443	24,8	1.643	4.928	27,9	1.529	4.746	34,0	1.244
92 05	7.646	7.492	26,3	2.318	7.127	30,1	1.757	7.472	29,2	1.881	7.163	35,0	1.398	7.927	40,8	1.073
93 01	18.914	18.869	28,2	2.337	18.535	26,4	2.251	21.668	31,8	1.805	20.622	29,4	1.976	18.138	34,5	1.386
93 02	3.654	3.761	24,8	1.558	3.683	24,8	1.587	3.816	25,3	1.491	3.611	29,8	1.206	3.628	35,4	997
93 03	184	184	58,8	541	184	58,8	541	184	58,8	541	184	58,8	541	185	58,7	541
93 04	7.598	7.152	29,0	1.909	7.895	30,3	1.786	7.668	35,6	1.244	7.415	43,0	864	8.360	51,7	664
93 71	14.364	14.619	25,0	2.593	14.307	29,2	1.976	13.454	31,5	1.843	15.453	34,7	1.596	14.465	42,6	1.111

Attachment B

Detailed Comparison of Sample Levels and Costs for Samples 1, 3, and 5

28/07/2000

			Sample 1									Sample 3									Sample 5								
		SE-96	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
11	01	11.528	11.694	31,5	10	13	15	98	108	131	3.002	10.847	36,0	3	7	9	32	45	68	1.491	11.118	40,0	3	4	6	23	45	68	1.035
11	02	10.501	10.180	26,1	3	12	13	22	70	73	2.469	10.215	30,5	1	9	10	3	63	66	1.966	11.129	41,4	1	5	6	3	38	41	1.121
11	03	38.170	35.497	20,2	4	21	25	32	165	197	4.759	36.353	23,9	4	15	19	32	112	144	3.344	36.333	27,4	4	12	16	32	108	140	2.850
11	04	34.782	36.208	17,1	6	24	29	95	315	391	6.640	34.939	20,5	5	16	21	76	214	290	4.465	36.487	23,7	5	12	17	76	171	247	3.448
11	05	79.633	82.936	12,7	13	51	58	157	539	612	12.872	82.085	14,4	9	40	47	102	371	444	9.604	78.003	17,3	9	28	35	73	268	341	6.802
11	06	25.103	24.552	20,3	3	20	22	24	158	181	4.636	26.889	28,3	2	11	13	13	95	108	2.574	19.805	46,9	2	5	7	13	41	54	1.149
11	07	42.769	41.765	19,8	7	22	28	38	156	186	4.926	39.399	24,4	6	15	21	32	114	146	3.363	46.374	28,1	6	11	17	32	92	124	2.546
11	08	35.294	34.593	20,8	8	19	23	77	161	187	4.417	36.752	24,3	6	16	20	47	152	178	3.876	35.200	26,3	4	12	16	28	110	136	2.869
11	09	104.561	103.736	11,5	31	67	78	403	773	919	17.527	103.508	12,8	17	51	62	231	581	727	13.271	103.015	14,5	15	39	50	146	420	566	9.918
11	10	47.837	51.304	17,6	12	37	42	140	351	388	8.958	47.632	19,8	6	24	29	56	246	295	6.004	46.844	22,2	5	17	22	47	136	183	3.876
11	11	73.600	73.983	12,9	15	51	58	152	488	522	12.198	74.757	14,4	12	40	47	112	418	472	10.061	72.178	17,0	8	30	37	54	287	341	7.286
11	71	32.221	32.647	19,6	2	20	22	25	274	299	5.643	32.627	23,6	2	15	17	25	216	241	4.332	36.977	35,8	2	8	10	25	109	134	2.251
11	72	2.413	2.405	22,9	0	8	8	0	74	74	1.919	2.437	23,8	0	8	8	0	79	79	1.966	2.205	39,3	0	5	5	0	33	33	1.073
12	01	58.947	59.059	12,2	22	49	65	214	429	575	11.523	60.015	13,9	17	35	51	160	321	467	8.369	56.684	17,0	16	23	39	146	178	324	5.187
12	02	34.474	34.205	20,2	11	16	27	82	131	189	3.960	33.015	24,6	10	12	21	71	74	132	2.527	34.766	28,8	9	8	17	58	57	115	1.767
12	03	70.164	69.473	12,7	27	45	66	239	373	540	10.383	69.858	14,2	24	34	55	208	308	475	8.094	67.245	17,1	21	24	45	167	160	327	5.168
12	04	28.969	27.423	20,4	7	19	25	56	146	191	4.275	27.663	22,2	7	16	22	56	134	179	3.705	26.450	26,1	6	12	18	45	88	133	2.660
12	05	44.279	42.606	16,4	18	30	43	134	227	310	6.716	41.627	19,7	14	19	32	97	140	223	4.218	42.167	21,3	14	16	29	83	122	205	3.591
12	06	25.218	24.773	20,6	12	19	27	102	144	202	4.258	25.713	28,6	8	10	18	58	86	144	2.337	26.626	34,4	8	7	15	58	60	118	1.634
12	07	58.102	58.608	12,0	17	48	64	127	417	534	11.257	59.933	13,9	16	37	53	117	340	457	8.854	56.937	17,2	16	25	41	117	195	312	5.652
12	08	85.507	86.356	12,0	20	53	73	167	573	740	13.499	84.524	14,1	20	40	60	187	404	571	9.916	85.942	15,0	20	35	55	167	396	563	9.092
12	09	52.813	51.940	12,2	16	52	67	129	402	520	11.723	52.654	14,0	16	40	55	129	319	437	9.110	52.106	17,1	15	29	44	118	222	340	6.517
12	10	16.998	18.610	25,0	4	12	16	29	102	131	2.793	15.658	31,3	4	8	12	29	54	83	1.729	17.713	45,7	4	4	8	29	38	67	989

			Sample 1									Sample 3									Sample 5								
		SE-96	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
12	11	21.316	21.227	20,4	7	18	25	41	150	191	4.161	21.194	25,3	7	13	20	41	107	148	2.992	22.396	26,0	7	12	19	41	109	150	2.859
12	12	174.119	177.291	8,0	44	109	151	426	1.017	1.443	26.229	173.012	10,0	42	73	115	401	633	1.034	17.109	176.124	12,2	42	51	93	401	468	869	12.198
12	13	83.988	85.021	12,1	23	49	71	195	426	607	11.495	83.243	14,0	23	37	59	195	312	493	8.568	85.478	17,2	22	26	48	181	228	409	6.118
12	71	12.332	11.847	25,6	1	11	12	11	139	150	2.992	13.952	31,5	1	9	10	11	123	134	2.538	11.482	46,2	1	4	5	11	59	70	1.168
12	72	14.361	14.856	24,7	3	12	15	31	137	168	3.125	14.986	32,7	3	8	11	31	94	125	2.109	16.056	35,4	3	7	10	31	79	110	1.814
12	73	26.202	26.723	19,8	5	19	24	37	225	262	5.025	27.415	21,4	5	17	22	37	183	220	4.322	27.404	30,9	5	9	14	37	117	154	2.479
12	74	15.706	16.432	21,2	4	12	16	45	139	184	3.144	14.673	31,5	4	8	12	45	84	129	2.014	14.455	37,2	4	6	10	45	67	112	1.540
12	75	228.074	228.974	8,0	45	118	163	554	1.389	1.943	31.131	229.715	10,0	45	81	126	554	1.029	1.583	22.087	238.309	12,0	45	59	104	554	758	1.312	16.169
12	76	27.788	27.978	20,2	6	18	24	79	237	316	4.987	28.032	29,7	6	10	16	79	128	207	2.736	32.033	37,9	6	6	12	79	72	151	1.595
13	01	1.281	1.286	56,6	3	4	4	15	18	18	779	1.250	56,7	3	4	4	15	16	16	760	1.176	60,9	1	3	3	0	13	13	579
13	02	38.984	39.628	16,1	5	38	42	42	288	320	8.512	38.325	19,2	4	24	28	32	220	252	5.738	39.966	21,1	4	20	24	32	193	225	4.873
13	03	40.384	40.233	16,5	7	33	38	48	235	267	7.248	39.715	19,8	6	22	27	37	150	182	4.769	43.657	21,3	6	19	24	32	142	174	4.237
13	04	38.256	39.061	20,1	4	21	24	39	214	235	5.225	40.196	23,5	3	16	19	21	191	212	4.246	39.655	26,3	3	13	16	21	138	159	3.287
13	05	56.283	56.879	12,2	14	61	56	137	497	542	12.473	57.139	14,1	6	41	46	58	387	442	10.003	57.075	16,9	5	30	35	45	283	328	7.248
13	06	75.253	76.124	12,1	9	54	60	83	588	642	13.794	74.506	14,1	6	41	47	54	414	468	10.165	72.819	17,4	6	29	35	54	269	323	6.963
13	07	68.481	65.801	12,5	10	54	61	141	618	706	14.079	67.395	14,1	10	42	49	141	557	645	11.675	65.274	17,1	7	31	38	88	385	473	8.369
13	08	52.576	50.791	12,7	14	54	58	184	677	626	13.669	50.972	14,2	7	42	46	88	478	527	10.925	51.456	16,1	5	33	37	49	366	415	8.493
13	09	61.508	61.041	12,7	18	49	56	298	425	507	11.465	60.644	14,3	13	39	46	160	324	405	9.008	59.785	17,2	7	28	35	82	225	307	6.393
13	71	93.201	90.478	12,1	8	56	64	97	603	700	14.240	94.687	14,0	8	44	52	97	504	601	11.476	96.301	17,0	8	31	39	97	413	510	8.635
13	72	9.830	9.843	25,3	0	11	11	0	144	144	3.040	9.513	31,4	0	8	8	0	111	111	2.270	8.571	35,7	0	6	6	0	90	90	1.767
13	73	8.913	7.050	24,3	0	11	11	0	131	131	2.918	6.338	38,3	0	6	6	0	55	55	1.434	6.340	45,9	0	4	4	0	47	47	1.054
13	74	7.852	8.114	24,1	1	10	11	6	136	142	2.812	7.826	36,3	1	6	7	6	67	73	1.548	6.105	52,8	1	3	4	6	38	44	817
13	75	23.003	24.802	25,7	2	13	15	25	201	226	3.895	22.563	28,0	2	10	12	25	156	181	3.002	24.427	32,9	2	8	10	25	119	144	2.346
13	76	18.539	18.834	25,0	1	13	14	16	160	176	3.495	21.495	31,6	1	10	11	16	126	142	2.717	17.229	38,2	1	7	8	16	81	97	1.833
14	01	23.413	23.336	21,7	4	18	22	26	181	197	4.607	22.195	25,0	3	14	17	16	106	122	3.135	25.224	27,1	3	12	15	16	113	129	2.897

			Sample 1									Sample 3									Sample 5								
		SE-86	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
14	02	30.761	30.969	20,4	5	19	22	40	234	258	5.111	29.162	24,2	5	14	17	40	144	188	3.496	27.316	35,8	3	7	10	24	59	83	1.624
14	03	35.101	35.158	20,0	6	20	26	36	185	221	4.797	32.052	24,4	6	14	20	38	116	152	3.230	36.379	25,7	6	12	18	36	125	161	3.011
14	04	6.624	6.944	26,5	5	11	12	32	72	78	2.358	6.640	31,0	2	8	9	13	55	61	1.738	7.079	35,0	1	6	7	6	52	58	1.406
14	05	16.989	16.798	25,5	4	10	13	76	110	187	2.565	18.267	29,2	3	8	11	57	79	136	1.966	16.921	36,1	3	5	8	57	55	112	1.282
14	06	27.704	27.875	20,3	4	21	25	24	181	205	4.911	25.461	29,5	4	10	14	24	94	118	2.413	25.634	35,8	4	7	11	24	55	79	1.586
14	07	14.628	14.331	25,3	2	13	14	26	102	119	2.945	13.664	29,4	1	10	11	17	63	80	2.118	14.675	35,9	1	7	8	17	64	81	1.572
14	08	37.246	36.515	22,3	4	18	22	69	223	250	5.006	36.955	25,1	4	15	18	59	170	197	3.895	34.597	30,7	4	11	14	27	134	161	2.945
14	09	24.983	25.329	20,8	6	19	21	67	197	211	4.750	24.726	22,8	5	16	18	57	149	163	3.847	25.618	30,3	2	9	11	14	97	111	2.288
14	10	26.263	26.924	19,8	3	21	24	17	215	232	5.234	28.949	24,0	3	15	18	17	172	189	3.914	28.678	34,8	3	8	11	17	84	101	2.014
14	11	15.339	15.271	25,3	1	13	14	7	124	131	3.154	14.764	33,1	1	9	10	7	86	93	2.185	13.709	37,5	1	7	8	7	50	57	1.539
14	12	4.132	3.887	25,3	0	10	10	0	96	96	2.432	4.126	35,8	0	6	6	0	68	68	1.558	3.481	56,3	0	3	3	0	19	19	638
14	71	72.710	73.152	11,6	7	49	56	89	716	805	14.250	72.439	14,0	7	39	46	89	556	645	11.210	73.199	16,9	7	29	36	89	411	500	6.312
14	72	34.358	33.622	20,0	3	18	21	24	333	357	5.899	34.868	22,3	3	15	18	24	382	406	5.909	34.893	26,6	3	13	16	24	285	309	4.683
14	73	22.220	21.660	25,4	3	12	15	40	130	170	3.058	23.581	29,4	3	9	12	40	95	135	2.270	22.254	36,0	3	6	9	40	57	97	1.453
15	01	32.233	31.649	20,2	14	19	33	100	166	266	4.466	30.688	21,6	14	16	30	100	167	267	4.018	33.123	23,7	14	14	28	100	158	258	3.610
15	02	26.665	26.638	20,2	12	18	28	94	179	242	4.436	27.364	27,9	10	9	19	63	61	124	1.947	25.445	35,8	10	5	15	63	51	114	1.244
15	03	26.935	26.181	20,0	11	14	25	112	157	269	3.619	25.316	24,5	11	9	20	112	111	223	2.422	27.398	28,5	11	7	18	112	83	195	1.652
15	04	16.649	15.324	25,0	7	9	16	80	101	181	2.327	17.267	30,1	7	7	14	80	79	159	1.814	17.683	35,8	7	5	12	80	62	142	1.349
15	05	18.232	18.811	24,6	7	13	19	61	103	164	2.954	15.909	26,7	6	9	15	51	71	122	2.042	17.135	34,0	6	6	12	51	56	107	1.444
15	06	19.848	20.704	25,1	10	10	20	75	115	190	2.612	23.061	29,5	10	8	18	75	94	169	2.109	21.036	36,1	10	5	15	75	55	130	1.282
15	07	16.528	17.188	25,3	7	9	16	79	125	204	2.555	15.650	30,6	7	6	13	79	70	149	1.577	19.156	34,2	7	5	12	79	54	133	1.273
15	08	14.526	14.867	25,6	9	9	17	63	94	140	2.261	15.442	29,3	9	7	15	63	76	122	1.786	15.223	35,2	8	5	13	46	44	90	1.178
15	09	26.883	26.525	20,2	11	13	24	153	126	279	3.173	23.454	21,9	11	10	21	153	87	240	2.346	23.715	28,4	11	6	17	153	50	203	1.367
15	71	48.668	45.726	15,6	18	31	49	180	315	495	7.704	49.290	19,0	18	21	39	180	233	413	5.405	48.811	21,7	18	18	36	180	185	365	4.493
16	01	92.913	93.400	12,1	13	56	69	96	541	611	13.651	93.971	13,9	11	44	54	77	437	507	10.839	95.215	17,1	10	30	40	70	320	390	7.609

			Sample 1									Sample 3									Sample 5								
		SE-96	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
16	02	93.935	96.136	11,7	10	57	64	123	719	790	15.494	95.076	12,9	8	49	56	89	540	620	12.663	97.914	14,1	8	38	46	71	505	576	10.725
16	03	59.345	59.914	12,3	13	51	57	126	514	568	12.635	58.610	14,1	8	40	46	84	387	441	9.756	62.304	17,0	6	29	35	54	289	343	7.153
16	04	61.256	62.001	13,6	10	51	57	94	462	513	12.141	60.728	15,3	8	39	45	74	336	387	9.120	62.644	17,5	7	29	35	51	285	336	7.115
16	05	75.199	76.272	12,1	6	51	55	51	679	708	14.202	76.933	14,0	5	40	44	38	542	571	11.229	73.386	17,2	4	28	32	29	356	385	7.638
16	06	83.517	87.183	12,1	14	58	68	121	476	552	13.338	87.661	14,1	10	44	54	76	356	432	10.070	86.675	16,9	10	31	41	76	255	331	7.134
16	07	56.303	55.380	12,6	12	50	55	178	591	646	13.214	54.314	14,5	9	40	46	136	439	494	10.250	55.964	17,0	5	30	35	55	360	415	7.980
16	08	18.758	18.421	24,9	1	14	15	7	119	126	3.258	16.705	30,0	1	10	11	7	90	97	2.375	19.000	35,8	1	7	8	7	72	79	1.746
16	71	162.911	162.832	8,0	10	115	125	170	1.510	1.680	31.825	163.978	10,0	10	80	90	170	1.077	1,247	22.391	159.044	12,0	10	60	70	170	757	927	16.311
16	72	18.005	18.394	24,8	0	13	13	0	156	156	3.458	19.522	29,4	0	10	10	0	112	112	2.584	17.446	41,1	0	6	6	0	70	70	1.577
17	01	27.774	26.762	20,0	11	20	31	59	143	202	4.398	26.271	24,0	11	14	25	59	105	164	3.125	25.668	29,8	11	9	20	59	68	127	2.014
17	02	37.896	36.520	20,3	10	18	28	86	183	269	4.474	34.515	28,8	10	9	19	86	94	180	2.261	41.288	33,9	10	7	17	86	63	149	1.662
17	03	34.383	35.169	19,6	11	19	30	82	142	224	4.237	32.851	24,8	11	12	23	82	91	173	2.688	31.678	29,9	11	8	19	82	55	137	1.738
17	71	32.804	31.738	19,8	8	18	26	96	275	371	5.348	33.222	23,9	8	14	22	96	211	307	4.132	37.314	33,5	8	8	16	96	109	205	2.251
18	01	32.864	34.192	21,3	5	20	22	66	168	194	4.636	35.565	28,3	3	11	13	41	97	123	2.598	31.753	35,3	2	7	9	26	58	84	1.615
18	02	80.629	82.680	12,3	10	50	59	124	540	627	12.730	81.987	14,2	9	39	47	108	409	496	9.613	79.932	16,3	9	31	39	87	340	427	7.942
18	03	113.628	116.472	10,7	18	74	81	195	693	738	17.831	117.757	12,2	12	57	64	120	582	602	14.003	113.811	14,2	8	43	50	40	426	466	10.593
18	04	122.498	121.573	10,2	24	74	89	323	916	1.096	19.950	121.875	11,9	16	57	72	197	679	959	15.114	123.843	13,0	16	48	63	180	576	756	12.768
18	05	134.036	135.928	10,2	14	79	93	169	911	973	20.662	137.120	11,8	13	62	69	158	753	815	16.577	135.500	13,0	8	51	58	62	590	652	13.357
18	06	45.830	45.552	16,1	8	31	35	78	325	353	7.799	49.417	19,9	5	22	26	41	216	244	5.336	45.949	24,1	4	15	19	28	165	193	3.847
18	07	25.966	25.604	20,7	7	20	24	46	127	144	4.246	25.303	29,4	4	10	14	17	79	96	2.270	31.610	38,8	4	6	10	17	49	66	1.377
18	08	43.020	46.318	16,4	8	34	37	78	282	320	7.847	43.780	19,5	4	22	25	42	173	211	4.987	42.634	21,9	3	18	21	38	137	175	4.037
18	71	99.753	98.221	12,1	6	56	62	79	619	698	16.292	100.170	14,0	6	44	50	79	591	670	12.302	98.211	17,1	6	31	37	79	398	477	8.493
18	72	19.318	20.010	24,2	0	14	14	0	196	196	3.990	18.113	30,4	0	11	11	0	142	142	3.021	16.828	33,8	0	8	8	0	102	102	2.185
31	71	286.701	284.006	6,0	36	199	233	424	2.539	2.931	54.368	285.094	7,9	34	126	160	392	1.675	2.067	35.084	295.587	10,0	34	86	120	392	1.121	1.513	23.721
31	72	335.118	337.912	6,0	34	205	238	409	3.062	3.421	60.249	337.088	8,0	34	128	161	409	1.914	2.273	37.639	335.884	10,1	33	88	121	359	1.263	1.622	25.374

		SE-96	Sample 1									Sample 3								Sample 5								
			Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
31	73	219.879	220.591	8,0	27	118	145	280	1.321	1.601	30.485	215.881	10,0	27	81	108	280	824	1.104	20.140	216.714	12,0	27	59	86	280	652	932	15.162
31	74	349.321	348.156	8,0	38	200	238	456	2.631	3.087	56.394	353.157	8,0	38	129	167	456	1.671	2.127	35.482	340.832	10,1	38	86	124	456	1.050	1.506	23.047
31	75	232.751	236.310	8,0	23	120	143	250	2.027	2.277	37.496	232.067	10,0	23	83	106	250	1.462	1.712	26.505	229.992	11,9	23	62	85	250	931	1.181	18.288
32	01	135.284	136.933	16,6	10	58	63	131	655	702	15.038	137.533	17,6	8	43	48	96	438	535	11.172	135.514	19,2	6	32	37	47	364	411	8.322
32	02	89.571	93.139	12,3	8	55	59	91	534	569	13.433	92.709	14,1	6	42	46	70	420	455	10.374	92.036	17,0	4	30	34	35	288	323	7.296
32	03	336.405	340.597	8,1	20	223	237	207	2.567	2.690	58.262	341.325	8,0	16	138	152	153	1.621	1.744	36.375	338.050	10,0	14	91	105	123	1.094	1.217	24.225
32	04	249.714	255.145	6,4	36	196	207	450	2.166	2.265	50.389	254.691	8,1	15	124	139	166	1.376	1.475	31.920	261.697	9,9	13	89	100	99	1.043	1.142	23.436
32	05	200.092	196.316	8,1	13	135	143	128	1.391	1.450	33.734	198.363	10,0	10	94	102	91	936	1.027	23.484	197.302	12,0	8	67	75	59	713	772	16.967
32	06	428.092	427.989	6,1	28	232	248	342	2.875	2.851	60.678	426.122	8,0	19	147	166	215	1.705	1.861	38.541	430.935	9,8	18	100	116	176	1.205	1.381	26.647
32	07	207.445	208.326	8,1	12	130	139	140	1.460	1.580	33.820	205.018	9,8	12	94	103	140	1.037	1.137	24.139	201.590	11,0	9	75	84	100	841	941	19.389
32	08	287.085	287.501	6,1	35	210	223	462	2.620	2.765	56.810	292.547	8,0	13	131	144	145	1.755	1.900	36.584	284.634	9,9	13	91	104	145	1.198	1.343	25.213
32	09	236.485	236.755	8,1	16	132	143	208	1.548	1.677	34.741	238.250	9,9	12	94	105	141	1.059	1.191	24.348	239.943	10,9	11	80	91	132	921	1.053	20.309
32	10	100.453	97.594	12,0	4	60	64	36	591	627	14.734	96.099	14,0	4	46	50	38	423	459	11.010	101.798	17,0	4	32	36	36	309	345	7.799
32	11	251.275	254.498	6,5	34	196	208	479	2.384	2.490	52.250	256.682	8,1	19	124	134	265	1.629	1.755	34.328	253.454	9,8	11	90	100	126	1.054	1.180	23.693
32	12	161.695	160.581	8,2	19	119	126	238	1.439	1.527	31.758	161.435	10,0	9	84	91	120	1.029	1.117	22.543	164.534	12,0	7	61	68	88	802	890	16.891
32	13	98.533	98.551	11,7	7	65	71	58	647	694	16.026	101.408	12,9	6	55	61	47	556	603	13.842	103.630	14,0	6	48	54	47	469	516	11.751
32	14	210.575	212.185	8,1	12	127	134	148	1.503	1.588	33.582	210.967	10,0	8	86	95	97	1.024	1.109	23.104	220.037	12,0	7	63	70	85	817	902	17.337
32	15	189.458	196.820	8,2	13	118	124	156	1.362	1.437	30.875	197.386	10,0	9	82	88	101	974	1.049	21.717	198.304	11,9	7	80	86	75	681	756	15.589
32	16	78.795	80.115	12,1	11	60	64	130	605	653	14.867	80.748	13,9	5	44	48	60	445	493	10.915	79.336	17,1	4	29	33	48	301	349	7.267
32	17	203.368	208.311	8,1	13	122	128	152	1.326	1.382	31.141	209.770	10,0	7	84	90	71	936	992	21.660	208.800	12,0	6	80	86	58	653	709	15.323
32	18	253.096	253.665	6,7	38	183	190	514	2.583	2.627	51.679	254.314	8,3	20	123	130	309	1.774	1.868	35.549	250.334	10,1	9	86	93	94	1.126	1.220	23.769
32	19	240.662	240.204	8,0	14	132	140	197	1.677	1.758	35.996	247.168	10,0	6	91	99	79	1.206	1.285	25.289	248.909	11,8	8	67	75	79	852	931	18.278
32	20	179.486	179.450	8,3	19	119	125	258	1.500	1.564	32.339	180.282	10,1	10	83	89	126	1.090	1.154	22.971	179.440	12,1	8	60	68	64	776	840	16.492
32	71	42.380	41.958	16,6	2	30	31	40	351	355	7.694	44.182	20,2	1	22	23	4	262	266	5.833	44.388	24,0	1	16	17	4	222	226	4.541
32	72	19.402	19.635	24,2	1	14	15	8	127	135	3.334	16.110	29,9	1	10	11	8	113	121	2.593	19.619	34,1	1	8	9	8	83	91	2.004

			Sample 1									Sample 3									Sample 5								
		SE-96	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
32	73	302.706	303.544	6,0	10	226	236	85	2.456	2.541	57.634	308.536	8,0	10	140	150	85	1.569	1.654	36.185	303.674	10,0	10	93	103	85	999	1.084	23.626
32	74	40.212	39.876	18,2	2	32	34	17	347	364	8.160	42.859	20,5	2	22	24	17	228	245	5.510	42.302	23,2	2	17	19	17	192	209	4.408
32	75	121.179	121.617	9,9	3	81	84	26	989	1.015	21.707	121.556	12,1	3	60	63	26	701	727	16.779	123.339	14,1	3	46	49	26	576	602	12.464
32	76	92.996	93.906	11,0	4	55	59	44	954	998	18.473	95.039	14,0	4	43	47	44	696	740	13.148	94.756	16,9	4	31	35	44	480	524	9.272
32	77	98.480	98.085	12,0	2	56	58	26	828	854	16.378	101.736	14,1	2	44	46	26	624	650	12.616	99.006	17,0	2	31	33	28	444	470	8.930
32	78	29.316	30.301	19,9	2	20	22	32	282	314	5.719	30.127	30,1	2	11	13	32	143	175	3.030	27.884	33,1	2	8	10	32	127	159	2.422
33	01	215.034	216.422	8,0	21	124	139	260	1.585	1.760	33.905	215.185	10,0	15	83	98	175	1.044	1.219	22.534	210.605	11,8	15	61	76	175	766	941	16.549
33	02	226.613	223.497	8,1	21	123	137	309	1.591	1.701	33.715	215.935	9,8	17	89	106	243	1.127	1.337	24.234	222.005	10,8	14	76	90	210	915	1.125	20.254
33	03	137.297	139.253	10,4	24	74	84	293	1.029	1.155	21.023	138.528	11,8	14	60	70	178	805	931	16.767	140.077	13,0	11	46	56	126	651	777	13.176
33	04	124.960	123.684	10,1	14	60	88	253	1.160	1.307	23.180	123.589	11,0	9	68	76	166	983	1.130	19.674	127.075	12,9	8	53	61	147	768	915	15.352
33	05	194.736	197.695	8,6	31	111	122	461	1.578	1.735	31.883	195.955	10,2	16	73	89	246	1.104	1.261	22.344	191.724	12,2	12	57	69	157	782	939	16.093
33	06	108.166	107.425	10,5	22	73	81	291	855	940	19.218	109.583	12,0	13	56	64	154	687	772	15.038	108.787	14,0	8	42	50	85	554	639	11.647
33	07	138.276	139.035	12,1	15	64	72	236	948	1.073	18.734	138.150	13,7	13	48	56	201	724	849	14.174	135.507	15,2	12	40	48	125	558	683	11.381
33	08	211.444	210.822	9,6	27	105	118	465	1.583	1.793	30.988	213.354	10,9	20	76	89	356	1.138	1.348	22.383	211.797	12,4	16	59	72	210	670	1.080	17.233
33	09	111.372	112.355	10,4	16	77	84	162	784	846	19.152	110.793	11,8	15	62	69	153	637	701	15.475	112.659	13,0	9	50	57	64	491	555	12.264
33	10	188.803	191.034	9,1	22	120	134	252	1.415	1.555	31.682	189.710	10,1	15	82	96	159	946	1.086	21.451	188.030	11,8	15	61	75	140	753	893	16.425
33	11	104.588	103.964	10,1	9	76	82	108	962	1.037	20.691	105.861	12,0	6	55	62	75	769	844	15.817	107.080	13,9	6	43	49	75	599	674	12.228
33	12	135.178	136.632	10,1	14	78	88	163	929	1.034	20.681	138.101	11,8	10	61	71	105	751	856	16.406	132.947	13,1	10	49	59	105	529	634	12.473
33	13	133.599	136.852	10,1	14	72	79	263	1.106	1.234	21.441	134.208	12,0	10	56	63	157	876	1.005	16.834	135.510	14,2	8	42	49	129	579	708	11.884
33	14	133.006	134.971	10,3	13	77	84	181	1.036	1.120	21.546	135.731	12,0	8	58	65	102	757	841	16.007	132.460	14,1	8	43	50	84	538	622	11.647
33	15	144.877	146.493	10,4	16	76	85	215	832	930	19.456	150.618	12,1	12	58	67	130	615	713	14.668	152.512	13,9	10	44	53	98	474	572	11.191
33	16	74.617	74.269	12,2	9	56	62	86	514	565	13.395	73.968	14,0	7	42	48	84	395	446	10.136	71.303	17,3	6	30	36	51	254	305	6.973
33	17	69.793	70.043	12,8	11	52	58	120	548	584	13.110	71.236	14,3	8	42	46	90	437	473	10.535	75.362	17,7	6	31	35	36	320	356	7.752
33	18	164.330	163.997	8,6	35	113	123	281	1.271	1.359	29.250	161.036	10,2	16	80	90	169	885	973	20.567	162.952	12,1	12	59	69	88	655	743	15.190
33	19	103.054	102.958	10,4	9	74	80	112	901	871	18.857	103.416	12,2	7	55	61	81	620	690	14.250	105.043	14,0	7	43	49	70	475	545	11.048

		SE-96	Sample 1									Sample 3									Sample 5									
			Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
33	20	181.685	183.078	9,7	20	105	117	273	1.562	1.712	30.799	181.662	10,4	17	75	87	234	1.066	1.216	21.527	182.283	11,2	15	63	75	150	977	1.127	18.857
33	21	107.810	107.528	10,1	10	78	88	142	899	987	20.396	109.703	12,0	6	58	63	95	711	799	15.570	109.367	13,8	5	45	50	88	561	649	12.160
33	22	116.497	116.274	10,5	18	74	92	228	989	1.077	20.643	115.239	12,1	11	58	66	120	726	816	15.732	114.043	14,0	8	44	52	88	531	619	11.732
33	23	250.537	251.325	10,7	52	98	114	929	1.505	1.810	29.193	252.975	11,4	36	74	90	685	1.176	1.481	22.420	251.655	11,9	30	61	77	305	982	1.287	18.801
33	24	102.050	101.860	10,2	14	78	85	181	868	945	20.102	100.503	12,1	8	58	65	92	610	687	14.611	103.381	14,0	7	44	51	77	499	576	11.428
33	25	104.504	108.196	10,6	13	71	77	203	993	1.076	20.226	107.052	12,3	9	54	60	136	733	816	15.171	107.143	14,1	6	41	47	83	629	712	12.207
33	26	123.190	124.287	10,5	16	71	78	292	975	1.105	20.054	124.846	11,3	13	61	68	240	897	1.027	17.793	125.254	13,1	9	47	54	130	690	820	13.699
33	27	120.874	118.653	10,0	10	83	90	109	889	972	21.061	118.533	12,0	7	62	69	83	669	752	15.779	118.212	14,0	7	47	54	93	515	598	12.036
33	28	173.042	173.161	8,1	14	116	126	169	1.472	1.600	31.616	173.629	9,9	11	81	91	137	1.024	1.152	22.040	169.563	12,0	10	58	68	128	696	824	15.428
33	29	165.550	168.631	8,2	20	121	131	241	1.308	1.414	30.818	167.643	10,1	13	83	93	147	887	993	21.042	168.524	11,9	11	61	71	106	657	763	15.513
33	71	23.728	21.562	20,9	1	18	19	17	217	234	4.797	22.027	24,7	1	14	15	17	158	175	3.629	20.396	36,0	1	7	8	17	69	86	1.719
33	72	103.486	103.481	10,0	6	78	84	58	875	933	20.168	104.122	12,0	6	59	65	58	657	715	15.209	102.828	14,0	6	45	51	58	489	547	11.485
33	73	25.889	26.809	19,9	1	20	21	5	247	252	5.386	26.771	24,1	1	15	16	5	146	151	3.687	28.368	34,0	1	8	9	5	100	105	2.166
33	74	180.161	179.671	8,0	10	129	139	69	1.349	1.418	32.423	180.740	10,0	10	89	99	69	975	1.044	22.790	180.609	12,0	10	65	75	69	667	736	16.216
33	75	47.259	47.833	19,4	3	23	26	27	276	303	6.118	46.385	22,0	3	19	22	27	233	260	5.101	47.873	25,5	3	14	17	27	181	208	3.847
33	76	38.681	38.346	19,9	2	21	23	22	239	261	5.462	38.125	24,3	2	15	17	22	150	172	3.705	35.613	36,5	2	7	9	22	78	100	1.805
34	01	61.915	62.645	12,6	14	46	53	191	605	718	12.739	62.097	14,2	11	38	45	156	491	604	10.440	63.190	15,1	8	33	40	113	402	515	8.835
34	02	161.699	160.220	8,3	23	110	128	349	1.713	1.971	32.993	161.031	10,1	22	78	96	333	1.202	1.460	23.275	156.134	12,1	21	57	75	258	898	1.156	17.195
34	03	100.028	100.870	10,5	19	70	83	185	721	851	17.489	100.604	12,2	16	51	64	157	524	654	12.730	99.857	14,1	15	39	52	130	386	516	9.585
34	04	145.753	144.305	10,1	19	77	94	267	934	1.164	20.577	147.206	12,0	17	56	73	230	723	953	15.380	146.741	14,0	17	42	59	230	538	768	11.495
34	71	106.817	108.712	10,0	12	74	86	159	916	1.075	19.960	109.524	11,9	12	55	67	159	733	892	15.323	110.281	13,9	12	43	55	159	568	727	11.932
35	01	69.600	70.484	13,1	10	43	53	98	554	605	12.559	68.796	14,5	9	40	45	92	484	535	10.078	70.096	15,5	7	34	39	51	404	455	9.006
35	02	120.776	122.709	10,3	17	74	92	183	942	1.022	20.197	125.687	12,1	11	59	67	121	838	889	16.653	122.651	13,1	9	47	55	80	610	690	12.939
35	03	104.915	106.268	10,4	15	58	75	222	982	1.076	19.665	104.903	12,4	10	50	57	138	658	752	13.851	106.517	14,1	8	39	46	94	545	639	11.165
35	04	193.437	188.809	8,4	20	109	120	297	1.661	1.835	32.347	188.951	10,2	16	77	93	246	1.134	1.308	22.477	187.183	12,0	12	57	68	174	895	1.069	17.166

			Sample 1									Sample 3									Sample 5								
		SE-96	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
35	05	184.565	191.385	8,5	17	111	120	264	1.575	1.674	31.834	189.572	10,2	12	77	86	142	1.003	1.182	21.992	180.617	12,3	11	67	66	99	736	835	15.656
35	06	195.233	200.774	8,4	18	116	127	174	1.315	1.408	30.124	199.067	10,2	17	80	91	166	903	996	20.739	204.733	12,0	14	59	70	93	675	768	15.389
35	07	260.491	296.756	6,1	30	208	223	363	2.483	2.657	54.900	292.146	8,0	18	128	146	191	1.501	1.675	33.715	302.665	10,0	17	84	101	174	1.069	1.243	22.923
35	08	126.326	129.229	10,0	13	79	86	143	945	1.013	20.085	126.301	12,0	7	58	65	68	669	737	15.171	132.512	14,0	7	44	51	68	512	580	11.552
35	09	259.098	267.521	6,2	32	205	221	358	2.357	2.517	53.551	269.021	8,0	21	130	146	224	1.538	1.698	34.371	257.307	10,0	17	87	103	160	926	1.086	22.021
35	10	242.998	247.656	8,2	20	119	132	271	1.606	1.790	33.345	250.683	9,6	16	89	102	217	1.223	1.407	25.146	245.309	11,0	14	68	81	184	846	1.030	18.373
35	11	92.892	97.360	11,9	6	58	64	67	619	686	14.696	95.723	14,5	6	42	48	67	420	487	10.374	94.310	16,1	6	32	38	67	334	401	9.037
35	12	87.014	89.081	12,0	7	57	62	60	650	695	14.839	86.460	13,9	6	44	49	53	456	509	11.069	85.574	17,0	5	32	37	45	330	375	7.999
35	13	126.353	128.993	11,4	13	72	79	138	756	808	18.116	127.582	13,0	12	53	60	120	568	641	13.642	125.258	14,7	10	41	48	58	462	515	10.821
35	14	185.946	190.483	8,1	14	120	130	163	1.526	1.637	32.727	190.847	10,0	11	81	91	121	1.046	1.158	22.249	190.385	11,9	10	58	68	112	718	830	15.637
35	15	207.992	210.252	8,1	14	121	131	175	1.720	1.857	34.732	210.439	10,1	12	85	95	153	1.142	1.278	23.769	214.415	12,0	11	62	72	137	873	1.010	17.717
35	16	131.062	132.705	10,0	12	77	89	167	925	1.046	20.491	133.336	12,0	9	56	65	121	665	786	14.829	130.249	14,1	9	41	50	121	487	608	10.658
35	17	165.316	173.207	8,3	12	113	121	171	1.433	1.555	30.789	172.302	10,1	10	79	87	151	966	1.068	21.185	169.209	12,0	9	58	66	122	676	798	15.238
35	18	166.907	161.290	18,0	16	93	101	184	1.062	1.155	24.225	162.452	16,7	11	68	74	130	810	903	17.727	161.581	19,7	9	49	57	93	610	703	13.243
35	19	76.141	78.605	12,3	7	55	59	74	591	630	13.974	77.168	14,2	6	41	45	62	428	467	10.298	77.408	16,9	4	30	34	39	297	336	7.381
35	20	104.181	104.381	10,3	14	74	81	211	1.017	1.112	20.909	105.793	12,0	9	55	62	127	796	891	15.922	104.780	14,0	7	41	48	96	606	701	11.989
35	21	118.794	119.578	12,0	16	61	68	224	716	789	16.074	117.599	13,6	13	47	54	164	519	592	12.074	118.536	14,8	9	37	44	73	433	506	9.737
35	22	139.249	139.067	11,0	21	75	84	241	905	994	19.997	140.047	12,1	16	61	70	180	748	837	16.376	141.398	13,3	12	49	58	89	615	704	13.299
35	23	125.627	129.520	10,3	10	75	82	122	880	959	19.760	125.730	12,3	8	55	62	92	647	726	14.506	129.572	14,1	8	41	48	79	493	572	10.915
35	24	154.207	153.866	8,9	57	127	138	618	1.623	1.768	34.722	158.351	10,2	22	81	92	284	1.043	1.188	22.220	158.807	12,0	14	55	66	145	626	771	14.307
35	25	149.861	153.779	9,1	27	106	114	586	1.581	1.701	31.131	153.277	10,6	13	75	63	195	1.128	1.246	22.097	151.456	12,5	10	53	61	120	810	930	15.751
35	26	101.984	103.564	11,3	30	90	97	422	1.132	1.242	24.434	100.981	13,0	17	59	66	233	730	840	15.903	104.834	14,1	12	44	51	110	540	650	11.818
35	27	76.323	76.365	13,3	16	45	50	210	500	562	11.590	76.921	14,7	14	37	42	195	434	495	9.747	76.229	17,1	8	27	32	62	337	399	7.305
35	28	86.118	87.367	12,6	14	49	56	178	580	645	12.958	86.456	14,5	12	38	45	144	445	510	10.003	86.955	16,9	9	28	35	65	349	414	7.571
35	29	160.081	168.151	8,2	20	107	119	257	1.435	1.576	29.896	167.477	10,0	15	74	88	165	944	1.065	20.216	165.219	11,1	13	62	74	141	792	933	16.948

			Sample 1									Sample 3								Sample 5							
		SE-96	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra	Pekerja		Wilkers			Units			Extra
Pr.	Kab.	Pekerja	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost	Est.	CV	N.S.	K.S.	Tot.	N.S.	K.S.	Tot.	Cost
35	71	48.722	45.867	15,6	1	34	35	10	453	463	9.471	49.077	21,8	1	23	24	10	339	349	6.716	45.931	21,4	1	20	21	10	258	268	5.491
35	72	26.030	25.579	20,0	2	19	21	21	266	287	5.415	27.202	23,4	2	15	17	21	207	228	4.246	26.049	32,2	2	9	11	21	119	140	2.498
35	73	134.899	134.889	10,0	8	82	88	68	1.064	1.130	22.572	130.859	12,1	8	60	66	66	758	824	16.321	133.320	14,1	6	46	52	66	593	659	12.625
35	74	31.352	31.939	19,7	2	21	23	32	235	267	5.424	33.195	23,6	2	16	18	32	177	209	4.113	31.392	28,3	2	11	13	32	116	148	2.774
35	75	38.126	36.026	20,5	1	21	22	15	297	312	6.013	38.951	21,9	1	18	19	15	278	293	5.377	37.616	25,6	1	14	15	15	226	241	4.275
35	76	23.910	24.639	24,1	1	14	15	9	171	180	3.752	24.417	31,3	1	9	10	9	141	150	2.707	24.969	35,6	1	7	8	9	101	110	2.023
35	77	34.104	33.682	20,3	3	19	22	50	234	284	5.111	34.668	21,6	3	17	20	50	205	255	4.531	33.197	26,5	3	12	15	50	142	192	3.173
35	78	450.963	464.578	6,0	21	236	257	281	2.978	3.259	64.163	449.131	8,0	21	142	163	261	1.811	2.092	38.786	464.927	10,0	21	97	118	281	1.245	1.526	26.571
51	01	36.748	38.698	20,9	5	19	22	64	235	299	5.348	37.207	21,5	3	17	20	40	213	253	4.607	35.089	23,6	3	14	17	40	171	211	3.752
51	02	48.827	52.423	19,3	6	23	28	87	266	361	6.327	44.994	23,9	6	15	20	87	199	262	4.170	47.509	26,1	5	12	17	63	136	199	3.116
51	03	76.346	74.068	12,5	8	44	49	128	742	813	13.737	75.409	14,1	7	36	41	104	634	705	11.495	76.562	17,6	6	27	32	71	460	531	8.474
51	04	105.321	109.103	11,3	27	53	64	676	1.181	1.436	19.275	109.999	12,6	19	42	53	446	959	1.214	15.494	108.867	14,2	14	33	44	255	729	984	11.941
51	05	22.329	22.917	24,1	2	14	16	27	187	214	3.904	23.926	28,7	2	11	13	27	136	163	2.964	23.312	31,7	2	8	10	27	105	132	2.213
51	06	31.741	33.166	20,3	6	20	24	107	322	401	6.099	30.938	22,3	5	16	20	94	231	310	4.626	30.837	27,2	4	10	14	79	157	236	3.011
51	07	47.986	49.650	16,6	12	32	37	188	522	594	9.823	48.267	18,4	8	23	26	139	347	419	6.792	51.862	22,3	7	16	21	72	256	328	4.864
51	08	60.839	69.127	12,5	14	54	62	168	545	645	13.385	69.385	14,2	11	40	48	133	438	538	10.241	69.352	15,2	9	35	43	100	369	469	8.825
51	71	78.491	77.568	12,1	8	55	63	75	577	652	13.841	77.741	14,1	8	42	50	75	453	528	10.687	79.912	16,9	8	31	39	75	351	426	8.046
52	01	83.746	82.393	12,1	8	54	61	95	710	796	14.953	82.282	13,9	7	42	49	86	501	587	11.143	84.507	17,4	7	31	38	66	371	437	8.236
52	02	103.812	101.913	11,2	28	68	82	306	919	1.086	19.030	101.853	12,8	16	51	65	186	647	814	13.898	102.294	14,4	16	39	53	167	547	714	11.124
52	03	148.619	145.469	10,7	21	85	97	332	1.136	1.319	23.712	146.840	12,3	15	57	69	210	807	990	16.330	142.024	14,3	15	42	54	183	611	794	12.188
52	04	46.355	48.405	19,4	11	27	30	111	292	327	6.878	43.897	22,7	7	19	22	65	181	216	4.607	46.704	24,0	6	18	19	35	169	204	4.037
52	05	16.089	15.486	24,9	2	12	14	33	168	201	3.420	17.066	32,6	2	10	12	33	129	162	2.745	17.554	40,8	2	6	8	33	71	104	1.586
52	06	60.289	57.598	16,5	8	47	53	94	556	626	12.426	58.886	17,5	8	36	44	94	467	557	10.402	57.845	19,7	6	27	33	70	339	409	7.324
52	71	52.529	52.081	12,0	3	48	51	69	558	627	12.597	52.607	14,1	3	38	41	69	468	537	10.222	52.538	17,0	3	28	31	69	358	427	7.657
53	01	31.958	33.059	18,5	6	20	26	90	308	398	5.966	33.470	22,5	6	16	22	90	249	339	4.797	32.584	25,5	6	11	17	90	162	252	3.211

USE OF POPULATION CENSUS 2000 TO IMPROVE HOUSEHOLD SURVEY FRAME

Report # 18

by
Sri Budianti Sukmadi

December, 2000

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 125

II. REVIEW OF CURRENT HOUSEHOLD SURVEYS 125
- **A. Socio-Economic Surveys (Susenas)** 125
- **B. Labor Force Surveys (Sakernas)** 125
- **C. Demographic and Health Surveys (SDKI)** 125
- **D. Agriculture Household Surveys** 126

III. REVIEW OF CURRENT ESTABLISHMENT SURVEYS 126

IV. CURRENT SITUATION WITH THE FRAME 127
- **A. Master Frame** 127
- **B. Population Census 2000** 128
  - **1. Preliminary Figures** 128
  - **2. Estimation Level** 130
  - **3. Mapping** 131

V. MASTER FRAME 133
- **A. Evaluation of Existing CB's** 134
- **B. Improvements for Next Census** 135

VI. AREA SAMPLING FOR AGRICULTURE CENSUS 136
- **A. Facts** 136
- **B. Methodology** 138

VII. CONCLUSIONS AND RECOMMENDATIONS 140

APPENDIX A FLOW CHART OF ACTIVITIES 141

APPENDIX B STATISTICAL TABLES 142

I. INTRODUCTION

Household Surveys in Indonesia have been conducted for a long time, starting with the Socio-Economic Survey in 1963 covering all provinces in Jawa only, and continuing with the second survey which covered all provinces in 1967 and was part of BPS activities with financial assistance from the UN. Since 1970 the Socio-Economic Surveys (Susenas) became part of regular BPS activities, and after 1976 they became routine annual surveys. Besides Susenas BPS conducts other household surveys such as Sakernas, Supas, Demographic and Health Surveys etc. using the household approach.

BPS needs to develop a more comprehensive sampling design for the collection of households surveys in an integrated way. If possible, integration should use Census Blocks as primary sampling units for selection whether following the household approach or the establishment approach (for small and household activities). The Population Census 2000 will provide basic information for the household survey frame and for sample designs with modification of the listing questionnaire for sample selection in subsequent steps (ultimate sampling units).

II. REVIEW OF CURRENT HOUSEHOLD SURVEYS

A. Socio-Economic Surveys (Susenas)

The Susenas surveys are yearly and they are divided into two parts:

a. a Core survey is conducted yearly and contains information on basic data for demographic, education, health and social welfare statistics.

b. three modules conducted once every three years :
- one covering detailed demographic questions
- one covering detailed social welfare questions, and
- one covering consumption and expenditure questions.

B. Labor Force Surveys (Sakernas)

The Sakernas surveys are conducted twice a year and contain the same information as in the Core Susenas especially for labour force items.

C. Demographic and Health Surveys (SDKI)

The last survey was conducted in 1998. Part of the questionnaire covers the same items as in Susenas and Sakernas but the survey's focus is on family planning and health characteristics.

D. Agriculture Household Surveys

In the current situation the weakest data are the agriculture data except for main food crops with yearly activities. Improvements in the agricultural design methodology are needed. Alternative designs need to evaluate whether to use a multi-purpose design for all agricultural activities or to integrate the primary sampling units only or independently for groups of commodities such as estate crops, food crops, livestock and poultry, land fishery, marine fishery and others.

III. REVIEW OF CURRENT ESTABLISHMENT SURVEYS

Establishment surveys cover all sectors except agriculture, are based on the 1996 Economic Census, and cover incorporated businesses. The sectors covered by the 1996 Economic Census (SE96) can be divided into 3 main groups:

a. Manufacturing, electricity, gas and water supply and construction
b. Wholesale and retail trade, hotels and restaurants, transport, storage and communications
c. Services (financial intermediates, real estate, education, health, etc.)

The legal status of establishments can be divided into :

a. With permanent location :
   - Corporate
   - Non Corporate
b. Without permanent location

Establishments in manufacturing are divided into large/medium industry and small scale/household industry.

Incorporated establishments include those in the large/medium category and are surveyed based on the establishment approach using an individual directory as a frame. Non-incorporated establishments and establishments without permanent location (household industry) are surveyed using area sampling based on enumeration areas. Currently an integrated design is used for small/household industries by using the same questionnaire for each industry. The economic census results form a base for a master frame with information by 4-digit code.

To attempt an integrated design between household surveys and small scale/household establishments, three questions need to be asked:

a. Is there any way to use the master frame from the Population Census 2000 for small scale/household industries as a master frame for these surveys?

b. Is there a possibility for integrating household surveys and small-scale/ household establishments, especially in first stage sampling units?

c. Is there a possibility for using the household approach for small-scale /household establishment surveys through household economic activities.

To answer these questions we have to evaluate the availability of the master frame and data used for sample design.

IV. CURRENT SITUATION WITH THE FRAME

In Indonesia the decennial Population Census constitutes the primary source for constructing a master frame for the agriculture census, the economic census and surveys conducted between two population censuses.

A. Master Frame

During the 1990 Population Census, enumeration areas (wilayah pencacahan/wilcah) were constructed for field operations as well as for the master frame. Each enumeration area contained about 300 households and enumeration areas were hand-sketched on a map. Using wilcah as a primary sampling unit for all censuses after the population census and for surveys cannot yield reliable results for all data especially for very specific characteristics (rare cases) and also for listing. Targeting results with a high precision will not meet expectations if one only uses the wilcah as the primary sampling unit without any further identification, especially for surveys of economic activities (e.g. agriculture and small establishment surveys). Most surveys cover the whole country, with three levels of estimation: national, provincial and kabupaten /kota.

The frame and its data are very useful for providing good sample design. Census data from the sampling frame are one of the best sources in Indonesia, since alternative sources such as administrative records are not available. Construction of the master frame from the Population Census 2000 has been accomplished by using listing questionnaires (L1) or individual questionnaires (L2). This has proven very useful for integrated household surveys, especially those using the household approach (such as Sakernas and Susenas).

B. Population Census 2000

The processing of the Population Census 2000 will be finished around June 2001 with complete count data. The evaluation of construction of the master sampling frame will have to wait for the finalization of data cleaning. For some provinces (especially the small ones) data will be finished in the first or the second month of 2001. Using data from these provinces will be necessary to determine in a preliminary way whether using the Population Census results for stratification is feasible and to define effective stratification and the CB frame for the master sampling frame as suggested by Mr. Vijay Verma.

1. Preliminary Figures

Based on the preliminary figures of the population census, Table 1 provides indicators of the condition of the total number of CB, households and population divided by urban and rural.

Table 1
Total Number of Common Census Blocks and Households by Group of Households

Item	Number of Households								Total
	0	1-20	21-40	41-80	81-120	121-150	151-200	>200	
Urban CB	18580	2680	6520	49447	85827	33099	17326	8064	204821
Average per CB		11	32	65	100	133	170	262	105
Rural CB	58630	2448	6289	56990	135613	54251	20639	4582	286675
Average per CB		11	32	66	100	132	168	243	103
Urban+Rural CB	77210	5128	12809	106437	221440	87350	37965	12646	491496
Average per CB		11	32	66	100	133	169	255	103

Source: preliminary figure, Population Census 2000

Table 2 presents the number of common CB's and total CB's by different groupings.

Table 2
Number of kabupaten/kota by grouping number of CB

Area	Grouping total number of CB							Total
	<=500	501-909	1000-1499	1500-1999	2000-2499	2500-2999	>=3000	
Common CB								
No. of kab/kota	89	81	49	35	28	22	36	340
Average CB	338	698	1189	1738	2239	2733	4529	1446
Total CB								
No. of kab/kota	64	81	54	28	32	21	60	340
Average CB	351	724	1212	1730	2222	2717	4464	1738

Source: preliminary figure, Population Census 2000

Table 2 shows that the total number of CB's varies from one kabupaten/kota to another. Some kabupaten/kota have too few CB's while others have too many. Since Indonesian household surveys are routine surveys, they usually have to tabulate results down to the provincial level by urban and rural, and down to the kabupaten/kota level for urban and rural combined. Based on past experience with households surveys (e.g. Socio-Economic Surveys, Labour Force and related household surveys) the inter-class correlation coefficient (roh) for general characteristics (such as health, labour force) is less than 0.30. Under such conditions the minimum sample size for domain tabulation is around 600 households (domain of kabupaten/kota) with the CB sample size (cluster size) between 6-16 households.

With the limited CB characteristics, i.e. that estimations have to be produced down to the kabupaten/kota level, the construction of the master sampling frame has to be done as simply as possible. There are two alternatives for the construction of household survey (such as Sakernas/ Susenas) master sampling frame:

a. **First alternative**: the construction of stratification using the cross-classification of variables to get area with potential if the level of estimation is by province only.

b. **Second alternative**: use a simple procedure by systematic arrangement and divide into two strata (e.g. agriculture and non-agriculture) using independent sub-samples, especially for Sakernas/ Susenas-type applications. The total number of CB's for the master frame are at least 5 times the minimum sample size. The suggestion is not to use replication design but to divide into 5-10 independent samples because we have to use kabupaten/ kota as the domain of study and if possible to divide into urban and rural areas for estimation purposes at the kabupaten/ kota level.

Table 3 shows the suggested sample size for the master sampling frame by size of CB.

Table 3
Suggested Sample Size for Master Sampling Frame by Size of CB

Size of CB Kab/Kota	Average No. of CB per Kab/Kota	No. of CB in Master Frame Sampling	Independent Sample
<=500	351	1/5	10
501 - 999	724	1/5	10
1000 - 1499	1212	1/5	10
1500 - 1999	1738	1/10	20
2000 - 2499	2222	1/10	20
2500 - 2999	2717	1/10	20
>=3000	4464	1/15	30
Total	1738	-	-

2. Estimation Level

The Government of Indonesia has adopted a policy of decentralization since 1999. This implies that there will be a shortage of statistical data on local administration. Data on provinces and kabupaten/kota are essential for Indonesia. Construction of a master sampling frame will become very useful and BPS has to construct one which is as simple as possible. Perhaps the National Government will be responsible for budgeting national and provincial estimates only and local governments will be responsible for estimations by regency. A simple method should be adopted, by using a simple selection, e.g. a systematic sampling arrangement by CB grouping with agriculture and non-agriculture strata as mentioned above. That would be easy to implement.

Using the ratio estimation method, regency offices can estimate levels for their own regency by using Population Census data as a benchmark. Changing characteristics of CB's will be easy to detect: in every listing we have to detect the characteristics between the census period and the period of the survey in the kabupaten/kota as part of the supervision of enumeration results.

The method of estimation is for each domain.

$$\hat{Y}_R = \frac{y}{x} X \qquad R = \frac{y}{x}$$

where y, x are the sample totals of Y and X respectively

$$v\left(\hat{Y}_R\right) = \frac{N^2\left(1-f\right)}{n}\left(s_y^2 + \hat{R}^2\ s_x^2 - 2\hat{R}\ s_{yx}\right)$$

on the basis of equal probability selection (systematic sampling), or modification by using stratification.

To simplify the calculation of variance, using independent sub-samples could be adopted, as :

$$\hat{Y} = \frac{1}{s}\sum_{i=1}^{s} \hat{Y}_i$$

$$v\left(\hat{Y}\right) = \frac{1}{s(s-1)}\sum_{i=1}^{s}\left(\hat{Y}_i - \hat{Y}\right)^2$$

where s is the number of independent sub-samples.

Using stratification the formulas are:

$$\hat{Y} = \sum_{h=1}^{H} \frac{1}{s} \sum_{i=1}^{s} \hat{Y}_{hi} = \frac{1}{s} \sum_{h=1}^{H} \sum_{i=1}^{s} \hat{Y}_{hi}$$

$$v(\hat{Y}) = \sum_{h=1}^{H} v\left(\hat{Y}_h\right)$$

$$= \sum_{h=1}^{H} \frac{1}{s(s-1)} \sum_{i=1}^{s} \left(\hat{Y}_{hi} - \hat{Y}_h\right)^2$$

Two independent samples may be enough for national or provincial estimates. The estimation procedure thus becomes simpler:

$$\hat{Y} = \frac{1}{2} \left(\hat{Y}_1 + \hat{Y}_2\right)$$

$$v(\hat{Y}) = \frac{1}{4} \left(\hat{Y}_1 - \hat{Y}_2\right)^2$$

Using stratification the formulas are:

$$\hat{Y} = \frac{1}{2} \sum_{s=1}^{H} \left(\hat{Y}_{s1} + \hat{Y}_{s2}\right)$$

$$v(\hat{Y}) = \frac{1}{4} \sum_{s=1}^{H} \left(\hat{Y}_{s1} - \hat{Y}_{s2}\right)^2$$

3. Mapping

Other than construction of a master sampling frame and choice of a sampling method, it is also important to monitor the CB's themselves including their mapping. Maintenance of the enumeration areas (Wilcah) in the last master frame was not so good and that resulted in a weak listing and the necessity of conducting a detailed evaluation for every survey. Some of the weakness were:

a. The boundary of the enumeration area (wilcah) was not clear and the wilcah area was too big (three times the size of the CB) containing around 300 households.

b. Filling of the maps was not good: no special unit in the regional offices was responsible for filling and monitoring maps and evaluating the existing conditions from time to time.

Since the new BPS organization in 1998, there is one unit responsible for maintenance of the map and the master frame. Some recommended activities for this unit are:

a. Filing the CB map whether in hard copy or in a computer file (using scanners utilized for processing the L2 population census)

b. Filing questionnaires (L1) together with the map and scan

c. Before listing for household surveys, use the CB map as a guide and print from the computer file.

d. Identify boundaries before listing.

e. After listing, the kabupaten/kota supervisor has to check with the ordinary listing results from the population census, and determine whether the enumerator did the listing in an appropriate way. Without good listing, the bias will be high even if the design is theoretically correct.

f. Monitoring of the availability of maps has to be conducted as a routine activity. After construction of the master sampling frame, the CB maps (including data) have to be checked and put in a database.

The main problems in the listing, which caused undercoverage, are the boundary problems and how seriously monitoring/supervising of the listing is conducted. In order to overcome these problems in the future, the unit who is responsible for the master frame and the maps has to do the following:

a. Identify the CB samples clearly and completely.

b. Identify the important mark of the CB at the kabupaten/kota level before distributing the questionnaire using a copy of the basic map, listing of the population census questionnaire. Marking in the copy map should include:
   - the building number, the middle and last name of the head of the household and the household number.
   - from the listing questionnaire, write down the name of the head of the household on the map
   - after listing, the enumerators have to sign in the listing questionnaire the location of the households and this has to be checked by the regional office against the master frame.

c. These activities are very important because in the last surveys enumerators sometimes went to the wrong CB. The difference of CB's between the sample list and the listing have a very big influence on the estimation (i.e. the probability becomes not fit).

V. MASTER FRAME

A good sample design requires balancing four criteria :

- Goal orientation
- Measurability
- Practicality
- Economy

Master frames with data which are highly correlated with survey objectives can serve as benchmarks for probability samples. Surveys are then designed to permit estimation of the sampling variability (ordinarily the standard error) of the estimates. What happened in the construction of CB's in the population Census 2000?

From Population Census 2000 there were 3 types of Census Blocks:

a. "Common" which refers to CB's used for buildings and places where household activities, offices, business activities, etc. take place.

b. "In preparation" which refers to empty CB's such as paddy fields, plantations, forests etc.

c. "Special" which refers to CB's out of reach of the public. That is, one needs to make special arrangements for entering: e.g. prisons, military locations etc

The total number of CB's by type in Indonesia is provided in Table 4:

Table 4
CB by Type

Type of CB	Urban	Rural	Total
(1)	(2)	(3)	(4)
Common	204821	286675	491496
Preparation	15473	81809	97282
Special	1516	679	2195
Total	221810	369163	590973

Source: Preliminary figure, Population Census 2000

The table shows that the total number of the "in preparation" Census Blocks is big: about 16 percent of the total number of CB's. It will make a difference whether one includes this category in the population frame or not. We have to consider the condition of these CB's and separate them into two categories: the first category would include those close to households or business areas, and the second category would include those from remote areas, e.g. forests, marshlands/swamps. In the short term, the first category of "in preparation" CB's may contain households or buildings, so we have to consider them as one stratum and give them special treatment. Ordinary strata would come from the "common" CB's using procedures mentioned above.

A. Evaluation of Existing CB's

Area sampling has to be used for household surveys and multistage sampling design needs to be used for selecting samples. CB's are the primary sampling units. The results of the evaluation of current CB's are as follows:

a. Updating of the CB's is very useful, especially for preparation of the agriculture census 2003 (boundary, size).

b. In urban areas of some provinces, there is a tendency to construct "in preparation" CB's in the smallest administrative units (Rukun Tetangga/RT) close to "common" CB's. This will have a big effect because these areas are close to areas with a dense population, which makes the boundaries between areas of dwelling units and areas of empty CB's unclear. This actually happened in Jakarta, and seems to be caused by the fact that the enumerators were paid according to the number of CB's. In this case, it is better to include such areas in one CB with a clear boundary.

c. In rural areas of some provinces, this also happened where empty CB's were close to an area of dwelling units, or where there was land for paddy, plantations, forests etc. In Central Kalimantan, for example, there were villages with many empty CB's containing forests (in one village there were 10 such CB's but only one "common" CB). The enumerators divided the forest area into more that one CB with a river or a road as boundary. In such cases, the CB boundaries were very nice and clear, but in other cases the boundaries were not clear and maybe difficult to detect (they were only imaginary). One interesting case was found in Riau where one village contained 20 CB's, but only 4 were "common" CB's. The rest (16 CB's) were areas of wet land of only 1,2,3,4,5, ha (which are considered very small) and without clear boundaries.

d. To use directly the results of mapping from Population Census 2000 and to examine them carefully one by one before using them as the basis for listing. The compensation method is the one cause of undercoverage or overcoverage. There are two alternatives to using the CB's from Population Census 2000 result:

- For the 2000 household surveys, using only "common" CB's may be reasonable because the incidence of people moving from one place to another is not so high (it will have a small effect on the estimation for the period June 2000 – January 2001).
- For the 2002 household surveys, the effect of people moving may increase, especially for "in preparation" CB's close to dense population areas. This has happened in urban areas or suburbs, for example in DKI Jakarta and surrounding kabupaten/kota (Botabek). DKI Jakarta consists of 248 "in preparation" CB's out of 19362, some of which are empty lands in one Rukun Tetangga with very small areas.

B. Improvements for Next Census

A Census is the main source for constructing a master frame or a master sampling frame. Based on the evaluation of the last census of construction, the Enumeration Areas during mapping ("Lingkungan" in 1960, "Block Census" in 1970, "Wilayah Pencacahan" in 1980 and 1990 and "Block Census" in 2000) always requires new boundary maps and it was difficult to compare enumeration areas from one census to another. No basic maps were constructed, only sketches. Villages (desa) have the lowest digitized sketch maps and these were always up-to-date during a census with regard to information on the total number of Enumeration District (Wilcah, Block Census depending on the Census). Filling of maps takes place in regional offices and headquarters down to the desa level with boundaries of the smallest enumeration areas. The map containing detailed enumeration areas (CB in 2000) is kept in regional offices to be used for field operations. Sometimes maps disappear and problems arise during field operations.

Three improvements can be made, starting with the 2003 Agricultural Census: overcoming the problem of sketching maps, standardizing household size of CB's and paying particular attention to the "in preparation" CB's. More specifically:

a. Pay close attention to the completeness of the maps. The CB maps have to be scanned in the regional offices (i.e. put in computer files) with unique identifiers. Scanners are available in the regional offices and can be used for any kind of survey.

b. The BPS Head Office should evaluate village maps divided into CB's, especially the boundaries, the contents of CB's, identification etc. One should assume that 80 percent of CB's should have clear boundaries and maps without clear boundaries should be identified.

c. Merge small CB's and link the "in preparation" CB's close to buildings/ households. Copies of these maps need to be sent to regional offices to update their records. As was done in the past, updated maps have to be sent to the Head Office with the original evaluation maps from the Head Office.

A flow chart of activities is provided in Appendix A.

VI. AREA SAMPLING FOR AGRICULTURE CENSUS

A. Facts

The Government of Indonesia has conducted the Agriculture Census since 1963, with a combination of the household approach (through area sampling) and the establishment approach (for incorporated establishments). The Agriculture Census covers all activities in agriculture in the whole country undertaken by households as well as establishments.

The scope of Agriculture Census 1993 units was:

a. All agricultural households, including those with and without land holding:
Land holding farm households covered households doing the following activities:
- Paddy and other main food crops
- Horticulture cultivation
- Plantation cultivation
- Animal husbandry
- Fish culture in fresh water ponds
- Fish culture in breakout water ponds

Agricultural households without landholding included households doing the following activities:
- Fish culture in marine water
- Fish culture in public water
- Fish catching in marine water
- Fish catching in public water

b. Agricultural establishments: enumeration for these establishments was conducted based on the latest agricultural establishment directory. These activities covered all establishments in agriculture that are legal corporations.

In addition, activities which were highly correlated with agriculture were also included, e.g. those of the Village Cooperative Units and the Village Potentialities Data.

The agriculture census units were the basic units defined and the list was carefully prepared with detailed characteristics, and whether to be conducted by complete count or on a sample basis.

The questions for the 2003 Census are:

a. Will the census be conducted on a complete count basis? If yes, then this raises a budget problem, especially the efficiency of the budget if we want to list all households. There are two ways to detect the concentration in urban areas: one uses the household approach and the other uses the area approach (e.g. location of plantations). The first approach will be difficult because there are no individual identifiers (there is no community of farmers). The second approach is easier to identify but is more difficult for interviews because respondents are spread across different urban and rural locations. Most urban areas do not have a potential in agriculture.

b. Will the census be conducted on a sample basis? If yes, then what is the best frame available for selecting the sample? This frame will also be used for subsequent surveys as a master sampling frame.

c. How to construct the listing questionnaire to produce the basic data and to be used subsequently for selecting samples based on more detailed characteristics?

d. How can the population census blocks be used as a frame for sample selection if sufficient budget is not available for conducting the agriculture census on a complete count basis.

e. Will the census be conducted on a complete count basis for the detailed structure of agricultural households or on a sample basis?

B. Methodology

Agricultural data cover various sub-sectoral activities each having different population characteristics spread across entire regions in Indonesia. The population and location of food crop households, for example, is almost spread all over the country, whereas the population of small-holder plantations and households doing fish cultivation is only located in certain regions.

Based on experience with the 1993 Agriculture Census, the sample of 20 percent of Enumeration Areas (wilcah) produced good results as a benchmark for the structure of agricultural activities and the small sample produced good results for detailed characteristics especially for national and provincial estimates. The 20 percent of Enumeration Areas was used as a master sampling frame for agriculture surveys.

The most important lesson from that census is the need to limit the scope. If the scope is too big, then we have to have a large number of enumerators and it will not be possible to get good quality data for detailed characteristics. That is why BPS needs to make a special effort to select only items related directly to the agriculture census in the master frame (population census). These items are the main activities in agriculture as a main occupation divided by:

- Food crops
- Horticulture
- Animal Husbandry
- Fishery
- Plantation

Employment status would be divided into:

- Self-employment
- Self employed assisted by the temporarily employed
- Employer
- Employee
- Unpaid family worker

Therefore, if the budget for the 2003 agriculture census is limited, the following alternatives are suggested:

a. Select 20 percent of census blocks (as was done in 1993) for rural areas and conduct a complete count for urban areas of concentration as mentioned above. The sample selection with probability proportional to size can be done using the total number of households in each census block as the measure of size. Alternatively, direct systematic CB selection by equal probability can be done.

b. Select 20 percent of census blocks (as was done in 1993) for rural areas and conduct a complete count for urban areas of

concentration as mentioned above, but using data from the 2000 Population Census for stratification. This would be very useful, especially for areas with fisheries and plantations. The allocation of 20 percent of census blocks will depend on the result of stratification (the primary sub-sectoral activity). Five strata can be used: food crops, horticulture, animal husbandry, fishery and plantation. The data used for stratification would be households with members engaging in activities in those sub-sectors and whose employment status would be:

- self-employed
- self-employed assisted by temporary employees
- employer (with an indication of whether in the same sub-sector or another one)

Households with more than one farmer with different employment status would then be classified as a farming household with the sub-sector corresponding to that of farmers whose status is "employer". If this status is not available then they would be classified according to the sub-sector corresponding to farmers whose status is "self-employed assisted by temporary workers." If this status is also not available then they would be classified according to the sub-sector corresponding to farmers whose status is "self-employed." If all farmers within the household have the same status, then they would be classified according to the main activity of farmers in that area of the CB. The data from Population Census 2000 should then be used only for stratification and for the selection of the 20 percent sample by systematic sampling in which census blocks are used as sampling units.

c. Use the same procedure as in b, but use a separate procedure for agricultural households without land holding especially for fisheries.

d. Use the same procedure as in b or c, but for stratification use "persons engaged as farmers" (total number of farmers in a sub-sector) rather than agricultural households. Farmers here are defined as members of a household engaged in agricultural activities with employment status as "self-employed" or "self-employed assisted by temporary employees" or "employers".

e. Use the same procedure as in b or c, but for stratification use all household members engaged in agricultural activities including employees. Stratification is used only for grouping whether the area is included in specific potential crops.

VII. CONCLUSIONS AND RECOMMENDATIONS

Before deciding on the sample design, it is important to evaluate the results of Population Census 2000 mapping and to reduce its weaknesses by using experience in existing surveys regarding enumeration areas. The master frame needs to be evaluated to overcome the weakness stemming from the lack of maps, their condition, the filing system and maintenance, coding and availability of data.

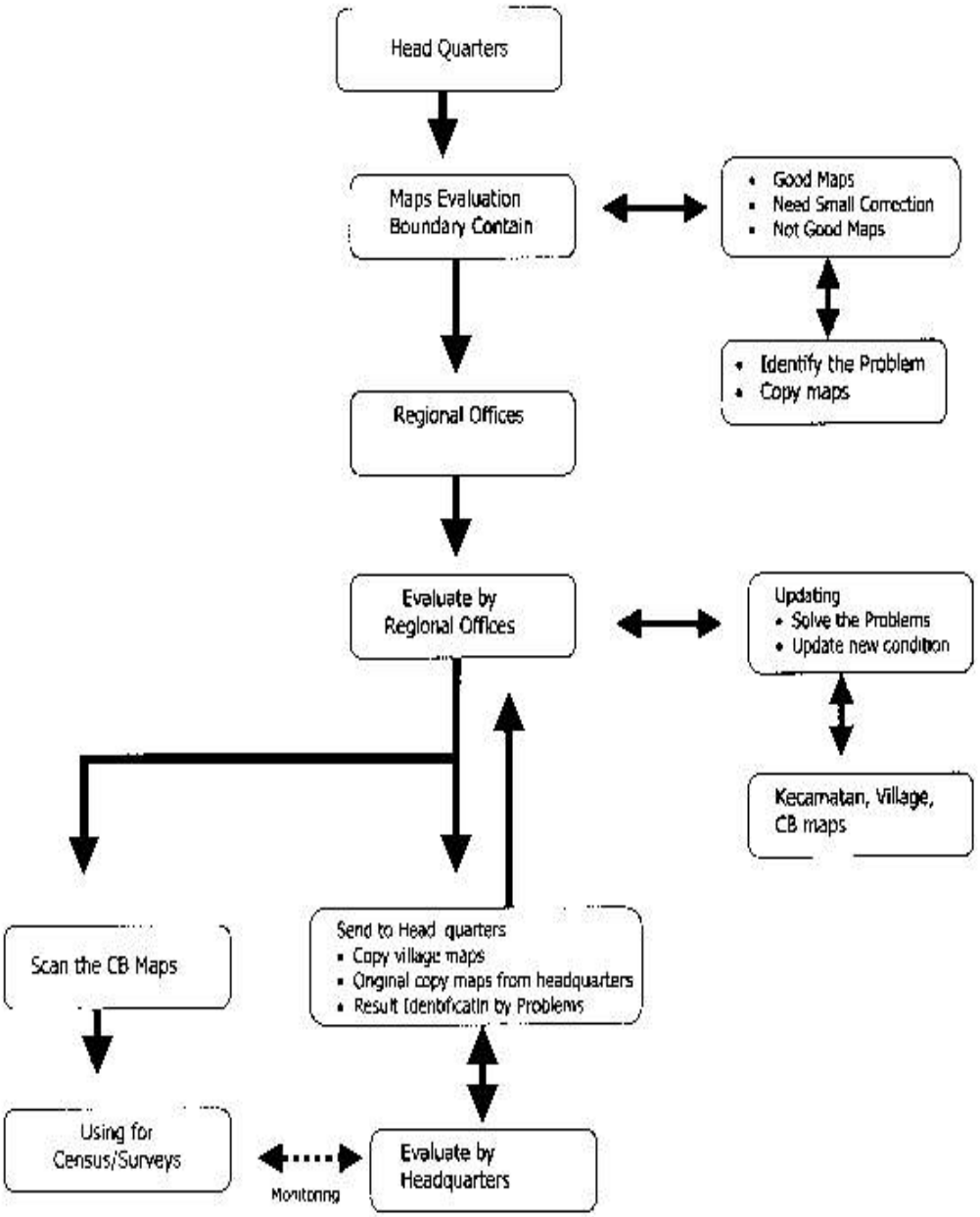
To construct one master sampling frame for household surveys such as Sakernas, Susenas, etc. one has to incorporate data from various sampling frames used for agricultural household census/surveys and the results of the Population Census 2000. For surveys of small establishments, it is recommended to continue using wilcah as was done in the past (Susi 1998 and 1999). These wilcah have to be maintained until we have a new master frame.

We need to build a master frame/master sampling frame **database**, which will be directly accessible to all subject matter divisions through an intranet system (i.e. one basic frame).

Past experience suggests that it is very important not to always change the enumeration area concept, such as the movement from Census Block to Wilcah then to Census Block again. Grid systems or Mesh Statistics have to be developed. With such systems, the preparation of the census, the construction of master frames and master sampling frames will be more accurate. This will also be very helpful for small area dissemination, especially if the government's decentralization policy is pursued.

APPENDIX A
FLOW CHART OF ACTIVITIES

APPENDIX A
FLOW CHART OF ACTIVITIES



APPENDIX B
STATISTICAL TABLES

Table B.1
Number of Kab/Kota, Sub Districts, Villages and Common Census Blocks

Province	Kab/Kota	Sub-Districts	Villages			Common Census Blocks		
			Urban	Rural	Total	Urban	Rural	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
01. Dista Aceh	12	123	385	4,253	4,638	1,828	6,763	8,591
02. Sumatera Utara	19	265	758	4,592	5,350	10,543	14,638	25,181
03. Sumatera Barat	15	120	396	1,780	2,176	2,989	7,798	10,787
04. Riau	15	96	205	1,263	1,468	4,842	6,046	10,888
05. Jambi	10	60	123	1,038	1,161	1,658	4,307	5,965
06. Sumatera Selatan	10	110	363	2,609	2,972	5,670	11,166	16,836
07. Bengkulu	4	32	118	1,044	1,162	892	2,336	3,228
08. Lampung	10	87	174	1,898	2,072	3,204	11,848	15,052
09. DKI Jakarta	5	43	265	0	265	19,038	0	19,038
10. Jawa Barat	28	543	2,233	4,991	7,224	51,734	50,112	101,846
11. Jawa Tengah	35	534	2,332	6,211	8,543	30,600	46,198	76,798
12. DI Yogyakarta	5	75	169	269	438	5,196	3,513	8,709
13. Jawa Timur	37	624	2,399	6,053	8,452	37,129	54,086	91,215
14. Bali	9	53	218	460	678	3,499	4,029	7,528
15. Nusa Tenggara Barat	7	62	168	535	703	3,121	6,041	9,162
16. Nusa Tenggara Timur	14	124	147	2,368	2,515	1,046	6,739	7,785
17. Kalimantan Barat	9	128	98	1,346	1,444	2,189	6,618	8,807
18. Kalimantan Tengah	6	85	58	1,271	1,329	1,196	3,449	4,645
19. Kalimantan Selatan	11	117	202	1,951	2,153	2,876	5,432	8,308
20. Kalimantan Timur	12	88	141	1,137	1,278	3,804	3,244	7,048
21. Sulawesi Utara	8	93	262	1,264	1,526	2,513	4,723	7,236
22. Sulawesi Tengah	8	68	84	1,337	1,421	913	3,786	4,699
23. Sulawesi Selatan	24	197	403	2,727	3,130	5,285	12,491	17,776
24. Sulawesi Tenggara	5	67	106	1,445	1,551	713	3,051	3,764
25. Maluku	5	29	69	776	845	704	2,077	2,781
26. Maluku Utara	3	27	53	671	724	355	1,538	1,893
27. Irja Barat	4	44	22	730	752	405	1,094	1,499
28. Irja Tengah	5	44	27	615	642	223	1,047	1,270
29. Irja Timur	5	85	37	1,437	1,474	656	2,505	3,161
INDONESIA	340	4,023	12,015	56,071	68,086	204,821	286,675	491,496

Table B.1a
Number of Kab/Kota. Sub Districts. Villages and Census Blocks

Province	Kab/Kota	Sub-Districts	Villages			Census Blocks		
			Urban	Rural	Total	Urban	Rural	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
01. Dista Aceh	12	123	385	4,253	4,638	1,975	9,558	11,533
02. Sumatera Utara	19	265	758	4,592	5,350	11,039	18,600	29,639
03. Sumatera Barat	15	120	396	1,780	2,176	3,075	8,716	11,791
04. Riau	15	96	205	1,263	1,468	5,111	8,264	13,375
05. Jambi	10	60	123	1,038	1,161	1,735	6,036	7,771
06. Sumatera Selatan	10	110	363	2,609	2,972	6,007	16,381	22,388
07. Bengkulu	4	32	118	1,044	1,162	938	3,378	4,316
08. Lampung	10	87	174	1,898	2,072	3,382	15,178	18,560
09. DKI Jakarta	5	43	265	0	265	19,362	0	19,362
10. Jawa Barat	28	543	2,233	4,991	7,224	55,748	60,079	115,827
11. Jawa Tengah	35	534	2,332	6,211	8,543	35,041	60,537	95,578
12. DI Yogyakarta	5	75	169	269	438	5,238	3,644	8,882
13. Jawa Timur	37	624	2,399	6,053	8,452	41,784	66,504	108,288
14. Bali	9	53	218	460	678	3,664	4,412	8,076
15. Nusa Tenggara Barat	7	62	168	535	703	3,342	7,005	10,347
16. Nusa Tenggara Timur	14	124	147	2,368	2,515	1,106	8,137	9,243
17. Kalimantan Barat	9	128	98	1,346	1,444	2,296	8,444	10,740
18. Kalimantan Tengah	6	85	58	1,271	1,329	1,327	6,120	7,447
19. Kalimantan Selatan	11	117	202	1,951	2,153	3,004	7,503	10,507
20. Kalimantan Timur	12	88	141	1,137	1,278	4,101	5,810	9,911
21. Sulawesi Utara	8	93	262	1,264	1,526	2,711	6,397	9,108
22. Sulawesi Tengah	8	68	84	1,337	1,421	964	5,579	6,543
23. Sulawesi Selatan	24	197	403	2,727	3,130	5,452	14,127	19,579
24. Sulawesi Tenggara	5	67	106	1,445	1,551	754	4,605	5,359
25. Maluku	5	29	69	776	845	829	3,348	4,177
26. Maluku Utara	3	27	53	671	724	406	2,645	3,051
27. Irja Barat	4	44	22	730	752	461	2,471	2,932
28. Irja Tengah	5	44	27	615	642	259	1,419	1,678
29. Irja Timur	5	85	37	1,437	1,474	699	4,266	4,965
INDONESIA	340	4,023	12,015	56,071	68,086	221,810	369,163	590,973

Table B.2
Total Number of Common Census Blocks and Households
by group number of households

ITEM	Number of Households								Total
	0	1-20	21-40	41-80	81-120	121-150	151-200	>200	
(1)	(2)	(3)	(4)	(50	(6)	(7)	(8)	(9)	(10)
Urban									
1. Census Blokcs (CB)	1,858	2,680	6,520	49,447	85,827	33,099	17,326	8,064	204,821
2. Households (HH)	0	30,640	210,376	3,223,665	8,549,843	4,405,481	2,938,954	2,115,970	21,474,929
3. Average per CB	0.00	11.43	32.27	65.19	99.62	133.10	169.63	262.40	104.85
Rural									
1. Census Blokcs (CB)	5,863	2,448	6,289	56,990	135,613	54,251	20,639	4,582	286,675
2. Households (HH)	0	27,422	202,565	3,769,530	13,609,728	7,199,734	3,464,246	1,112,481	29,385,706
3. Average per CB	0.00	11.20	32.21	66.14	100.36	132.71	167.85	242.79	102.51
Urban + Rural									
1. Census Blokcs (CB)	7,721	5,128	12,809	106,437	221,440	87,350	37,965	12,646	491,496
2. Households (HH)	0	58,062	412,941	6,993,195	22,159,571	11,605,215	6,403,200	3,228,451	50,860,635
3. Average per CB	0.00	11.32	32.24	65.70	100.07	132.86	168.66	255.29	103.48

Table B.2a
Total Number of Census Blocks and Households
by group number of households

ITEM	Number of Households								Total
	0	1-20	21-40	41-80	81-120	121-150	151-200	>200	
(1)	(2)	(3)	(4)	(50	(6)	(7)	(8)	(9)	(10)
Urban									
1. Census Blokcs (CB)	17,401	3,238	6,697	49,724	86,031	33,176	17,379	8,164	221,810
2. Households (HH)	0	33,881	215,659	3,240,223	8,569,997	4,415,815	2,948,015	2,148,311	21,571,901
3. Average per CB	-	10.46	32.20	65.16	99.62	133.10	169.63	263.14	97.25
Rural									
1. Census Blokcs (CB)	87,730	2,663	6,416	57,124	135,680	54,269	20,672	4,609	369,163
2. Households (HH)	0	29,346	206,318	3,777,451	13,616,342	7,202,131	3,469,782	1,120,797	29,422,167
3. Average per CB	-	11.02	32.16	66.13	100.36	132.71	167.85	243.18	79.70
Urban + Rural									
1. Census Blokcs (CB)	105,131	5,901	13,113	106,848	221,711	87,445	38,051	12,773	590,973
2. Households (HH)	0	63,227	421,977	7,017,674	22,186,339	11,617,946	6,417,797	3,269,108	50,994,068
3. Average per CB	-	10.71	32.18	65.68	100.07	132.86	168.66	255.94	86.29

Table B.3a
Total Number of Common Census Blocks and HouseHolds by Group Number of HouseHolds and Provinces

Urban + Rural

Province	<= 40			41 - 120			>=121			Total		
	CB	HH 0	Average per CB	CB	HH 0	Average per CB	0 CB	HH 0	Average CB	CB	HH 0	Average CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 D I ACEH	4058	10576	2.61	3429	287394	83.81	1,104	173681	157.32	8591	471651	54.90
2 SUMATERA UTARA	1725	39673	23.00	16391	1406122	85.79	7,065	1085346	153.62	25181	2531141	100.52
3 SUMATER A BARAT	429	12628	29.44	8231	695109	84.45	2,127	313785	147.52	10787	1021522	94.70
4 RIAU	859	19865	23.13	6712	566481	84.40	3,317	563590	169.91	10888	1149936	105.62
5 JAMBI	379	8689	22.93	4083	350034	85.73	1,503	230803	153.56	5965	589526	98.83
6 SUMATER A SELATAN	1271	24446	19.23	11044	944797	85.55	4,521	702012	155.28	16836	1671255	99.27
7 BENGKULU	133	3015	22.67	2058	180006	87.47	1,037	161790	156.02	3228	344811	106.82
8 LAMPUNG	595	11222	18.86	9927	899338	90.60	4,530	681618	150.47	15052	1592178	105.78
9 DKI JAKARTA	1218	30119	24.73	10674	890082	83.39	7,146	1323212	185.17	19038	2243413	117.84
10 JAWA BARAT	4450	94219	21.17	58727	5285953	90.01	38,669	6114021	158.11	101846	11494193	112.86
11 JAWA TENGAH	1165	21843	18.75	58665	5442812	92.78	16,968	2371898	139.79	76798	7836553	102.04
12 D I YOGYAKARTA	132	3619	27.42	6322	574739	90.91	2,255	343617	152.38	8709	921975	105.86
13 JAWA TIMUR	2354	53475	22.72	65072	5919128	90.96	23,789	3511288	147.60	91215	9483891	103.97
14 BALI	217	6272	28.90	5203	452628	86.99	2,108	341246	161.88	7528	800146	106.29
15 NUSA TENGGARA BARAT	224	4824	21.54	6030	551086	91.39	2,908	429062	147.55	9162	984972	107.51
16 NUSA TENGGARA TIMUR	178	4615	25.93	5607	496375	88.53	2,000	291647	145.82	7785	792637	101.82
17 KALIMANTAN BAR AT	648	14753	22.77	6115	511340	83.62	2,044	317388	155.28	8807	843481	95.77
18 KALIMANTAN TENGAH	461	11514	24.98	2861	236618	82.70	1,323	221966	167.77	4645	470098	101.21
19 KALIMANTAN SELATAN	572	12170	21.28	5957	502173	84.30	1,779	270287	151.93	8308	784630	94.44
20 KALIMANTAN TIMUR	983	23861	24.27	4807	388135	80.74	1,258	198045	157.43	7048	610041	86.56
21 SULAWESI UTARA	298	6473	21.72	4747	417820	88.02	2,191	330986	151.07	7236	755279	104.38
22 SULAWESI TENGAH	292	4472	15.32	3214	282019	87.75	1,193	175712	147.29	4699	462203	98.36
23 SULAWESI SELATAN	728	17875	24.55	12951	1129314	87.20	4,097	608807	148.60	17776	1755996	98.78
24 SULAWESI TENGGARA	110	3213	29.21	2519	221461	87.92	1,135	175457	154.59	3764	400131	106.30
25 MALUKU	497	7332	14.75	1716	139957	81.56	568	88611	156.01	2781	235900	84.83
26 MALUKU UTARA	542	3756	6.93	989	81431	82.34	362	56359	155.69	1893	141546	74.77
27 IRJA BARAT	296	5827	19.69	927	73830	79.64	276	44433	160.99	1499	124090	82.78
28 IRJA TENGAH	248	4202	16.94	834	64818	77.72	188	30229	160.79	1270	99249	78.15
29 IRJA TIMUR	596	6455	10.83	2065	161766	78.34	500	79970	159.94	3161	248191	78.52
Total	**25658**	**471003**	**18.36**	**327877**	**29152766**	**88.91**	**137,961**	**21236866**	**153.93**	**491496**	**50860635**	**103.48**

Table B.3e
Total Number of Common Census Blocks and HouseHolds by Group Number of HouseHolds and Provinces

Urban

Province	<= 40			41 - 120			>=121			Total		
	CB	HH 0	Average per CB	CB	HH 0	Average per CB	0 CB	HH 0	Average CB	CB	HH 0	Average CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 D I ACEH	726	1,892	2.61	911	79,259	87.00	338	52,359	154.91	1,975	133,510	67.60
2 SUMATERA UTARA	1,087	16,230	14.93	7,090	604,448	85.25	2,862	453,264	158.37	11,039	1,073,942	97.29
3 SUMATER A BARAT	224	4,317	19.27	2,216	184,319	83.18	635	94,514	148.84	3,075	283,150	92.08
4 RIAU	809	12,360	15.28	2,917	240,305	82.38	1,385	248,318	179.29	5,111	500,983	98.02
5 JAMBI	194	3,158	16.28	1,178	99,507	84.47	363	57,647	158.81	1,735	160,312	92.40
6 SUMATER A SELATAN	776	8,178	10.54	3,775	320,038	84.78	1,456	232,206	159.48	6,007	560,422	93.29
7 BENGK ULU	79	746	9.44	544	47,374	87.08	315	49,244	156.33	938	97,364	103.80
8 LAMPUNG	383	5,411	14.13	2,125	181,712	85.51	874	136,312	155.96	3,382	322,435	95.63
9 DKI JAKARTA	1.,499	30,606	20.42	10,708	892,715	83.37	7,155	1,325,067	185.19	19,362	2,248,388	116.12
10 JAWA BARAT	7,011	71,704	10.23	30,458	2,636,035	86.55	18,279	3,017,771	165.09	55,748	5,725,510	102.70
11 JAWA TENGAH	4,964	13,691	2.76	23,808	2,159,919	90.72	6,269	894,222	142.64	35,041	3,067,832	87.55
12 D I YOGYAKARTA	132	3,044	23.06	3,598	321,320	89.31	1,508	240,483	159.47	5,238	564,847	107.84
13 JAWA TIMUR	5,813	32,938	5.67	26,699	2,366,516	88.64	9,272	1,420,970	153.25	41,784	3,820,424	91.43
14 BALI	302	4,326	14.32	2,299	196,349	85.41	1,063	185,005	174.04	3,664	385,680	105.26
15 NUSA TEN GGARA BARAT	292	2,170	7.43	1,985	179,322	90.34	1,065	160,497	150.70	3,342	341,989	102.33
16 NUSA TENGGARA TIMUR	61	677	11.10	664	59,467	89.56	381	59,736	156.79	1,106	119,880	108.39
17 KALIMANTAN BAR AT	185	2,306	12.46	1,590	134,532	84.61	521	84,166	161.55	2,296	221,004	96.26
18 KALIMANTAN TEN GAH	228	2,358	10.34	703	59,978	85.32	396	71,091	179.52	1,327	133,427	100.55
19 KALIMANTAN SELATAN	325	4,382	13.48	1,968	165,683	84.19	711	111,602	156.96	3,004	281,667	93.76
20 KALIMANTAN TIMUR	710	10,956	15.43	2,641	215,909	81.75	750	119,520	159.36	4,101	346,385	84.46
21 SULAWESI UTARA	333	3,712	11.15	1,659	140,158	84.48	719	114,333	159.02	2,711	258,203	95.24
22 SULAWESI TEN GAH	76	872	11.47	680	57,556	84.64	208	31,165	149.83	964	89,593	92.94
23 SULAWESI SE LATAN	512	8,512	16.63	3,712	313,953	84.58	1,228	192,151	156.47	5,452	514,616	94.39
24 SULAWESI TENGGARA	64	665	10.39	426	38,225	89.73	264	44,933	170.20	754	83,823	111.17
25 MALUKU	234	1,997	8.53	407	33,323	81.87	188	31,323	166.61	829	66,643	80.39
26 MALUKU UTARA	76	249	3.28	189	15,991	84.61	141	23,194	164.50	406	39,434	97.13
27 IRJA BARAT	63	257	4.08	228	20,237	88.76	170	27,313	160.66	461	47,807	103.70
28 IRJA TENGAH	52	608	11.69	156	12,691	81.35	51	9,141	179.24	259	22,440	86.94
29 IRJA TIMUR	126	1,218	9.67	421	33,379	79.29	152	24,594	161.80	699	59,191	84.68
Total	**27,336**	**249,540**	**9.13**	**135,755**	**11,810,220**	**87.00**	**58,719**	**9,512,141**	**161.99**	**221,810**	**21,571,901**	**97.25**

Table B.3c
Total Number of Common Census Blocks and HouseHolds by Group Number of HouseHolds and Provinces

Rural

Province	<= 40			41 - 120			>=121			Total		
	CB	HH	Average	CB	HH	Average	CB	HH	Average	CB	HH	Average

		0	per CB		0	per CB	0	0	CB		0	CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 D I ACEH	3470	8784	2.53	2523	208481	82.63	770	122002	158.44	6763	339267	50.17
2 SUMATERA UTARA	1092	23781	21.78	9333	804081	86.15	4213	634649	150.64	14638	1462511	99.91
3 SUMATERA BARAT	285	8431	29.58	6019	511150	84.92	1494	219661	147.03	7798	739242	94.80
4 RIAU	292	7900	27.05	3818	328087	85.93	1936	316137	163.29	6046	652124	107.86
5 JAMBI	251	5553	22.12	2914	251274	86.23	1142	173482	151.91	4307	430309	99.91
6 SUMATERA SELATAN	809	16413	20.29	7286	626000	85.92	3071	470951	153.35	11166	1113364	99.71
7 BENGKULU	97	2275	23.45	1516	132751	87.57	723	112749	155.95	2336	247775	106.07
8 LAMPUNG	381	5957	15.64	7810	718187	91.96	3657	545512	149.17	11848	1269656	107.16
9 DKI JAKARTA	0	0	0.00	0	0	0.00	0	0	0	0	0	0.00
10 JAWA BARAT	1306	23403	17.92	28358	2657432	93.71	20448	3109449	152.07	50112	5790284	155.55
11 JAWA TENGAH	574	9449	16.46	34900	3285791	94.15	10724	1484777	138.45	46198	4780017	103.47
12 D I YOGYAKARTA	24	654	27.25	2739	254599	92.95	750	103868	138.49	3513	359121	102.23
13 JAWA TIMUR	1057	23339	22.08	38463	3558953	92.53	14566	2101922	144.30	54086	5684214	105.10
14 BALI	70	2058	29.40	2910	256706	88.22	1049	156959	149.63	4029	415723	103.18
15 NUSA TENGGARA BARAT	139	3097	22.28	4054	372357	91.85	1848	269293	145.72	6041	644747	106.73
16 NUSA TENGGARA TIMUR	164	4226	25.77	4953	437571	88.34	1622	232344	143.25	6739	674141	100.04
17 KALIMANTAN BARAT	561	12546	22.36	4532	377342	83.26	1525	233497	153.11	6618	623385	94.20
18 KALIMANTAN TENGAH	362	9194	25.40	2160	176793	81.85	927	150875	162.76	3449	336862	97.67
19 KALIMANTAN SELATAN	363	7874	21.69	3995	336943	84.34	1074	159671	148.67	5432	504488	92.87
20 KALIMANTAN TIMUR	550	13036	23.70	2182	173542	79.53	512	79190	154.67	3244	265768	81.93
21 SULAWESI UTARA	149	2889	19.39	3098	278523	89.90	1476	217559	147.40	4723	498971	105.65
22 SULAWESI TENGAH	261	3625	13.89	2539	224769	88.53	986	144931	146.99	3786	373325	98.61
23 SULAWESI SELATAN	358	9553	26.68	9259	816985	88.24	2874	417878	145.40	12491	1244416	99.63
24 SULAWESI TENGGARA	87	2578	29.63	2093	183236	87.55	871	130524	149.86	3051	316338	103.68
25 MALUKU	383	5380	14.05	1311	106810	81.47	383	57743	150.77	2077	169933	81.82
26 MALUKU UTARA	517	3538	6.84	800	65440	81.80	221	33165	150.07	1538	102143	66.41
27 IRJA BARAT	278	5573	20.05	706	54167	76.72	110	17926	162.96	1094	77666	70.99
28 IRJA TENGAH	216	3640	16.85	683	52499	76.87	148	23726	160.31	1047	79865	76.28
29 IRJA TIMUR	504	5241	10.40	1649	128789	78.10	352	56021	159.15	2505	190051	75.87
Total	**14600**	**229987**	**15.75**	**192.603**	**17.379.258**	**90.23**	**79472**	**11776461**	**148.18**	**286675**	**29385706**	**102.51**

Table B.3d
Total Number of Census Blocks and HouseHolds by Group Number of HouseHolds and Provinces

Urban + Rural

Province	<= 40			41 - 120			>=121			Total		
	CB	HH 0	Average per CB	CB	HH 0	Average per CB	CB	HH 0	Average CB	CB	HH 0	Average CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 D I ACEH	6985	10676	1.53	3439	288024	83.75	1109	174778	157.60	11533	473478	41.05
2 SUMATERA UTARA	6119	40177	6.57	16440	1409755	85.75	7080	1088847	153.79	29639	2538779	85.66
3 SUMATER A BARAT	1426	12901	9.05	8235	695469	84.45	2130	314315	147.57	11791	1022685	86.73
4 RIAU	3314	20263	6.11	6737	568518	84.39	3324	565024	169.98	13375	1153805	86.27
5 JAMBI	2163	8835	4.08	4098	351257	85.71	1510	232630	154.06	7771	592722	76.27
6 SUMATER A SELATAN	6792	24617	3.62	11068	946691	85.53	4528	703336	155.33	22388	1674644	74.80
7 BENGK ULU	1216	3081	2.53	2062	180318	87.45	1038	161993	156.06	4316	345392	80.03
8 LAMPUNG	4084	11520	2.82	9941	900408	90.58	4535	682811	150.56	18560	1594739	85.92
9 DKI JAKARTA	1499	30606	20.42	10708	892715	83.37	7155	1325067	185.19	19362	2248388	116.12
10 JAWA BARAT	18264	95507	5.23	58823	5293974	90.00	38740	6129555	158.22	115827	11519036	99.45
11 JAWA TENGAH	19870	23919	1.20	58715	5446118	92.76	16993	2378999	140.00	95578	7849036	82.12
12 D I YOGYAKARTA	287	3698	12.89	6337	575919	90.88	2258	344351	152.50	8882	923968	104.03
13 JAWA TIMUR	19167	59233	3.09	65258	5932020	90.90	23863	3528396	147.86	108288	9519649	87.91
14 BALI	754	6399	8.49	5210	453153	86.98	2112	341964	161.91	8076	801516	99.25
15 NUSA TEN GGARA BARAT	1391	5332	3.83	6043	551938	91.34	2913	429790	147.54	10347	987060	95.40
16 NUSA TENGGARA TIMUR	1622	5124	3.16	5618	497155	88.49	2003	292080	145.82	9243	794359	85.94
17 KALIMANTAN BAR AT	2564	14919	5.82	6128	512354	83.61	2048	317992	155.27	10740	845265	78.70
18 KALIMANTAN TEN GAH	3261	11552	3.54	2863	236771	82.70	1323	221966	167.77	7447	470289	63.15
19 KALIMANTAN SELATAN	2754	12385	4.50	5966	502848	84.29	1787	271581	151.98	10507	786814	74.88
20 KALIMANTAN TIMUR	3814	24014	6.30	4829	389925	80.75	1268	199785	157.56	9911	613724	61.92
21 SULAWESI UTARA	2149	6678	3.11	4763	419152	88.00	2196	332016	151.19	9108	757846	83.21
22 SULAWESI TEN GAH	2125	4568	2.15	3222	282564	87.70	1196	176440	147.53	6543	463572	70.85
23 SULAWESI SE LATAN	2494	18165	7.28	12980	1131633	87.18	4105	610544	148.73	19579	1760342	89.91
24 SULA WESI TENGGARA	1701	3263	1.92	2522	221670	87.89	1136	175695	154.66	5359	400628	74.76
25 MALUKU	1887	7440	3.94	1718	140133	81.57	572	89214	155.97	4177	236787	56.69
26 MALUKU UTARA	1697	3787	2.23	991	81560	82.30	363	56651	156.06	3051	141998	46.54
27 IRJA BARAT	1716	5838	3.40	936	74613	79.71	280	45239	161.57	2932	125690	42.87
28 IRJA TENGAH	639	4248	6.65	839	65190	77.70	200	33177	165.88	1678	102615	61.15
29 IRJA TIMUR	2391	6459	2.70	2070	162168	78.34	504	80615	159.95	4965	249242	50.20
Total	**124145**	**485204**	**3.91**	**328559**	**29204013**	**88.89**	**138269**	**21304851**	**154.08**	**590973**	**50994068**	**86.29**

Table B.3b
Total Number of Census Blocks and HouseHolds by Group Number of HouseHolds and Provinces

Urban

Province	<= 40			41 - 120			>=121			Total		
	CB	HH 0	Average per CB	CB	HH 0	Average per CB	CB	HH 0	Average CB	CB	HH 0	Average CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 D I ACEH	588	1,792	3.05	906	78,913	87.10	334	51,679	154.73	1,828	132,384	72.42
2 SUMATERA UTARA	633	15,892	25.11	7,058	602,041	85.30	2,852	450,697	158.03	10.543	1,068.630	101.36
3 SUMATER A BARAT	144	4,197	29.15	2,212	183,959	83.16	633	94,124	148.70	2,989	282,280	94.44
4 RIAU	567	11,965	21.10	2,894	238,394	82.38	1,381	247,453	179.18	4,842	497,812	102.81
5 JAMBI	128	3,136	24.50	1,169	98,760	84.48	361	57,321	158.78	1,658	159,217	96.03
6 SUMATER A SELATAN	462	8,033	17.39	3,758	318,797	84.83	1,450	231,061	159.35	5,670	557,891	98.39
7 BENGK ULU	36	740	20.56	542	47,255	87.19	314	49,041	156.18	892	97,036	108.78
8 LAMPUNG	214	5,265	24.60	2,117	181,151	85.57	873	136,106	155.91	3,204	322,522	100.66
9 DKI JAKARTA	1,218	30,119	24.73	10,674	890,082	83.39	7,146	1,323,212	185.17	19,038	2,243,413	117.84
10 JAWA BARAT	3,144	70,816	22.52	30,369	2,628,521	86.55	18,221	3,004,572	164.90	51,734	5,703,909	110.25
11 JAWA TENGAH	591	12,394	20.97	23,765	2,157,021	90.76	6,244	887,121	142.08	30,600	3,056,536	99.89
12 D I YOGYAKARTA	108	2,965	27.45	3,583	320,140	89.35	1,505	239,749	159.30	5,196	562,854	108.32
13 JAWA TIMUR	1,297	30,136	23.24	26,609	2,360,175	88.70	9,223	1,409,366	152.81	37,129	3,799,677	102.34
14 BALI	147	4,214	28.67	2,293	195,922	85.44	1,059	184,287	174.02	3,499	384,423	109.87
15 NUSA TEN GGARA BARAT	85	1,727	20.32	1,976	178,729	90.45	1,060	159,769	150.73	3,121	340,225	109.01
16 NUSA TENGGARA TIMUR	14	389	27.79	654	58,804	89.91	378	59,303	156.89	1,046	118,496	113.28
17 KALIMANTAN BAR AT	87	2,207	25.37	1,583	133,998	84.65	519	83,891	161.64	2,189	220,096	100.55
18 KALIMANTAN TEN GAH	99	2,320	23.43	701	59,825	85.34	396	71,091	179.52	1,196	133,236	111.40
19 KALIMANTAN SELATAN	209	4,296	20.56	1,962	165,230	84.22	705	110,616	156.90	2,876	280,142	97.41
20 KALIMANTAN TIMUR	433	10,825	25.00	2,625	214,593	81.75	746	118,855	159.32	3,804	344,273	90.50
21 SULAWESI UTARA	149	3,584	24.05	1,649	139,297	84.47	715	113,427	158.64	2,513	256,308	101.99
22 SULAWESI TEN GAH	31	847	27.32	675	57,250	84.81	207	30,781	148.70	913	88,878	97.35
23 SULAWESI SE LATAN	370	8,322	22.49	3,692	312,329	84.60	1,223	190,929	156.12	5,285	511,580	96.80
24 SULAWESI TENGGARA	23	635	27.61	426	38,225	89.73	264	44,933	170.20	713	83,793	117.52
25 MALUKU	114	1,952	17.12	405	33,147	81.84	185	30,868	166.85	704	65,967	93.70
26 MALUKU UTARA	25	218	8.72	189	15,991	84.61	141	23,194	164.50	355	39,403	110.99
27 IRJA BARAT	18	254	14.11	221	19,663	88.97	166	26,507	159.68	405	46,424	114.63
28 IRJA TENGAH	32	562	17.56	151	12,319	81.58	40	6,503	162.57	223	19,384	86.92
29 IRJA TIMUR	92	1,214	13.20	416	32,977	79.27	148	23,949	161.82	656	58,140	88.63
Total	**11,058**	**241,016**	**21.80**	**135,274**	**11,773,508**	**87.03**	**58,489**	**9,460,405**	**161.75**	**204,821**	**21,474,929**	**104.85**

Table B.3f.
Total Number of Census Blocks and HouseHolds by Group Number of HouseHolds and Provinces

Rural

Province	<= 40			41 - 120			>=121			Total		
	CB	HH 0	Average per CB	CB	HH 0	Average per CB	CB	HH 0	Average CB	CB	HH 0	Average CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 D I ACEH	6,259	8,784	1.40	2,528	208,765	82.58	771	122,419	158.78	9,558	339,968	35.57
2 SUMATERA UTARA	5,032	23,947	4.76	9,350	805,307	86.13	4,218	635,583	150.68	18,600	1,464,837	78.75
3 SUMATERA BARAT	1,202	8,584	7.14	6,019	511,150	84.92	1,495	219,801	147.02	8,716	739,535	84.85
4 RIAU	2,505	7,903	3.15	3,820	328,213	85.92	1,939	316,706	163.33	8,264	652,822	79.00
5 JAMBI	1,969	5,677	2.88	2,920	251,750	86.22	1,147	174,983	152.56	6,036	432,410	71.64
6 SUMATERA SELATAN	6,016	16,439	2.73	7,293	626,653	85.93	3,072	471,130	153.36	16,381	1,114,222	68.02
7 BENGKULU	1,137	2,335	2.05	1,518	132,944	87.58	723	112,749	155.95	3,378	248,028	73.42
8 LAMPUNG	3,701	6,109	1.65	7,816	718,696	91.95	3,661	546,499	149.28	15,178	1,271,304	83.76
9 DKI JAKARTA	0	0	0	0	0	0	0	0	0.00	0	0	0
10 JAWA BARAT	11,253	23,803	2.12	28,365	2,657,939	93.70	20,461	3,111,784	152.08	60,079	5,793,526	96.43
11 JAWA TENGAH	14,906	10,228	0.69	34,907	3,286,199	94.14	10,724	1,484,777	138.45	60,537	4,781,204	78.98
12 D I YOGYAKARTA	155	654	4.22	2,739	254,599	92.95	750	103,868	138.49	3,644	359,121	98.55
13 JAWA TIMUR	13,354	26,295	1.97	38,559	3,565,504	92.47	14,591	2,107,426	144.43	66,504	5,699,225	85.70
14 BALI	452	2,073	4.59	2,911	256,804	88.22	1,049	156,959	149.63	4,412	415,836	94.25
15 NUSA TENGGARA BARAT	1,099	3,162	2.88	4,058	372,616	91.82	1,848	269,293	145.72	7,005	645,071	92.09
16 NUSA TENGGARA TIMUR	1,561	4,447	2.85	4,954	437,688	88.35	1,622	232,344	143.25	8,137	674,479	82.89
17 KALIMANTAN BARAT	2,379	12,613	5.30	4,538	377,822	83.26	1,527	233,826	153.13	8,444	624,261	73.93
18 KALIMANTAN TENGAH	3,033	9,194	3.03	2,160	176,793	81.85	927	150,875	162.76	6,120	336,862	55.04
19 KALIMANTAN SELATAN	2,429	8,003	3.29	3,998	337,165	84.33	1,076	159,979	148.68	7,503	505,147	67.33
20 KALIMANTAN TIMUR	3,104	13,058	4.21	2,188	174,016	79.53	518	80,265	154.95	5,810	267,339	46.01
21 SULAWESI UTARA	1,816	2,966	1.63	3,104	278,994	89.88	1,477	217,683	147.38	6,397	499,643	78.11
22 SULAWESI TENGAH	2,049	3,696	1.80	2,542	225,008	88.52	988	145,275	147.04	5,579	373,979	67.03
23 SULAWESI SELATAN	1,982	9,653	4.87	9,268	817,680	88.23	2,877	418,393	145.43	14,127	1,245,726	88.18
24 SULAWESI TENGGARA	1,637	2,598	1.59	2,096	183,445	87.52	872	130,762	149.96	4,605	316,805	68.80
25 MALUKU	1,653	5,443	3.29	1,311	106,810	81.47	384	57,891	150.76	3,348	170,144	50.82
26 MALUKU UTARA	1,621	3,538	2.18	802	65,569	81.76	222	33,457	150.71	2,645	102,564	38.78
27 IRJA BARAT	1,653	5,581	3.38	708	54,376	76.80	110	17,926	162.96	2,471	77,883	31.52
28 IRJA TENGAH	587	3,640	6.20	683	52,499	76.87	149	24,036	161.32	1,419	80,175	56.50
29 IRJA TIMUR	2,265	5,241	2.31	1,649	128,789	78.10	352	56,021	159.15	4,266	190,051	44.55
Total	**96,809**	**235,664**	**2.43**	**192,804**	**17,393,793**	**90.21**	**79,550**	**11,792,710**	**148.24**	**369,163**	**29,422,167**	**79.70**

Table B.4
Number of Kab/Kota by grouping number of Common Census Blocks

ITEM	Number of Common Census Blocks							Total
	<= 500	501-999	1000-1499	1500-1999	2000-2499	2500-2999	>=3000	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Urban								
1. No. of Kab/Kota	233	42	26	9	8	5	13	336
2. Number of CB	40,859	27,962	31,506	14,533	17,777	13,542	58,642	204,821
3. Average per CB	175.36	665.76	1,211.77	1,614.78	2,222.12	2,708.40	4,510.92	609.59
Rural								
1. No. of Kab/Kota	125	77	45	30	19	8	9	313
2. Number of CB	29,133	56,483	55,295	51,370	41,167	21,596	31,631	286,675
3. Average per CB	233.06	733.55	1,228.78	1,712.33	2,166.68	2,699.50	3,514.56	915.89
Urban + Rural								
1. No. of Kab/Kota	89	81	49	35	28	22	36	340
2. Number of CB	30,047	56,505	58,278	60,824	62,686	60,127	163,029	491,496
3. Average per CB	337.61	697.59	1,189.35	1,737.83	2,238.79	2,733.05	4,528.58	1,445.58

Table B.4a
Number of Kab/Kota by grouping number of Census Blocks

ITEM	Number of Census Blocks							Total
	<= 500	501-999	1000-1499	1500-1999	2000-2499	2500-2999	>=3000	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Urban								
1. No. of Kab/Kota	221	49	23	17	6	6	14	336
2. Number of CB	38,318	32,558	27,438	28,382	13,818	16,570	64,726	221,810
3. Average per CB	173.38	664.45	1,192.96	1,669.53	2,303.00	2,761.67	4,623.29	660.15
Rural								
1. No. of Kab/Kota	94	77	46	33	26	15	22	313
2. Number of CB	20,177	56,790	56,946	57,100	58,416	40,855	78,879	369,163
3. Average per CB	214.65	737.53	1,237.96	1,730.30	2,246.77	2,723.67	3,585.41	1,179.43
Urban + Rural								
1. No. of Kab/Kota	64	81	54	28	32	21	60	340
2. Number of CB	22,446	58,663	65,435	48,450	71,097	57,063	267,819	590,973
3. Average per CB	350.72	724.23	1,211.76	1,730.36	2,221.78	2,717.29	4,463.65	1,738.16

Table B.5a:
No. of Kab/Kota by grouping number of Common Census Blocks by provinces

Urban + Rural

Province	Grouping no. of Common Census Blocks								Total	
	<= 500		501 - 999		1000 - 1999		>= 2000			
	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
01. Dista Aceh	5	291.00	4	743.00	3	1388.00	0	0.00	12	715.92
02. Sumatera Utara	4	329.00	7	659.14	5	1766.60	3	3472.67	19	1,325.32
03. Sumatera Barat	6	165.83	4	909.25	5	1231.00	0	0.00	15	719.13
04. Riau	4	304.50	6	614.17	5	1197.00	0	0.00	15	725.87
05. Jambi	5	475.60	4	633.75	1	1052.00	0	0.00	10	596.50
06. Sumatera Selatan	1	305.00	1	557.00	4	1417.00	4	2576.50	10	1683.60
07. Bengkulu	0	0.00	4	807.00	0	0.00	0	0.00	4	807.00
08. Lampung	1	232.00	2	768.00	4	1565.00	3	2341.33	10	1505.20
09. DKI Jakarta	0	0.00	0	0.00	0	0.00	5	3807.60	5	3807.60
10. Jawa Barat	0	0.00	3	666.67	2	1632.50	23	4199.17	28	3637.36
11. Jawa Tengah	2	351.50	2	591.50	11	1757.36	20	2779.05	35	2194.23
12. DI Yogyakarta	0	0.00	1	978.00	2	1508.50	2	2357.00	5	1741.80
13. Jawa Timur	4	381.00	2	551.00	8	1697.00	23	3261.43	37	2465.27
14. Bali	2	413.00	4	790.50	3	1180.00	0	0.00	9	836.44
15. Nusa Tenggara Barat	1	414.00	1	693.00	4	1427.75	1	2344.00	7	1308.86
16. Nusa Tenggara Timur	7	383.57	6	653.50	1	1179.00	0	0.00	14	556.07
17. Kalimantan Barat	1	450.00	2	656.00	6	1174.17	0	0.00	9	978.56
18. Kalimantan Tengah	3	437.33	1	694.00	2	1319.50	0	0.00	6	774.17
19. Kalimantan Selatan	3	402.67	5	688.20	3	1219.67	0	0.00	11	755.27
20. Kalimantan Timur	7	281.14	2	615.50	3	1283.00	0	0.00	12	587.33
21. Sulawesi Utara	3	368.33	1	735.00	3	1095.33	1	2110.00	8	904.50
22. Sulawesi Tengah	4	333.00	3	618.33	1	1512.00	0	0.00	8	587.38
23. Sulawesi Selatan	8	343.13	12	730.67	3	1183.33	1	2713.00	24	740.67
24. Sulawesi Tenggara	1	381.00	2	635.50	2	1056.00	0	0.00	5	752.80
25. Maluku	4	388.00	0	0.00	1	1229.00	0	0.00	5	556.20
26. Maluku Utara	2	300.00	0	0.00	1	1293.00	0	0.00	3	631.00
27. Irja Barat	3	317.33	1	547.00	0	0.00	0	0.00	4	374.75
28. Irja Tengah	5	254.00	0	0.00	0	0.00	0	0.00	5	254.00
29. Irja Timur	3	373.67	1	848.00	1	1192.00	0	0.00	5	632.20
INDONESIA	**89**	**337.61**	**81**	**697.59**	**84**	**1417.88**	**86**	**3323.74**	**340**	**1445.58**

Table B.5b
No. of Kab/Kota by grouping number of Common Census Blocks by provinces

Urban

Province	Grouping no. of Common Census Blocks								Total	
	<= 500		501 - 999		1000 - 1999		>= 2000			
	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
01. Dista Aceh	12	152.33	0	0.00	0	0.00	0	0.00	12	152.33
02. Sumatera Utara	15	209.07	2	534.50	0	0.00	2	3169.00	19	554.89
03. Sumatera Barat	14	103.71	0	0.00	1	1537.00	0	0.00	15	199.27
04. Riau	13	171.77	0	0.00	2	1304.50	0	0.00	15	322.80
05. Jambi	9	72.89	0	0.00	1	1002.00	0	0.00	10	165.80
06. Sumatera Selatan	9	327.00	0	0.00	0	0.00	1	2727.00	10	567.00
07. Bengkulu	4	223.00	0	0.00	0	0.00	0	0.00	4	223.00
08. Lampung	8	178.63	0	0.00	1	1775.00	0	0.00	9	356.00
09. DKI Jakarta	0	0.00	0	0.00	0	0.00	5	3807.60	5	3807.60
10. Jawa Barat	2	261.00	10	735.80	7	1309.71	9	3854.00	28	1847.64
11. Jawa Tengah	11	395.73	13	669.23	9	1346.11	2	2716.00	35	874.29
12. DI Yogyakarta	2	136.50	0	0.00	2	1390.50	1	2142.00	5	1039.20
13. Jawa Timur	16	357.31	10	656.90	6	1315.50	5	3390.00	37	1003.49
14. Bali	6	217.00	2	519.00	1	1159.00	0	0.00	9	388.78
15. Nusa Tenggara Barat	4	228.75	3	735.33	0	0.00	0	0.00	7	445.86
16. Nusa Tenggara Timur	14	74.71	0	0.00	0	0.00	0	0.00	14	74.71
17. Kalimantan Barat	8	137.88	0	0.00	1	1086.00	0	0.00	9	243.22
18. Kalimantan Tengah	6	199.33	0	0.00	0	0.00	0	0.00	6	199.33
19. Kalimantan Selatan	10	142.50	0	0.00	1	1451.00	0	0.00	11	261.45
20. Kalimantan Timur	9	154.89	0	0.00	2	1205.00	0	0.00	11	317.00
21. Sulawesi Utara	6	158.67	1	508.00	1	1053.00	0	0.00	8	314.13
22. Sulawesi Tengah	7	57.00	1	514.00	0	0.00	0	0.00	8	114.13
23. Sulawesi Selatan	23	114.65	0	0.00	0	0.00	1	2648.00	24	220.21
24. Sulawesi Tenggara	5	142.60	0	0.00	0	0.00	0	0.00	5	142.60
25. Maluku	5	140.80	0	0.00	0	0.00	0	0.00	5	140.80
26. Maluku Utara	3	118.33	0	0.00	0	0.00	0	0.00	3	118.33
27. Irja Barat	4	101.25	0	0.00	0	0.00	0	0.00	4	101.25
28. Irja Tengah	4	55.75	0	0.00	0	0.00	0	0.00	4	44.60
29. Irja Timur	4	164.00	0	0.00	0	0.00	0	0.00	4	131.20
INDONESIA	**233**	**175.36**	**42**	**665.76**	**35**	**1315.40**	**26**	**3460.04**	**336**	**609.59**

Table B.5c
No. of Kab/Kota by grouping number of Common Census Blocks by provinces

Rural

Province	Grouping no. of Common Census Blocks								Total	
	<= 500		501 - 999		1000 - 1999		>= 2000			
	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
01. Dista Aceh	5	260.60	4	761.75	2	1206.50	0	0.00	11	614.82
02. Sumatera Utara	3	181.33	5	628.80	6	1491.50	1	2001.00	15	975.87
03. Sumatera Barat	6	104.33	7	878.29	1	1024.00	0	0.00	14	557
04. Riau	11	261.00	3	689.67	1	1106.00	0	0.00	15	403.07
05. Jambi	7	362.57	3	589.67	0	0.00	0	0.00	10	430.7
06. Sumatera Selatan	3	104.33	1	920.00	3	1145.33	3	2165.67	10	1116.6
07. Bengkulu	1	14.00	3	774.00	0	0.00	0	0.00	4	584
08. Lampung	2	34.00	3	793.33	3	1653.33	2	2220.00	10	1184.8
09. DKI Jakarta	0	0.00	0	0.00	0	0.00	0	0.00	0	0
10. Jawa Barat	5	50.40	1	965.00	7	1795.29	12	3027.33	25	2004.48
11. Jawa Tengah	3	81.67	3	704.67	18	1399.78	8	2330.38	32	1443.69
12. DI Yogyakarta	1	433.00	2	684.50	1	1711.00	0	0.00	4	878.25
13. Jawa Timur	3	50.00	1	535.00	18	1495.33	10	2648.50	32	1690.19
14. Bali	5	335.20	3	784.33	0	0.00	0	0.00	8	503.63
15. Nusa Tenggara Barat	1	340.00	2	825.00	3	1350.33	0	0.00	6	1006.83
16. Nusa Tenggara Timur	9	329.67	4	673.25	1	1079.00	0	0.00	14	481.36
17. Kalimantan Barat	2	441.50	3	795.33	3	1116.33	0	0.00	8	827.25
18. Kalimantan Tengah	4	321.75	1	980.00	1	1182.00	0	0.00	6	574.83
19. Kalimantan Selatan	5	246.80	6	699.67	0	0.00	0	0.00	11	493.82
20. Kalimantan Timur	9	197.11	2	735.00	0	0.00	0	0.00	11	294.91
21. Sulawesi Utara	4	150.75	3	839.33	1	1602.00	0	0.00	8	590.38
22. Sulawesi Tengah	5	255.80	2	542.00	1	1423.00	0	0.00	8	473.25
23. Sulawesi Selatan	12	295.17	11	697.36	1	1278.00	0	0.00	24	520.46
24. Sulawesi Tenggara	1	74.00	3	648.67	1	1031.00	0	0.00	5	610.20
25. Maluku	4	250.50	0	0.00	1	1075.00	0	0.00	5	415.40
26. Maluku Utara	2	180.50	0	0.00	1	1177.00	0	0.00	3	512.67
27. Irja Barat	4	273.50	0	0	0	0.00	0	0.00	4	273.50
28. Irja Tengah	5	209.40	0	0	0	0.00	0	0.00	5	209.40
29. Irja Timur	3	204.33	1	752	1	1140.00	0	0.00	5	501.00
INDONESIA	**125**	**233.06**	**77**	**733.55**	**75**	**1422.20**	**36**	**2622.06**	**313**	**915.89**

Table B.5d
No. of Kab/Kota by grouping number of Census Blocks by provinces

Urban + Rural

Province	Grouping no. of Census Blocks								Total	
	<= 500		501 - 999		1000 - 1999		>= 2000			
	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
01. Dista Aceh	2	116.50	4	581.50	6	1495.67	0	0.00	12	961.08
02. Sumatera Utara	3	277.00	6	672.50	3	1238.33	7	3008.29	19	1559.95
03. Sumatera Barat	6	177.67	2	940.50	7	1263.43	0	0.00	15	786.07
04. Riau	3	352.33	6	755.00	6	1298.00	0	0.00	15	891.67
05. Jambi	0	0.00	8	682.25	2	1156.50	0	0.00	10	777.10
06. Sumatera Selatan	1	311.00	1	714.00	1	1672.00	7	2813.00	10	2238.80
07. Bengkulu	0	0.00	1	517.00	3	1266.33	0	0.00	4	1079.00
08. Lampung	1	241.00	1	847.00	3	1441.00	5	2629.80	10	1856.00
09. DKI Jakarta	0	0.00	0	0	0	0.00	5	3872.40	5	3872.40
10. Jawa Barat	0	0.00	3	713.67	1	1753.00	24	4663.88	28	4136.68
11. Jawa Tengah	2	374.00	2	635.00	1	1549.00	30	3067.03	35	2,730.80
12. DI Yogyakarta	0	0.00	0	0.00	3	1379.67	2	2371.50	5	1,776.40
13. Jawa Timur	4	408.50	2	623.50	6	1766.50	25	3792.32	37	2,926.70
14. Bali	2	432.00	4	860.50	3	1256.67	0	0.00	9	897.33
15. Nusa Tenggara Barat	1	475.00	1	715.00	4	1659.75	1	2518.00	7	1478.14
16. Nusa Tenggara Timur	5	417.80	7	686.86	2	1173.00	0	0.00	14	660.21
17. Kalimantan Barat	0	0.00	3	746.67	6	1416.67	0	0.00	9	1193.33
18. Kalimantan Tengah	2	487.50	1	723	1	1201.00	2	2274.00	6	1241.17
19. Kalimantan Selatan	1	350.00	5	750.80	5	1280.60	0	0.00	11	955.18
20. Kalimantan Timur	6	389.50	1	802.00	5	1354.40	0	0.00	12	825.92
21. Sulawesi Utara	2	377.50	2	696.00	3	1303.67	1	3050.00	8	1138.5
22. Sulawesi Tengah	3	367.00	3	769.00	1	1052.00	1	2083.00	8	817.88
23. Sulawesi Selatan	8	370.63	11	770.36	4	1343.00	1	2768.00	24	815.79
24. Sulawesi Tenggara	1	396.00	2	806.5	2	1675.00	0	0.00	5	1,071.80
25. Maluku	1	406.00	3	708.33	1	1646.00	0	0.00	5	835.40
26. Maluku Utara	1	304.00	1	555.00	0	0.00	1	2192.00	3	1,017.00
27. Irja Barat	2	390.50	0	0.00	2	1075.50	0	0.00	4	733.00
28. Irja Tengah	5	335.60	0	0.00	0	0.00	0	0.00	5	335.60
29. Irja Timur	2	424.50	1	747.00	1	1304.00	1	2065.00	5	993.00
INDONESIA	**64**	**350.72**	**81**	**724.23**	**82**	**1388.84**	**113**	**3504.24**	**340**	**1738.16**

Table B.5e
No. of Kab/Kota by grouping number of Census Blocks by provinces

Urban

Province	Grouping no. of Census Blocks								Total	
	<= 500		501 - 999		1000 - 1999		>= 2000			
	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
01. Dista Aceh	11	133.45	1	507.00	0	0.00	0	0.00	12	164.58
02. Sumatera Utara	15	225.00	2	568.00	0	0.00	2	3264.00	19	581.00
03. Sumatera Barat	14	109.36	0	0.00	1	1544.00	0	0.00	15	205.00
04. Riau	12	156.50	1	528.00	2	1352.50	0	0.00	15	340.73
05. Jambi	9	80.00	0	0.00	1	1015.00	0	0.00	10	173.50
06. Sumatera Selatan	8	343.13	1	502.00	0	0.00	1	2760.00	10	600.70
07. Bengkulu	3	145.00	1	503.00	0	0.00	0	0.00	4	234.50
08. Lampung	8	195.88	0	0.00	1	1815.00	0	0.00	9	375.78
09. DKI Jakarta	0	0.00	0	0.00	0	0	5	3872.40	5	3872.40
10. Jawa Barat	2	273.00	8	761.00	9	1341.67	9	4115.44	28	1991.00
11. Jawa Tengah	7	401.00	15	674.73	11	1454.73	2	3055.50	35	1001.17
12. DI Yogyakarta	2	143.00	0	0.00	2	1399.00	1	2154.00	5	1047.60
13. Jawa Timur	13	364.31	12	670.33	7	1505.57	5	3693.00	37	1129.30
14. Bali	5	179.00	3	536.00	1	1161.00	0	0.00	9	407.11
15. Nusa Tenggara Barat	4	249.00	3	782.00	0	0.00	0	0.00	7	477.43
16. Nusa Tenggara Timur	14	79.00	0	0.00	0	0.00	0	0.00	14	79.00
17. Kalimantan Barat	8	150.13	0	0.00	1	1095.00	0	0.00	9	255.11
18. Kalimantan Tengah	6	221.17	0	0.00	0	0.00	0	0.00	6	221.17
19. Kalimantan Selatan	10	153.10	0	0.00	1	1473.00	0	0.00	11	273.09
20. Kalimantan Timur	9	176.56	0	0.00	2	1256.00	0	0.00	11	372.82
21. Sulawesi Utara	6	164.83	1	636.00	1	1086.00	0	0.00	8	338.88
22. Sulawesi Tengah	7	60.71	1	539.00	0	0.00	0	0.00	8	120.50
23. Sulawesi Selatan	23	119.87	0	0.00	0	0.00	1	2695.00	24	227.17
24. Sulawesi Tenggara	5	150.80	0	0.00	0	0.00	0	0.00	5	150.80
25. Maluku	5	165.80	0	0.00	0	0.00	0	0.00	5	165.80
26. Maluku Utara	3	135.33	0	0.00	0	0.00	0	0.00	3	135.33
27. Irja Barat	4	115.25	0	0.00	0	0.00	0	0.00	4	115.25
28. Irja Tengah	4	64.75	0	0.00	0	0.00	0	0.00	4	64.75
29. Irja Timur	4	174.75	0	0.00	0	0.00	0	0.00	4	174.75
INDONESIA	**221**	**173.38**	**49**	**664.45**	**40**	**1395.50**	**26**	**3658.23**	**336**	**660.15**

Table B.5f
No. of Kab/Kota by grouping number of Census Blocks by provinces

Rural

Province	Grouping no. of Census Blocks								Total	
	<= 500		501 - 999		1000 - 1999		>= 2000			
	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg	Kab/Kota	Avg
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
01. Dista Aceh	3	211.67	3	700.67	5	1364.20	0	0.00	11	868.91
02. Sumatera Utara	2	32.50	6	767.00	5	1679.60	2	2767.50	15	1240.00
03. Sumatera Barat	6	112.33	4	896.75	4	1113.75	0	0.00	14	622.57
04. Riau	8	254.38	5	727.60	2	1295.50	0	0.00	15	550.93
05. Jambi	3	307.00	6	673.17	1	1076.00	0	0.00	10	603.60
06. Sumatera Selatan	3	148.33	0	0.00	3	1534.33	4	2833.25	10	1638.10
07. Bengkulu	1	14.00	1	957.00	2	1203.50	0	0.00	4	844.50
08. Lampung	2	35.50	1	807.00	3	1378.67	4	2541.00	10	1517.80
09. DKI Jakarta	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
10. Jawa Barat	5	57.80	0	0.00	1	1282.00	19	3079.37	25	2403.16
11. Jawa Tengah	3	98.00	2	725.00	12	1622.67	15	2621.40	32	1891.78
12. DI Yogyakarta	1	438.00	2	716.50	1	1773.00	0	0.00	4	911.00
13. Jawa Timur	3	56.67	1	761.00	12	1573.75	16	2918.00	32	2078.25
14. Bali	5	373.20	3	848.67	0	0.00	0	0.00	8	551.50
15. Nusa Tenggara Barat	1	395.00	0	0.00	5	1322.00	0	0.00	6	1167.50
16. Nusa Tenggara Timur	5	310.20	8	668.50	1	1238.00	0	0.00	14	581.21
17. Kalimantan Barat	0	0.00	3	645.33	5	1301.60	0	0.00	8	1055.50
18. Kalimantan Tengah	2	254.00	2	797.00	1	1878.00	1	2140.00	6	1020.00
19. Kalimantan Selatan	3	175.33	6	779.00	2	1151.50	0	0.00	11	682.09
20. Kalimantan Timur	7	248.00	2	755.50	2	1281.50	0	0.00	11	528.18
21. Sulawesi Utara	4	189.75	1	810.00	2	1207.00	1	2414.00	8	799.63
22. Sulawesi Tengah	4	269.50	2	745.00	2	1505.50	0	0.00	8	697.38
23. Sulawesi Selatan	10	281.80	12	711.58	2	1385.00	0	0.00	24	588.63
24. Sulawesi Tenggara	1	78.00	2	717.50	2	1546.00	0	0.00	5	921
25. Maluku	2	247.50	2	690.50	1	1472.00	0	0.00	5	669.60
26. Maluku Utara	1	89.00	1	509.00	0	0.00	1	2047.00	3	881.67
27. Irja Barat	2	212.50	1	960.00	1	1086.00	0	0.00	4	617.75
28. Irja Tengah	5	283.80	0	0.00	0	0.00	0	0.00	5	283.80
29. Irja Timur	2	191.50	1	681.00	2	1601.00	0	0.00	5	853.20
INDONESIA	**94**	**214.65**	**77**	**737.53**	**79**	**1443.62**	**63**	**2827.78**	**313**	**1179.43**

Table B.6a
Number of Census Blocks

Urban + Rural

Province	Census Blocks			Total
	Common	Preparation	Special	
(1)	(2)	(3)	(4)	(5)
01. Dista Aceh	8,591	2,888	54	11,533
02. Sumatera Utara	25,181	4,335	123	29,639
03. Sumatera Barat	10,787	965	39	11,791
04. Riau	10,888	2,428	59	13,375
05. Jambi	5,965	1,767	39	7,771
06. Sumatera Selatan	16,836	5,499	53	22,388
07. Bengkulu	3,228	1,076	12	4,316
08. Lampung	15,052	3,471	37	18,560
09. DKI Jakarta	19,038	248	76	19,362
10. Jawa Barat	101,846	13,719	262	115,827
11. Jawa Tengah	76,798	18,491	289	95,578
12. DI Yogyakarta	8,709	148	25	8,882
13. Jawa Timur	91,215	16,479	594	108,288
14. Bali	7,528	523	25	8,076
15. Nusa Tenggara Barat	9,162	1,124	61	10,347
16. Nusa Tenggara Timur	7,785	1,391	67	9,243
17. Kalimantan Barat	8,807	1,899	34	10,740
18. Kalimantan Tengah	4,645	2,795	7	7,447
19. Kalimantan Selatan	8,308	2,163	36	10,507
20. Kalimantan Timur	7,048	2,820	43	9,911
21. Sulawesi Utara	7,236	1,832	40	9,108
22. Sulawesi Tengah	4,699	1,820	24	6,543
23. Sulawesi Selatan	17,776	1,711	92	19,579
24. Sulawesi Tenggara	3,764	1,585	10	5,359
25. Maluku	2,781	1,374	22	4,177
26. Maluku Utara	1,893	1,143	15	3,051
27. Irja Barat	1,499	1,417	16	2,932
28. Irja Tengah	1,270	381	27	1,678
29. Irja Timur	3,161	1,790	14	4,965
INDONESIA	**491,496**	**97,282**	**2,195**	**590,973**

\

Table B.6b
Number of Census Blocks

Urban

Province	Census Blocks			Total
	Common	Preparation	Special	
(1)	(2)	(3)	(4)	(5)
01. Dista Aceh	1,828	117	30	1,975
02. Sumatera Utara	10,543	411	85	11,039
03. Sumatera Barat	2,989	59	27	3,075
04. Riau	4,842	217	52	5,111
05. Jambi	1,658	58	19	1,735
06. Sumatera Selatan	5,670	295	42	6,007
07. Bengkulu	892	40	6	938
08. Lampung	3,204	160	18	3,382
09. DKI Jakarta	19,038	248	76	19,362
10. Jawa Barat	51,734	3,791	223	55,748
11. Jawa Tengah	30,600	4,215	226	35,041
12. DI Yogyakarta	5,196	17	25	5,238
13. Jawa Timur	37,129	4,329	326	41,784
14. Bali	3,499	142	23	3,664
15. Nusa Tenggara Barat	3,121	172	49	3,342
16. Nusa Tenggara Timur	1,046	21	39	1,106
17. Kalimantan Barat	2,189	85	22	2,296
18. Kalimantan Tengah	1,196	124	7	1,327
19. Kalimantan Selatan	2,876	105	23	3,004
20. Kalimantan Timur	3,804	270	27	4,101
21. Sulawesi Utara	2,513	170	28	2,711
22. Sulawesi Tengah	913	41	10	964
23. Sulawesi Selatan	5,285	107	60	5,452
24. Sulawesi Tenggara	713	38	3	754
25. Maluku	704	111	14	829
26. Maluku Utara	355	47	4	406
27. Irja Barat	405	44	12	461
28. Irja Tengah	223	10	26	259
29. Irja Timur	656	29	14	699
INDONESIA	**204.821**	**15.473**	**1.516**	**221.810**

Table B.6c:
Number of Census Blocks

Rural

Province	Census Blocks			Total
	Common	Preparation	Special	
(1)	(2)	(3)	(4)	(5)
01. Dista Aceh	6,763	2,771	24	9,558
02. Sumatera Utara	14,638	3,924	38	18,600
03. Sumatera Barat	7,798	906	12	8,716
04. Riau	6,046	2,211	7	8,264
05. Jambi	4,307	1,709	20	6,036
06. Sumatera Selatan	11,166	5,204	11	16,381
07. Bengkulu	2,336	1,036	6	3,378
08. Lampung	11,848	3,311	19	15,178
09. DKI Jakarta	0	0	0	0
10. Jawa Barat	50,112	9,928	39	60,079
11. Jawa Tengah	46,198	14,276	63	60,537
12. DI Yogyakarta	3,513	131	0	3,644
13. Jawa Timur	54,086	12,150	268	66,504
14. Bali	4,029	381	2	4,412
15. Nusa Tenggara Barat	6,041	952	12	7,005
16. Nusa Tenggara Timur	6,739	1,370	28	8,137
17. Kalimantan Barat	6,618	1,814	12	8,444
18. Kalimantan Tengah	3,449	2,671	0	6,120
19. Kalimantan Selatan	5,432	2,058	13	7,503
20. Kalimantan Timur	3,244	2,550	16	5,810
21. Sulawesi Utara	4,723	1,662	12	6,397
22. Sulawesi Tengah	3,786	1,779	14	5,579
23. Sulawesi Selatan	12,491	1,604	32	14,127
24. Sulawesi Tenggara	3,051	1,547	7	4,605
25. Maluku	2,077	1,263	8	3,348
26. Maluku Utara	1,538	1,096	11	2,645
27. Irja Barat	1,094	1,373	4	2,471
28. Irja Tengah	1,047	371	1	1,419
29. Irja Timur	2,505	1,761	0	4,266
INDONESIA	**286,675**	**81,809**	**679**	**369,163**

SUSI ESTIMATION PROCEDURE

Report # 20

by
Vijay Verma

February, 2001

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 163

II. EXISTING SUSI SAMPLE DESIGN 163

III. ESTIMATION PROCEDURE USED SO FAR 163
- **A. Estimates of Aggregates** 163
  - **1. The 1996 Economic Census** 164
  - **2. SUSI 1996** 164
  - **3. SUSI 1998** 165
  - **4. SUSI 1999** 165
- **B. Variance Estimation** 165

IV. RATIO ESTIMATION 166

V. VARIANCE ESTIMATION 170

VI. CONSISTENCY OF RATIO ESTIMATES OF TOTALS 172

VII. NUMERICAL ILLUSTRATION OF THE ESTIMATION PROCEDURE 175

I. INTRODUCTION

The following procedure is described for the SUSI sample as hitherto selected. It is proposed that this be used to re-compute SUSI estimates and their standards errors.

Depending on the results of this investigation, it may be considered whether it would be necessary and desirable to republish some of the results. In any case, it is important to consider the following procedure for the publication of SUSI 1999.

II. EXISTING SUSI SAMPLE DESIGN

The SUSI sample has been selected as follows.

The sampling frame is provided by the 1996 Economic Census, which itself is based on a sample of areas (wilcah). The census covered all wilcah in urban areas and other areas of concentration of non-agricultural activity, while a 20% sample was taken in other areas. Overall, more than 50% of the wilcah were included.

Following the 1996 Economic Census, the first three applications of SUSI (1996, 1998, 1999) were sub-sampled from the census areas as follows.

- For SUSI 1996, a sample of areas (wilcah) was selected with probability proportional to a measure of size (PPS), systematically from lists ordered by geographical location, and all establishments within each selected wilcah were taken (compact cluster sampling).

- In SUSI 1998 a second sampling stage was introduced: within each sample area, establishments were listed and divided into three groups of ISIC (i=1,2,3), and a certain number selected with equal probability from each group. In both the above surveys, fieldwork was confined to a single period during the year.

- The design was similar for SUSI 1999, except that the sample wilcah were divided into 4 replications (representative sub-samples), and one replication was enumerated during each quarter of the year. Thus SUSI became a quarterly survey. However, considering the year as a whole, the sample structure of the remains unchanged.

III. ESTIMATION PROCEDURE USED SO FAR

A. Estimates of Aggregates

For the 1996 Economic Census as well as for each round of SUSI, an 'unbiased linear

estimator' of the following type has been used:

The contribution of a census/sample area k to the estimate of an aggregate y is given by

[1] $$Y_{k,i} = y_{k,i} * W_{k,i}$$

where

y an aggregate quantity such as the number of establishments, employment, or output in a particular ISIC or group of ISIC's (i)

$y_{k,i}$ the quantity measured in area (wilcah) k for ISIC(s) i

$W_{k,i}$ weighting or inflation factor for area k for ISIC(s) i, as defined below.

Summed over census/sample areas in any domain (e.g. province or the whole country, groups of ISIC's) gives an estimate of the total for the domain (d):

[2] $$Y_{(d)} = SY_k, \text{ where the sum is over } k,i \text{ e } d$$

The weighting or inflation factor $W_{k,i}$ is defined as follows.

1. The 1996 Economic Census

[3] $$W_{k,i} = 1 / (P_{1k} * R_{1k})$$

where

P_{1k} the probability of selection of area k into the census (=1.0 for urban and other areas of concentration; =0.2 for other areas)

$R_{1,k}$ response rate in area k (proportion of establishments successfully enumerated) in the census

Note that W is the same for all ISIC's, its value depending only on the wilcah.
[verify whether R was used, or assumed 1.0 throughout]

2. SUSI 1996

This was carried out over a sub-sample of census areas, with complete coverage within each area selected.

[4] $$W_{k,i} = 1 / (P_{1k} * P_{2k} * R_{2k})$$

where

P_{2k} the probability of selection of census area k into the survey

R_{2k} response rate in area k (proportion of establishments successfully enumerated in the survey)

[verify whether R was used, or assumed 1.0 throughout]

3. SUSI 1998

Up to three different sub-sampling rates ($P_{3k,i}$, i=1-3) were used within each sample area for the selection of establishment. (Subscript i represent a group of ISIC's)

[5] $$W_{k,i} = 1 / (P_{1k} * P_{2k} * P_{3k,i} * R_{3k,i})$$

where

$P_{3k,i}$ sub-sampling rate for establishment in ISIC group i, in sample area k

$R_{3k,i}$ response rate for ISIC group i in area k (proportion of establishments successfully enumerated in the survey)

[verify whether R was used, or assumed 1.0 throughout]

4. SUSI 1999

The same as above, except that for estimating from each of the 4 replications, we have

[6] $$W_{k,i} = 4 / (P_{1k} * P_{2k} * P_{3k,i} * R_{3k,i})$$

[verify whether R was used, or assumed 1.0 throughout]

B. Variance Estimation

It is understood that variance for SUSI has been estimated as follows:

$$\mathbf{var}(Y_{[d,i]}) = \sum_{k \in [d]} \left\{ w_k . (w_k - 1) . Y_{k,i}^2 \right\}$$

where k refers to a sample area in a geographic domain of estimation (d), such as a province or region. Subscript i indicates an individual ISIC or a set of ISIC's for which the estimation is being made. The weight w_k depends only on the area (wilcah k), independent of the ISIC's concerned since no account is taken of any differences by ISIC in the sub-sampling applied within the same area.

There are two limitations of this variance estimation procedure.

- It disregards the stratification present in the actual design, which results from systematic sampling of areas from geographically ordered lists.

- Also disregarded is the effect of subsampling within sample areas. This helps to control the sample size per cluster in the actual design, and hence to reduce sampling error.

It is expected that on both these counts, reported variances have been over-estimated. In other words, the SUSI design may be better than it appears from the reported sampling errors.

A more appropriate variance estimation procedure will be described below: it is nothing but the most commonly used for ratio estimate in surveys of diverse types.

IV. RATIO ESTIMATION

Notation

k subscript identifying a sample area (wilcah).

The following quantities are defined *for each area* (subscript k has been dropped for convenience; it will be added when necessary):

P probability of selection of the area (wilcah); this is the *actual* (as distinct from a relative) overall probability, taking also into account that the frame from the 1996 Economic Census is itself based on a sample.

$P=P_1*P_2$, the product of the area's probability of selection into the census (P_1), and the probability of selection from the census frame into the survey (P_2).

i subscript identifying groups of ISIC's (i=1,2,3), as distinguished at the second stage of sampling

j individual ISIC (at the most detailed level at which the survey results are reported)

j(i) identifying that an individual ISIC j belongs to the group i; used to indicate quantities which vary by group i, but are the same for all individual ISIC's j in the group

N_i' number of establishments of ISIC *group* i listed for SUSI in the area (i=1-3)

n_i' number of establishments of ISIC group i *enumerated* in SUSI in the area (this may be less than the number selected due to non-response)

W_i sample weight for ISIC group i in the area, $W_i = \frac{1}{P} \cdot \frac{N_i'}{n_i'}$.

Note: This quantity may also be identified as $W_{j(i)}$ when necessary for an individual ISIC j in group i, indicating that its value is the same for all j in the group.

Note. The procedure is described assuming the SUSI 1998 structure. This applies to the other surveys as well, with the weights W above appropriately defined.

SUSI 1966:

$$W_i = \frac{1}{P}$$

SUSI 1999, each replication: $$W_i = \frac{4}{P} \cdot \frac{N_i'}{n_i'}$$

n_j' number enumerated establishments of *individual* ISIC j in the area

e_j' total employment in enumerated establishments of individual ISIC j in the area

v_j' total value (output) of enumerated establishments of individual ISIC j in the area

n_j number of establishments of *individual* ISIC j in the area according to the 1996 Economic Census

e_j employment in ISIC j in the area according to the 1996 Economic Census

v_j value (production) of ISIC j in the area according to the 1996 Economic Census

We begin with the 'unbiased linear estimator' exactly as described above. This is then used to construct the more precise ration estimator. Using the notation introduced above, the estimation is as follows.

The contribution of the area to estimates of the number of establishments, employment and output in SUSI are, respectively, as follows. (Note. This is the same as [1] and [5] above, except for the new notation introduced here).

[7] $$N_j' = W_{j(i)} n_j'$$ from SUSI

$$E_j' = W_{j(i)} e_j'$$

$$V_j' = W_{j(i)} v_j'$$

where subscript j(i) indicates that the weights are determined by the group i to which ISIC j belongs. It is the same for all ISIC's in the group, and equals W_i defined above.

Similarly, the contribution of the same area to estimates of the number of establishments, employment and output from the 1996 Economic Census are, respectively:

[8] $$N_j = n_j / P$$ from the 1996 Economic Census, *using SUSI areas only*

$$E_j = e_j / P$$

$$V_j = v_j / P$$

The above quantities may be summed up over all sample areas (k) in a province to provides estimates (of the number of establishments, employment and output) by ISICxProvince. They may be further aggregated to other domains (d), such as by province to obtain national-level estimates by ISIC, or by ISIC to obtain total (all ISIC) provincial estimates, or both by ISIC and

province to obtain total national estimates.

From SUSI [7], we have estimates of such domain totals as (this is the same as [2] above):

[9]
$$N'_{(d)} = \sum_{k,j \in d} N'_{k,j} ,$$

$$E'_{(d)} = \sum_{k,j \in d} E'_{k,j} ,$$

$$V'_{(d)} = \sum_{k,j \in d} V'_{k,j}$$

where the sum is over all sample areas (k) and/or ISIC's (j) in the domain (d).
Similarly, the 1996 Economic Census gives the following estimates <u>from SUSI areas</u> in the domain:

[10]
$$N_{(d)} = \sum_{k,j \in d} N_{k,j} ,$$

$$E_{(d)} = \sum_{k,j \in d} E_{k,j} ,$$

$$V_{(d)} = \sum_{k,j \in d} V_{k,j}$$

Note: In either of the above sets, any areas selected but not enumerated in SUSI are excluded from the sum. However, any areas enumerated but found to be empty (no establishments) are to be included with zero value.

We also have the full the 1996 Economic Census estimates for the domain (i.e. from *all census areas* in the domain – not just those in SUSI sample, i.e. the census data weighted by inverse of the census selection probabilities P_1 defined above). Let these be:

[11]
$$N^{EC}_{(d)}, \quad E^{EC}_{(d)}, \quad V^{EC}_{(d)},$$

for the number of establishments, employment and output, respectively.

<u>SUSI ratio estimates</u>

Therefore, the corresponding *ratio estimates* from SUSI are:

[12]
$$N^{S}_{[d]} = \left(\frac{N^{t}_{[d]}}{N_{[d]}}\right).N^{EC}_{[d]} ,$$

$$E^{S}_{[d]} = \left(\frac{E^{t}_{[d]}}{E_{[d]}}\right).E^{EC}_{[d]} ,$$

$$V^{S}_{[d]} = \left(\frac{V^{t}_{[d]}}{V_{[d]}}\right).V^{EC}_{[d]} .$$

Estimates of proportionate change from the 1996 Economic Census are, obviously:

[13]
$$\frac{N^{S}_{[d]}}{N^{EC}_{[d]}} = \left(\frac{N^{t}_{[d]}}{N_{[d]}}\right) ,$$

$$\frac{E^{S}_{[d]}}{E^{EC}_{[d]}} = \left(\frac{E^{t}_{[d]}}{E_{[d]}}\right),$$

$$\frac{V^{S}_{[d]}}{V^{EC}_{[d]}} = \left(\frac{V^{t}_{[d]}}{V_{[d]}}\right).$$

Note. If a denominator in [12] is zero, i.e. all SUSI areas in the domain were empty at the time of the 1996 Economic Census, [7] is used to provide the required estimate, instead of [12]. No estimate of the proportionate change [13] is possible. By contrast, if a domain is empty in SUSI, the estimates are all zero, of course.

Note. It has been pointed out that reliable estimates of value from the 1996 Economic Census are not available. If so, we may use census employment (rather than output) as the auxiliary variable in estimating output in [4]:

[14]
$$V^{S}_{[d]} = \left(\frac{V^{t}_{[d]}}{E_{[d]}}\right).E^{EC}_{[d]} ,$$

while there is no point in computing proportionate change in V from [13].

V. VARIANCE ESTIMATION

We consider the combined ratio estimator of two aggregates y and x

[15] $$r = \frac{y}{x} = \frac{\sum_h y_h}{\sum_h x_h} = \frac{\sum_h \sum_k y_{hk}}{\sum_h \sum_k x_{hk}} = \frac{\sum_h \sum_k \sum_j w_{hkj} \cdot y_{hkj}}{\sum_h \sum_k \sum_j u_{hkj} \cdot x_{hkj}}$$

where

$w_{hkj}.y_{hkj}$ weighted aggregate of variable y for ISIC j for SUSI area (wilcah) k in stratum h, as estimated from SUSI;

$u_{hkj}.x_{hkj}$ the same, for SUSI area k, but as estimated from the 1996 Economic Census

S_j summed over any set of ISIC's (or a single ISIC) of interest

S_k summed over sample areas in stratum h. Note. The systematic sample of SUSI may be approximated as consisting of two independent selections per stratum. That is, areas in the order in which they were selected may be paired, each pair forming a stratum for variance computations, and

S_h summed over strata h in any estimation domain of interest, such as a province, region or whole country.

Variance of r is estimated as:

[16] $$var(r) = \frac{1}{x^2}.[var(y) + r^2.var(x) - 2r.cov(x,y)]$$

where

[17] $$var(y) = \sum_h \left[(1 - f_h).\frac{a_h}{a_h - 1}.\sum_k \left(y_{hk} - \frac{y_h}{a_h} \right)^2 \right].$$

$$var(x) = \sum_h \left[(1 - f_h).\frac{a_h}{a_h - 1}.\sum_k \left(x_{hk} - \frac{x_h}{a_h} \right)^2 \right].$$

$$cov(x,y) = \sum_h \left[(1 - f_h).\frac{a_h}{a_h - 1}.\sum_k \left(x_{hk} - \frac{x_h}{a_h} \right)\left(y_{hk} - \frac{y_h}{a_h} \right) \right].$$

A useful simplification is obtained by introducing the computational variable

[18]
$$z_{hk} = \frac{1}{x} \cdot (y_{hk} - r \cdot x_{hk}); \quad z_h = \sum_k z_{hk}$$

which reduces var(z) to the same form as var(y) of a simple total such as y or x:

[19]
$$var(r) = \sum_h \left[(1 - f_h) \cdot \frac{a_h}{a_h - 1} \cdot \sum_k \left(z_{hk} - \frac{z_h}{a_h} \right)^2 \right].$$

A program such as CLUSTERS or PC-CARP may be used for performing the above computations.

Note. This simple and flexible method is based on the following assumtions about the sample design:

The sample selection is independent between strata.

1 Two or more primary selections are drawn from each stratum.

2 These primary selections are drawn at random, independently and with replacement.

3 The number of primary selections is large enough for valid use of the ratio estimator and the linearisation approximation involved in the standard expression for its variance.

4 The quantities x_{hi} in the denominator are reasonably uniform in size within strata.

The last mentioned requirement is concerned with keeping the bias of the ratio estimator small. In practical terms the requirement is that relative variance, var(x)/x, should be ideally below 0.1, and anyway should not exceed 0.2 when ratio estimation is used.

Note. With the model of two independent selections per stratum (a_h=2), as can be assumed for SUSI, and with the finite population correction neglected, the expressions become particularly simple:

[20]
$$var(y) = \sum_h (y_{h1} - y_{h2})^2$$

$$var(x) = \sum_h (x_{h1} - x_{h2})^2$$

$$cov(x,y) = \sum_h (x_{h1} - x_{h2}) \cdot (y_{h1} - y_{h2})$$

or introducing the variable z_{hk} as above,

$$var(r) = \sum_h (z_{h1} - z_{h2})^2.$$

VI. CONSISTENCY OF RATIO ESTIMATES OF TOTALS

Consistency of estimates produced from the survey is obviously a desirable property. A special aspect of consistency is the *additivity* of estimates produced at different level of aggregation. Consistency requires that statistics for mutually exclusive domains should add up to the estimates given for their sum, for instance provincial estimates adding up exactly to the national level estimate provided, or separate estimates by ISIC adding up to the estimate for all sectors combined.

In this sense, ratio estimates, though more precise than simple unbiased estimates (i.e. direct or 'linear' estimates of totals), may not be additive. A slight modification of the weights at the estimation stage may be introduced to ensure such consistency, as explained below. Of course, often such refinement is of no substantive significance, but it is a desirable feature in publishing official results.

Suppose ratio estimates of totals of some variable are reported for a matrix of domains (i,j). For example, i may refer to ISIC and j to province. Thus we are required to produce estimates of totals of some variable Y by ISIC and province (Y_{ij}); also by ISIC for the whole country ($Y_{i.}$); by province for all ISIC's combined ($Y_{.j}$); and the overall total for all ISIC's combined for the whole country ($Y_{..}$).

With linear estimates, it is automatically ensured that the Y_{ij} summed over provinces equal $Y_{i.}$, and summed over ISIC's equal $Y_{.j}$. Similarly, the $Y_{i.}$ summed over ISIC's and the $Y_{.j}$ summed over provinces both equal the overall total $Y_{..}$.

Non-linear estimates such as ratio estimates may not be exactly consistent in the above sense.

Consider a ratio estimate for a particular ISIC (i) and province (j)

[21] $$Y''_{ij} = \frac{y_{ij}}{x_{ij}} . X_{ij} .$$

Its sum over provinces for a given ISIC is

[22] $$Y''_{i.} = \sum_{.j} \left(\frac{y_{ij}}{x_{ij}} . X_{ij} \right);$$

the sum over ISIC's for a given province is

[23] $$Y''_{.j} = \sum_{i} \left(\frac{y_{ij}}{x_{ij}} . X_{ij} \right);$$

and the total over all ISIC's and provinces is

[24] $$Y''_{..} = \sum_i \sum_j \left(\frac{y_{ij}}{x_{ij}} . X_{ij} \right).$$

Normally, in place of [22] and [23], totals by ISIC and province are estimated using 'combined' ratios:

[25] $$Y'_{i.} = \left(\frac{\Sigma_j y_{ij}}{\Sigma_j x_{ij}} \right) . X_{i.} \text{ and}$$

[26] $$Y'_{.j} = \left(\frac{\Sigma_i y_{ij}}{\Sigma_i x_{ij}} \right) . X_{.j},$$

which are not exactly the same as [22] and [23].

Similarly, [25] and [26] summed to the whole country

[27] $$Y'_{(I)} = \Sigma_i Y'_{i.} = \sum_i \left(\frac{\Sigma_j y_{ij}}{\Sigma_j x_{ij}} . X_{i.} \right)$$

[28] $$Y'_{(J)} = \Sigma_j Y'_{.j} = \sum_j \left(\frac{\Sigma_i y_{ij}}{\Sigma_i x_{ij}} . X_{.j} \right)$$

may differ from each other, from [24], and also from the proper 'combined' overall ratio estimate for the country:

[29] $$Y = \left(\frac{\Sigma_i \Sigma_j y_{ij}}{\Sigma_i \Sigma_j x_{ij}} \right) . X_{..}.$$

In short, estimates for individual ISIC's may not add up exactly to the 'combined ratio' estimate for all ISIC's combined, or those for individual provinces may not add up exactly to the estimate for the whole country. The discrepancies, even if normally small or even trivial, can be irritating and confusing to the non-statistical user.

How to make these estimates exactly consistent?

1. First make the marginal estimates [25] and [26] to exactly add up to [29]. This is achieved by replacing estimator [25] by

[25.a]
$$Y_{i.} = Y'_{i.} \cdot \left(\frac{Y}{\Sigma_i Y'_{i.}} \right) = Y'_{i.} \cdot \left(\frac{Y}{Y'_{[I]}} \right),$$

which summed over i give the correct total Y.

Similarly, replace [26] by

[26a]
$$Y_{j} = Y'_{.j} \cdot \left(\frac{Y}{\Sigma_i Y'_{j}} \right) = Y'_{.j} \cdot \left(\frac{Y}{Y'_{[J]}} \right),$$

which summed over j also give the correct total Y.

2. Next, for each row in turn, the cell estimates [21] can be re-scaled to add-up to the corrected row total [25.a]:

[21.a]
$$Y_{ij}^{(1)} = Y''_{ij} \cdot \left(\frac{Y_{i.}}{\Sigma_j Y''_{ij}} \right) = Y''_{ij} \cdot \left(\frac{Y_{i.}}{Y''_{i.}} \right).$$

Using the results of [21.a], these cell estimates can be re-scaled, for each column in turn, to add-up to the corrected column total [25.b]:

[21.b]
$$Y_{ij}^{(2)} = Y_{ij}^{(1)} \cdot \left(\frac{Y_{j}}{\Sigma_i Y_{ij}^{(1)}} \right).$$

3. The above is the first cycle of an *iterative procedure*. In general for cycle t = 2, 3, … :

[27.a]
$$Y_{ij}^{[2t+1)} = Y_{ij}^{[2t)} \cdot \left(\frac{Y_{i.}}{\Sigma_j Y_{ij}^{[2t)}} \right).$$

[27.b]
$$Y_{ij}^{[2t+2)} = Y_{ij}^{[2t+1)} \cdot \left(\frac{Y_{j}}{\Sigma_i Y_{ij}^{[2t+1)}} \right).$$

The iterative procedures may be continued till all terms in parentheses in [27.a, 27.b] become close to 1.0 within a small margin, such as 0.05%. Normally, only a very small number of iterations are needed.

If a single main variable Y and a particular classification of the reporting domains only is of interest, the above adjustments can be incorporated into the sample weights at the micro-level, so that strict additivity is automatically ensured when the data are tabulated. However, generally it is more convenient to apply the above manipulations to the tabulated results.

VII. NUMERICAL ILLUSTRATION OF THE ESTIMATION PROCEDURE

The following detailed results are meant only to provide illustrations of the estimation procedure and data-related issues which may be important. These should not be compared with any officially published figures. The use of wrong data and/or computational mistakes in the following illustrations, prepared within a very short span of time, cannot be ruled out.

TABLE 1. SUSI 1999 Sample number of wilcah/establishments/employment - by province, ISIC and subsample

Unweighted frequencies – numbers of wilcah, establishments and workers – in various domains of the sample provided the natural starting point for investigation of the sample implementation and outcome. The important result shown in Table 1 is that - at least in the data file available for the present analysis – in a significant proportion (7% overall in the three SUSI99 subsamples) of the selected wilcah not a single establishment was enumerated in the survey. The proportion is particularly high (13%) in subsample 3, with the lowest figures in provinces 14 (30%), 61, 12, 82 and 81 (100% missing in the last case). While some of these wilcah may be genuinely empty at the time of the survey, these figures strongly suggest a failure to enumerate them (non-response). In the case of non-response, the weights of the enumerated areas must be appropriately adjusted; otherwise, the results will be under-estimated.

The appropriate adjustment for non-response is to increase the inflation factors $F_2=1/P_2$ by the ratio of

- Estimated 1996 Economic Census results from the areas selected in the survey, to
- Estimated 1996 Economic Census results from the areas enumerated in the survey (including areas found empty)

These estimates are of the type shown in Table 2; the estimation procedure has been described earlier.

TABLE 2. 1996 Economic Census estimated number of establishments/employment, by province and ISIC (1 digit)

Estimates of numbers of establishments and employment have been constructed from the full 1996 Economic Census by Mr Kadarmanto of BPS. The data from each census wilcah are inflated by $1/P_1$, the inverse of the selection probability of the wilcah into the census. These are

the factors by which ratios from the survey are multiplied to estimate aggregates (totals).

TABLE 3 1996 Economic Census estimates - from SUSI99 areas

The census estimates from SUSI areas are constructed by weighting the data from each wilcah in SUSI by $F_1*F_2 = 1/(P_1*P_2)$, where P_2 is the probability of selection of the census area into the SUSI sample. Estimate have been constructed separately by SUSI subsample. Two points should be noted. One, the factors F_2 must be adjusted for non-response at the wilcah level as noted in relation to Table 1. (Apparently this has not been done in the data used here – hence these are under-estimates). Secondly, it turns out that the factors as coded in the present data need to be multiplied by 4/3 to get the results shown (similarly the data for each subsample have been multiplied by 4). This may be because the survey was originally designed to have 4 subsamples., and the factors computed with that assumption have not been adjusted to the fact that only 3 subsamples wee actually enumerated. *But please see also notes to Table 8 below.*

TABLE 4. Estimation for SUSI 1999 (based on 3 quarters): summary by ISIC and by province

TABLE 5. Estimation for SUSI 1999 (based on 3 quarters): cross-tabulation by ISIC and province

These tables show the actual results: linear or 'unbiased' estimates, and the improved ratio estimates by varying degrees of detail. Please note 'unofficial' nature of the data files used for this illustration. The BPS should repeat this exercise using the final data.

The linear or 'unbiased' estimates by inflating SUSI data by $F_1*F_2*F_{3.i} = 1/(P_1*P_2*P_{3,i})$, where $P_{3,i}$ is the sampling rate applied to ISIC group i for the selection of establishments within the wilcah. As above, this factor must be adjusted for non-response (which apparently has not been done in the present data). The adjustment requires multiplying the factor by inverse of the response rate within the wilcah. Or, the factor equals the ratio of the number of establishments *listed*, to the number actually *enumerated*.

TABLE 6. SUSI99 Estimates by subsample

TABLE 7. The direction of change from one quarter to the next - SUSI99 subsamples

SUSI99 has been designed to provide quarterly estimates. Table 7 in particular shows the effect of ratio estimation on the indicated direction of change in the level from one quarter to the

next. Among the 26 provinces, for instance, linear or 'unbiased' estimate indicates a consistent decline from subsample 1 to subsample 2, and from subsample 2 to subsample 3, while in a majority (18) provinces the two changes are in opposite directions. The (more precise) ratio estimates indicate, by contrast, a consistent decline in 8 cases but a consistent increase in another 6.

TABLE 8. Ratio of census estimates from 1998 SUSI areas, to estimates from the 1996 Economic Census

Estimates similar to Table 3 above have been made for SUSI98.

A puzzling feature of the results is that the estimates fall short of the expected census values by 25% (a figure almost identical to that for SUSI99 in Table 3, before the results were multiplied by 4/3). This large difference is unlikely to be the result of general data errors, since the pattern is practically identical across the provinces (last two columns of Table 8). Most likely, there is some common discrepancy in the inflation factors in the data file. Could this be true of SUSI99 as well (but hidden here by the introduction of the factor 4/3)? It is important that this is investigated further by BPS.

ANNEX TABLES

A.09 SUSI99 Estimates by subsample and ISIC (details on Table 6)

A.10 The direction of quarterly change – by ISIC (details on Table 7)

Very detailed tables such as these are useful in identifying problems and patterns across the diverse reporting domains (ISIC, by province, by quarter etc.)

Table 1. Unweighted sample numbers - SUSI 1999 (3 quarters)

A. Wilcah

Number of wilcah in SUSI 1999 sample					Number of non-empty wilcah in SUSI99					proportion non-empty			
	Subsample					Subsample				Subsample			
Prov	1	2	3	1-3	Prov	1	2	3	1-3	1	2	3	1-3
11	56	56	56	168	11	49	47	46	142	0,88	0,84	0,82	0,85
12	262	262	261	785	12	257	256	123	636	0,98	0,98	0,47	0,81
13	53	52	53	158	13	53	52	53	158	1,00	1,00	1,00	1,00
14	42	43	43	128	14	39	38	30	105	0,93	0,84	0,70	0,82
15	41	40	40	121	15	41	39	39	119	1,00	0,98	0,98	0,98
16	57	58	58	173	16	53	56	55	166	0,93	1,00	0,95	0,96
17	39	39	39	117	17	39	39	38	116	1,00	1,00	0,97	0,99
18	56	56	56	166	18	54	56	54	163	0,96	1,00	0,98	0,98
31	154	155	156	464	31	150	154	149	453	0,97	0,99	0,96	0,98
32	186	185	185	556	32	182	185	184	551	0,98	1,00	0,99	0,99
33	283	284	284	851	33	282	281	282	845	1,00	0,99	0,99	0,99
34	67	67	67	201	34	67	67	67	201	1,00	1,00	1,00	1,00
35	288	288	288	864	35	287	288	282	857	1,00	1,00	0,98	0,99
51	49	49	49	147	51	47	49	46	142	0,96	1,00	0,94	0,97
52	47	47	47	141	52	47	45	46	138	1,00	0,96	0,98	0,98
53	42	41	41	124	53	42	41	41	124	1,00	1,00	1,00	1,00
61	45	45	45	135	61	44	43	30	117	0,98	0,96	0,67	0,87
62	41	41	40	122	62	41	38	38	117	1,00	0,93	0,95	0,96
63	44	43	44	131	63	44	35	44	123	1,00	0,81	1,00	0,94
64	39	40	40	119	64	38	39	38	115	0,97	0,98	0,95	0,97
71	39	39	39	117	71	38	35	33	106	0,97	0,90	0,85	0,91
72	39	39	38	116	72	39	38	38	115	1,00	0,97	1,00	0,99
73	57	57	58	172	73	56	57	57	170	0,98	1,00	0,98	0,99
74	36	36	36	108	74	36	33	36	105	1,00	0,92	1,00	0,97
81	42	41	41	124	81	13			13	0,31	0,00	0,00	0,10
82	39	39	39	117	82	34	27	13	74	0,87	0,69	0,33	0,63
Total	2.143	2.141	2.141	6.425	Total	2.072	2.037	1.862	5.972	0,97	0,95	0,87	0,93

not in census 1

Table 1, Unweighted sample numbers - SUSI 1999 (3 quarters)
B, Number of establishments
SUSI 1999 Sample number of establishments - by province, ISIC and subsample

	ISIC																			
	Sum of Utot subsample				Sum of U2 (units in ISIC 2)				Sum of U3				Sum of U4				Sum of U5			
Prov	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	642	629	571	1.842	2	1	1	4	261	229	192	682					14	8	18	40
12	2.849	2.922	1.265	7.056	10	7	2	19	470	481	207	1.158	1	6	1	8	18	31	7	56
13	693	603	678	1.974	2	2	7	11	194	199	230	623					27	15	17	59
14	472	476	453	1.401	3	3	1	7	72	110	50	232	1	3	12	16	11	5	3	19
15	525	477	524	1.526	4	3	24	31	54	85	65	204	1	1	2	4	21	10	17	48
16	756	736	684	2.176	2	3	3	8	184	192	159	535	2	2	3	7	16	19	9	44
17	371	453	397	1.221	1	3	1	5	80	70	58	208					6	13	8	27
18	790	747	700	2.237	2	2	1	5	299	290	227	816	1	1	1	3	12	11	4	27
31	2.256	2.204	2.298	6.758					160	184	213	557					4	3	9	16
32	2.667	2.554	2.603	7.824	37	8	10	55	654	651	723	2.028	5	1	1	7	9	10	4	23
33	4.776	4.644	4.540	13.960	178	107	22	307	2.035	2.010	2.020	6.065	1	1	1	3	16	2	6	24
34	1.061	1.108	1.056	3.225	38	32	18	88	406	396	440	1.242					2	4	3	9
35	4.344	4.311	4.309	12.964	87	66	73	226	1.300	1.489	1.358	4.147	2	1	1	4	10	6	4	20
51	904	876	760	2.540	11	2	2	15	444	401	375	1.220					2	2	2	6
52	721	750	863	2.334	17	33	34	84	268	387	382	997					21	13	11	45
53	609	634	660	1.903	14	12	2	28	343	357	358	1.058	3	1	1	5	6	11	22	39
61	484	396	337	1.217	17	6	12	35	124	122	97	343	2	2	1	5	7	5	8	20
62	561	716	543	1.820	38	63	48	149	133	137	88	358					20	28	17	65
63	720	544	714	1.978	37	13	15	65	236	194	222	711	1	1	1	3	23	25	36	84
64	551	481	456	1.488	31	20	2	53	71	58	60	189	1	1	1	3	11	6	11	28
71	621	494	566	1.681	9	4	2	15	238	194	262	694					10	15	14	39
72	480	536	508	1.524	1	4	1	6	127	121	151	399	2	2	1	5	38	30	20	88
73	782	790	884	2.446	2	12	6	20	234	244	327	805	1	3	2	6	22	10	12	44
74	419	373	421	1.213	15	7	3	25	112	119	140	371	1	1	10	12	12	9	9	30
81	106			106					32			32					7			7
82	357	265	119	741	6	14	1	21	67	33	31	131					14	8	2	24
Total	29.527	28.719	26.909	85.156	564	427	291	1.282	8.657	8.733	8.413	25.803	25	27	39	91	359	299	273	931

Table 1. Unweighted sample numbers - SUSI 1999 (3 quarters)

C. Number of establishments (cont)

SUSI 1999 Sample number of establishments - by province, ISIC and subsample (cont.)

	ISIC															
	Sum of U6				Sum of U7				Sum of U8				Sum of U9			
	subsample															
Prov	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	284	242	237	743	43	86	47	176	3	5	5	13	55	69	71	194
12	1,322	1,279	587	3,188	429	516	249	1,194	93	69	21	183	506	533	211	1,250
13	274	242	243	759	70	37	69	176	24	29	45	98	102	79	67	248
14	200	211	161	572	88	46	68	202	26	12	64	102	71	86	94	251
15	228	227	210	665	100	69	83	252	15	24	14	53	102	58	109	269
16	312	302	278	892	168	125	121	414	25	28	26	79	47	65	85	197
17	167	205	175	547	49	58	60	166	25	34	50	109	44	70	47	161
18	289	293	304	886	110	73	65	248	9	11	16	36	68	66	82	216
31	1,057	936	997	2,990	249	280	271	800	405	424	383	1,212	381	377	425	1,183
32	992	976	927	2,895	673	627	646	1,946	90	64	103	257	207	217	189	613
33	1,583	1,595	1,500	4,678	485	454	503	1,442	72	84	60	216	406	391	428	1,225
34	361	389	348	1,098	80	76	71	229	71	80	75	226	103	129	101	333
35	1,555	1,555	1,566	4,676	681	575	668	1,924	157	132	129	418	542	487	516	1,547
51	289	296	250	835	35	52	38	125	43	39	28	110	80	84	65	229
52	247	237	265	749	111	88	143	322	14	6	10	32	43	24	38	105
53	182	173	197	552	22	25	16	63	12	22	19	53	27	33	45	105
61	210	186	142	538	38	30	35	103	11	4	5	20	75	41	37	153
62	209	276	202	691	118	110	102	330	6	17	14	37	43	85	72	200
63	230	176	247	655	87	76	109	272	3	5	6	14	44	52	78	174
64	231	204	189	624	65	78	92	235	46	51	43	140	95	63	58	216
71	241	173	181	595	68	58	67	191	5	7	7	19	52	43	33	128
72	196	239	205	640	43	75	59	177	19	13	12	44	54	52	59	165
73	319	328	308	965	114	122	148	384	6	7	9	24	92	64	52	208
74	169	164	173	506	56	38	51	145	16	9	4	29	38	26	31	95
81	47			47	5			5	3			3	12			12
82	166	127	62	355	37	36	9	82	21	6	4	33	46	39	10	95
Total	11,344	11,038	9,946	32,324	4,021	3,792	3,790	11,603	1,222	1,186	1,152	3,560	3,395	3,222	3,005	9,562

Number of Establishments by Province and 2 Digit ISIC Based on Integrated Survey Frame (SUSI Frame) - INFLATED

TABLE 2

A. 1996 Economic Census estimated number of establishments, by province and ISIC (1 digit)

Prov	c2	c3	c4	c5	c6	c7	c8	c9	Total
11	2,690	58,160	190	7,815	164,136	22,772	6,443	23,987	335,183
12	1,789	56,373	624	6,451	428,004	88,158	22,069	67,450	661,190
13	2,923	60,411	229	3,975	207,758	22,221	18,985	32,894	352,596
14	1,240	20,301	1,099	2,720	141,815	22,767	14,608	20,611	225,244
15	4,089	17,880	305	3,601	88,075	15,409	5,541	11,317	146,627
16	4,751	64,905	1,248	5,000	270,920	49,561	23,029	35,908	450,896
17	874	9,658	75	2,075	49,649	6,873	5,215	5,704	80,871
18	2,011	60,492	198	8,623	275,361	43,555	13,892	37,448	440,411
31	49	33,218	50	5,701	500,947	72,877	101,621	79,146	794,608
32	14,702	304,572	884	17,325	1,999,592	507,137	143,094	215,302	3,254,228
33	48,267	683,787	786	14,080	1,559,580	248,377	70,696	201,455	2,827,987
34	8,032	81,453	23	2,639	176,420	20,037	25,898	28,782	345,520
35	47,371	577,153	1,487	30,167	1,050,275	377,593	134,164	321,607	3,389,788
51	3,947	38,088	19	2,031	168,649	18,880	18,852	18,995	300,612
52	7,943	73,374	6	6,618	166,324	33,781	5,712	14,389	309,067
53	2,113	77,884	116	2,502	74,605	3,518	5,264	3,881	171,883
54	3,522	39,243	460	3,585	126,262	16,283	4,786	17,091	212,642
62	10,090	16,988	72	7,381	79,711	15,732	3,625	7,874	140,787
63	3,514	62,094	129	10,365	165,520	37,347	7,792	25,724	317,474
64	1,449	12,418	190	2,860	92,514	18,050	14,395	11,337	151,359
71	4,540	54,382	82	9,065	134,094	23,768	3,768	15,893	261,438
72	1,466	29,420	214	4,412	91,123	6,131	3,067	8,637	147,052
73	1,770	108,293	457	5,878	354,672	68,177	14,292	28,356	585,085
74	1,601	20,079	276	2,287	60,533	5,737	1,469	3,951	95,932
81	2,609	19,490	72	2,520	75,384	5,541	4,140	6,057	115,273
82	1,153	8,552	48	1,206	38,695	4,100	2,489	3,444	60,946
Total	109,780	2,729,222	9,568	176,351	8,413,529	1,719,217	679,025	1,292,690	16,160,583

Number of Establishments by Province and 2 Digit ISIC Based on Integrated Survey Frame (SUSI Frame) - INFLATED

TABLE 2 (cont)
B. 1996 Economic Census estimated number of workers, by province and ISIC (1 digit)

Prov.	c2	c3	c4	c5	c6	c7	c8	c9	Total
11	4.797	146.716	280	19.891	288.188	31.872	7.938	35.954	538.384
12	3.968	137.288	1.280	18.279	692.526	01.029	31.309	114.314	1.078.247
13	4.049	121.275	321	10.509	337.643	32.764	25.458	95.094	539.905
14	2.344	51.476	1.604	5.835	247.748	26.400	17.860	37.510	392.477
15	5.198	42.545	480	7.272	145.838	20.065	7.282	18.094	246.755
16	6.467	124.128	2,177	11.341	410.079	68.058	30.801	57.318	730.201
17	1.243	22.663	120	6.171	77.858	9.391	5.573	9.761	132.698
18	4.844	157.554	317	15.246	414.482	57.480	20.060	53.780	724.050
31	87	165.040	105	15.151	853.221	64.334	129.739	170.020	1.423.767
32	25.217	852.336	1.657	41.043	2.851.790	580.780	158.830	351.793	4.921.672
33	64.958	1.540.392	1.805	36.577	2.219.942	305.551	126.908	308.821	4.606.847
34	10.440	164.506	32	10.480	270.798	28.880	38.099	46.537	577.144
35	69.100	1.149.622	2.410	61.063	2.634.502	460.340	208.348	478.131	5.408.904
51	6.527	173.003	30	9.940	240.772	20.728	22.065	37.910	520.042
52	12.485	146.889	14	17.635	240.724	40.072	9.684	23.784	506.047
53	4.350	126.825	222	8.581	125.355	12.919	6.205	12.538	298.475
61	18.437	77.714	577	7.983	223.505	21.310	8.738	28.884	382.168
62	22.617	38.895	108	12.446	136.569	20.869	3.831	13.591	247.105
63	10.658	118.370	180	24.442	247.258	42.787	10.405	36.275	492.981
64	2.245	27.858	336	7.108	164.707	19.568	18.051	20.782	260.705
71	6.268	128.244	101	20.216	215.787	30.299	13.264	24.643	438.817
72	2.748	53.769	358	11.530	138.054	12.843	5.502	14.370	244.765
73	2.740	192.928	745	18.806	624.727	79.518	22.571	49.446	988.274
74	1.925	41.020	429	5.839	95.834	10.912	2.250	6.538	165.672
81	3.268	36.081	148	5.551	137.531	9.904	6.026	10.181	208.697
82	2.151	21.454	104	3.463	71.844	7.698	3.841	8.707	117.062
Total	286.378	5.092.169	16.863	418.755	14.003.282	2.111.960	963.029	2.027.405	26.809.032

TABLE 3
1996 Economic Census estimates - from SUSI99 areas
A. Number of establishments ISIC all(2-9)

	full EC96 estimate	Estimates of 1996 Economic Census variables, from SUSI 1999 census areas incl. found empty in SUSI)			
		Subsample			
Prov	Total	1	2	3	1-3
11	335.183	333.668	328.920	284.022	315.543
12	651.136	638.395	668.805	675.618	680.920
13	352.896	355.652	328.489	355.224	345.768
14	225.244	212.490	262.455	241.372	238.772
15	146.627	178.808	138.673	158.590	158.690
16	480.646	400.317	467.116	394.396	449.943
17	80.871	72.684	90.374	65.002	82.667
18	449.411	439.542	390.178	404.946	411.556
31	794.609	744.312	734.085	777.945	752.114
32	3.284.228	3.425.023	3.302.198	3.007.885	3.245.035
33	2.827.867	2.755.490	2.858.099	2.756.276	2.790.222
34	346.528	320.176	369.080	330.956	340.071
35	3.389.796	3.239.995	3.118.819	3.293.835	3.217.550
51	300.612	283.740	319.085	268.689	290.571
52	309.057	301.952	316.047	316.852	311.617
53	171.893	169.099	171.912	176.377	173.129
61	212.642	243.879	203.649	198.898	215.475
62	140.767	140.104	172.058	129.745	147.303
63	317.474	290.921	323.399	319.478	311.283
64	151.359	162.653	160.036	145.117	155.936
71	251.439	245.160	220.254	242.314	235.909
72	147.052	146.183	157.217	141.099	148.166
73	586.585	578.687	604.106	556.428	580.407
74	95.382	103.444	81.463	86.651	91.186
81	116.273	108.588	118.067	122.198	116.284
82	60.986	60.283	55.688	72.825	62.932
Total	16.166.593	16.039.206	15.969.062	15.548.940	15.949.069

February 13, 2001 SUSI Estimation Procedure

TABLE 3

1996 Economic Census estimates - from SUSI99 areas (cont)

B. Number of workers ISIC all(2-9)

	full EC96 estimate	Estimates of 1996 Economic Census variables, from SUSI 1999 census areas incl. found empty in SUSI)			
		Subsample			
Prov	Total	1	2	3	1-3
11	538.384	555.134	516.330	494.177	521.880
12	1.078.247	1.059.000	1.080.495	1.105.821	1.081.772
13	599.905	601.792	540.782	589.896	577.490
14	392.477	375.358	431.114	413.311	406.584
15	246.755	278.000	247.904	260.910	264.938
16	730.791	771.748	742.200	682.296	732.082
17	132.698	120.451	149.434	135.184	135.023
18	724.080	719.669	676.313	695.188	696.724
31	1.423.757	1.384.337	1.348.424	1.398.718	1.378.493
32	4.921.672	5.270.574	4.793.002	4.529.623	4.864.399
33	4.606.647	4.601.479	4.594.951	4.492.114	4.562.848
34	577.144	566.077	596.009	556.297	572.794
35	5.106.904	5.047.051	4.856.230	5.063.896	4.989.059
51	520.042	500.818	529.528	483.554	504.633
52	505.047	464.629	520.172	498.782	494.528
53	299.475	308.537	299.546	303.621	303.234
61	382.498	427.932	368.417	358.671	385.007
62	247.105	232.293	295.442	246.506	258.080
63	492.391	481.292	487.731	500.445	483.156
64	250.708	293.508	252.840	229.940	258.782
71	438.817	414.815	418.592	430.900	420.803
72	244.793	230.024	265.170	225.021	240.072
73	889.274	890.643	928.136	889.710	895.497
74	163.572	165.110	157.055	156.375	159.713
81	208.697	183.248	196.825	238.488	205.188
82	117.062	111.277	107.708	129.648	116.211
Total	25.829.032	26.031.893	25.394.952	25.097.091	25.507.979

TABLE 4

Estimation for SUSI 1999 (based on 3 quarters): summary by ISIC and by province

	Number of establishments				Number of workers			
		ratio to census				ratio to census		
ISIC	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation
all(2-9)	16.166.583	0,980	0,966	0,985	25.829.032	0,968	1,114	1,128
2	189.780	0,841	0,730	0,868	296.579	0,906	0,985	1,088
3	2.729.222	0,984	0,984	1,022	5.892.159	0,996	1,119	1,132
4	9.568	1,002	0,181	0,181	15.863	1,009	0,109	0,109
5	176.551	0,942	0,652	0,586	418.755	0,961	0,731	0,780
6	9.413.529	0,987	0,990	1,004	14.103.282	0,988	1,179	1,193
7	1.718.217	1,006	1,049	1,043	2.111.960	1,011	1,054	1,042
8	679.026	0,965	0,628	0,651	983.029	0,986	0,670	0,680
9	1.250.690	0,970	0,915	0,943	2.027.405	0,976	1,022	1,047
Province								
11	335.183	0,941	0,975	1,036	536.384	0,969	1,104	1,139
12	661.136	1,015	0,840	0,827	1.076.247	1,008	0,965	0,962
13	362.896	0,980	0,958	0,977	589.905	0,979	1,157	1,182
14	225.244	1,060	0,860	0,811	392.477	1,036	1,048	1,012
15	146.627	1,082	1,102	1,018	246.755	1,074	1,287	1,199
16	450.646	0,998	0,984	0,985	730.701	1,002	1,256	1,254
17	80.871	1,022	1,070	1,047	132.698	1,018	1,190	1,169
18	449.411	0,916	0,980	0,981	724.050	0,962	1,044	1,085
31	794.609	0,947	0,936	0,989	1.423.757	0,987	1,022	1,057
32	3.254.228	0,997	0,983	0,986	4.921.872	0,988	1,113	1,127
33	2.827.867	0,997	1,015	1,029	4.606.647	0,990	1,136	1,145
34	346.526	0,981	1,018	1,037	577.144	0,992	1,078	1,086
35	3.399.796	0,949	0,953	1,004	5.106.904	0,977	1,148	1,175
51	300.612	0,967	0,969	1,002	520.042	0,970	1,098	1,131
52	309.057	1,008	1,009	1,001	505.047	0,979	1,056	1,079
53	171.693	1,007	1,062	1,055	299.475	1,013	1,260	1,245
61	212.642	1,013	1,011	0,997	382.498	1,007	1,216	1,207
62	140.787	1,046	1,070	1,022	247.105	1,044	1,205	1,154
63	317.474	0,980	0,908	0,926	492.391	0,981	1,089	1,110
64	151.359	1,030	1,059	1,028	250.708	1,032	1,343	1,301
71	251.439	0,938	0,922	0,983	438.817	0,959	0,972	1,014
72	147.052	1,008	1,055	1,047	244.793	0,981	1,173	1,197
73	596.585	0,989	1,011	1,022	889.274	1,007	1,251	1,242
74	95.382	0,956	0,964	1,008	163.572	0,976	1,107	1,134
81	116.273	1,000	0,090	0,090	208.697	0,988	0,099	0,100
82	60.908	1,032	0,731	0,709	117.062	0,993	0,819	0,825
Total	16.166.583	0,980	0,966	0,985	25.829.032	0,968	1,114	1,128

TABLE 4

Estimation for SUSI 1999 (based on 3 quarters): summary by ISIC and by province

	Number of establishments				Number of workers			
		ratio to census				ratio to census		
ISIC	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation
all(2-9)	16.166.563	0,980	0,966	0,985	25.829.032	0,968	1,114	1,128
2	169.780	0,841	0,730	0,868	296.579	0,905	0,885	1,089
3	2.729.222	0,984	0,984	1,022	5.892.159	0,986	1,119	1,132
4	9.568	1,002	0,181	0,191	15.863	1,008	0,109	0,109
5	176.551	0,942	0,652	0,586	418.755	0,961	0,731	0,780
6	9.413.529	0,987	0,990	1,004	14.103.282	0,986	1,179	1,193
7	1.718.217	1,006	1,049	1,043	2.111.960	1,011	1,054	1,042
8	679.026	0,965	0,628	0,651	963.029	0,986	0,670	0,680
9	1.250.690	0,970	0,915	0,943	2.027.405	0,976	1,022	1,047
Province								
11	335.183	0,941	0,975	1,036	538.384	0,969	1,104	1,139
12	661.136	1,015	0,840	0,827	1.078.247	1,005	0,965	0,962
13	362.896	0,980	0,958	0,977	589.905	0,979	1,157	1,182
14	225.244	1,060	0,860	0,811	392.477	1,036	1,048	1,012
15	146.627	1,082	1,102	1,018	246.756	1,074	1,287	1,199
16	450.646	0,996	0,984	0,995	730.701	1,002	1,256	1,254
17	90.871	1,022	1,070	1,047	132.698	1,018	1,190	1,169
18	449.411	0,916	0,960	0,981	724.060	0,962	1,044	1,085
31	794.609	0,947	0,936	0,989	1.423.757	0,987	1,022	1,057
32	3.254.228	0,997	0,983	0,986	4.921.872	0,989	1,113	1,127
33	2.827.867	0,997	1,015	1,029	4.606.847	0,990	1,136	1,146
34	346.526	0,981	1,018	1,037	577.144	0,992	1,078	1,086
35	3.399.796	0,949	0,953	1,004	5.106.904	0,977	1,148	1,175
51	300.612	0,967	0,969	1,002	520.042	0,970	1,098	1,131
52	309.057	1,008	1,009	1,001	505.047	0,979	1,056	1,079
53	171.693	1,007	1,062	1,055	239.475	1,013	1,260	1,245
61	212.642	1,013	1,011	0,997	362.498	1,007	1,216	1,207
62	140.787	1,046	1,070	1,022	247.105	1,044	1,205	1,164
63	317.474	0,980	0,908	0,926	492.391	0,981	1,089	1,110
64	151.359	1,030	1,059	1,028	250.708	1,032	1,343	1,301
71	251.439	0,938	0,922	0,983	438.817	0,959	0,972	1,014
72	147.052	1,008	1,055	1,047	244.793	0,981	1,173	1,197
73	596.565	0,989	1,011	1,022	889.274	1,007	1,251	1,242
74	95.382	0,986	0,964	1,008	163.572	0,976	1,107	1,134
81	116.273	1,000	0,090	0,090	208.697	0,988	0,099	0,100
82	60.908	1,032	0,731	0,709	117.062	0,993	0,819	0,825
Total	16.166.563	0,980	0,966	0,985	25.829.032	0,968	1,114	1,128

TABLE 5

Estimation for SUSI 1999 (based on 3 quarters): crosstabulation by ISIC and province

A. Number of establishments

ISIC	all(2-9)				2				3				4			
	full EC96 estimate	EC96/SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96/SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96/SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96/SUSI areas	SUSI simple expansion	SUSI ratio estimation
Prov																
11	235.183	0,941	0,975	1,036	2.690	0,954	0,258	0,270	58.150	0,906	0,869	0,979	150	0,117	0,000	0,000
12	661.136	1,015	0,840	0,827	1.785	0,844	0,545	0,646	58.373	1,024	0,932	0,910	324	1,118	0,140	0,125
13	352.896	0,960	0,958	0,977	2.923	0,785	0,670	0,726	69.411	1,059	1,079	1,020	228	0,298	0,000	0,000
14	225.244	1,060	0,860	0,811	1.240	0,876	0,692	0,722	20.361	1,099	0,800	0,808	1.009	0,849	0,187	0,221
15	146.627	1,082	1,102	1,018	4.089	1,031	0,852	0,833	17.660	0,851	0,768	0,900	335	0,967	0,240	0,257
16	460.646	0,998	0,994	0,995	4.754	0,626	0,353	0,562	54.905	1,015	0,863	0,870	1.246	0,615	0,099	0,160
17	80.871	1,022	1,070	1,047	574	1,454	0,315	0,217	9.806	0,867	0,873	1,007	75	0,488	0,000	0,000
18	449.411	0,916	0,880	0,961	2.911	1,462	0,266	0,182	55.492	1,028	1,064	1,035	198	0,717	0,322	0,449
31	784.609	0,947	0,926	0,988	49	1,012	0,000	0,000	33.218	0,979	0,971	0,992	60	0,874	0,000	0,000
32	3.254.228	0,997	0,983	0,988	14.702	0,668	0,481	0,723	364.572	0,935	0,915	0,979	694	1,431	0,075	0,052
33	2.627.867	0,987	1,016	1,029	46.287	0,972	0,623	0,647	988.767	0,955	1,049	1,086	786	1,077	0,138	0,130
34	346.526	0,981	1,018	1,037	8.012	0,804	0,995	1,088	31.419	1,010	1,141	1,130	23	2,111	0,000	0,000
35	3.359.796	0,949	0,953	1,004	47.371	0,717	0,608	0,848	577.113	0,905	0,964	1,062	1,497	1,185	0,096	0,081
51	300.612	0,967	0,969	1,002	3.947	0,724	0,342	0,473	88.038	0,874	1,011	1,157	19	0,000	0,000	
52	309.067	1,008	1,000	1,001	7.043	0,783	1,154	1,474	73.174	1,065	1,123	1,052	6	0,000	0,000	
53	171.898	1,007	1,062	1,055	2.113	1,306	1,122	0,859	77.684	0,955	0,975	0,990	118	1,769	1,826	1,032
61	212.842	1,013	1,041	0,997	3.522	0,900	0,846	1,061	39.243	1,009	1,047	1,038	460	1,014	0,373	0,171
62	140.787	1,048	1,070	1,022	10.090	0,775	0,844	1,088	16.996	1,137	1,169	1,029	72	1,752	0,000	0,000
63	317.474	0,980	0,908	0,926	8.514	0,798	1,022	1,281	62.034	1,159	1,016	0,877	123	0,232	0,234	1,011
64	151.358	1,030	1,059	1,028	1.449	0,578	2,319	3,422	12.416	0,956	1,007	1,020	198	1,526	0,269	0,190
71	251.439	0,938	0,922	0,983	4.540	0,606	0,448	0,740	54.382	1,009	1,138	1,128	62	0,845	0,000	0,000
72	147.052	1,003	1,055	1,047	1.458	0,336	0,291	0,889	29.420	0,998	0,892	0,894	214	1,337	0,431	0,322
73	558.585	0,989	1,011	1,022	1.770	1,128	2,221	1,968	106.293	0,982	0,915	0,931	457	0,777	0,407	0,627
74	85.382	0,956	0,954	1,008	1.051	1,463	1,218	0,832	20.079	0,937	0,944	0,900	275	1,042	0,478	0,458
81	116.273	1,000	0,090	0,090	2.599	0,800	0,000	0,000	19.490	1,088	0,122	0,112	72	0,609	0,000	0,000
82	60.900	1,032	0,731	0,709	1.153	1,488	1,018	0,878	9.652	0,964	0,476	0,495	49	4,844	0,000	0,000
Total	18.166.589	0,980	0,986	0,985	180.780	0,841	0,730	0,868	2.729.222	0,964	0,984	1,022	9.568	1,002	0,101	0,161
CV across provinces																
	1,505	0,039	0,201	0,199	1,667	0,317	0,743	0,784	1,591	0,078	0,229	0,223	1,134	0,873	1,791	1,336

TABLE 5 (cont)

Estimation for SUSI 1999 (based on 3 quarters): crosstabulation by ISIC and province (cont)

A. Number of establishments (cont)

ISIC 5					6				7				8				9			
full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation
7.816	1,102	0,646	0,585	184.126	0,970	1,069	1,091	22.772	0,881	1,079	1,225	5.443	0,880	0,296	0,434	23.987	0,816	0,902	0,986
9.451	0,985	0,436	0,492	428.004	1,004	0,830	0,826	66.156	1,076	1,002	0,931	22.059	1,001	0,424	0,424	57.450	1,028	0,850	0,827
8.975	0,960	0,858	0,894	207.736	0,976	0,969	0,992	22.221	0,946	1,024	1,083	18.085	0,938	0,575	0,613	32.394	0,935	0,909	0,972
2.720	1,101	0,713	0,647	141.618	1,011	0,828	0,817	22.767	1,206	1,024	0,849	14.608	1,274	0,724	0,568	20.611	1,059	1,047	0,989
3.601	0,869	1,097	1,167	88.675	1,124	1,156	1,029	15.409	1,097	1,219	1,121	5.641	0,936	0,620	0,664	11.317	1,261	1,491	1,183
5.000	0,619	1,240	2,130	279.820	1,017	1,028	1,012	45.684	1,059	1,400	1,328	29.329	1,000	0,531	0,532	35.906	0,073	0,746	0,857
2.875	1,049	0,463	0,441	40.849	1,097	1,141	1,039	6.873	0,989	1,156	1,171	5.216	0,927	0,972	1,048	5.704	0,970	1,193	1,230
6.523	0,860	0,547	0,635	278.381	0,909	0,900	0,990	43.666	0,894	0,817	0,913	19.892	0,597	0,370	0,652	37.448	0,887	0,773	0,871
6.701	1,141	0,205	0,180	500.947	0,938	0,926	0,987	72.277	0,990	0,906	0,915	101.621	0,916	0,967	1,055	79.146	0,868	1,034	1,069
17.325	0,834	0,423	0,506	1.000.592	1,003	1,014	1,011	507.151	1,076	1,123	1,043	143.594	0,835	0,660	0,814	215.392	0,983	0,848	0,862
14.680	0,943	0,244	0,259	1.359.539	1,004	1,049	1,045	243.377	0,943	0,992	1,052	70.096	0,924	0,540	0,585	201.455	1,004	0,939	0,936
2.633	0,833	0,246	0,295	178.428	0,972	1,037	1,067	20.237	1,001	1,019	1,018	29.898	0,985	0,706	0,709	20.762	0,861	0,910	0,948
30.167	0,924	0,117	0,127	1.000.275	0,955	0,966	1,032	377.552	0,963	1,037	1,077	134.164	1,074	0,646	0,602	321.657	0,959	0,901	0,942
2.831	1,050	0,148	0,141	153.549	1,001	1,035	1,034	16.380	0,920	0,705	0,766	18.852	1,217	0,608	0,500	18.996	0,975	1,039	1,065
5.618	0,931	0,912	0,979	168.324	1,010	1,013	1,002	33.791	1,041	0,984	0,945	6.712	0,722	0,440	0,610	14.389	0,878	0,688	0,784
2.502	0,658	1,344	1,567	74.695	1,030	1,141	1,108	3.516	0,910	1,268	1,384	5.234	0,841	0,702	0,834	5.981	1,157	1,268	1,094
3.965	0,791	0,607	0,742	128.292	1,027	1,057	1,030	15.283	0,965	0,864	0,895	4.736	1,037	0,503	0,485	17.091	1,034	0,994	0,962
7.881	1,257	0,538	0,409	79.711	1,029	1,083	1,053	15.736	1,031	1,299	1,260	2.825	1,249	0,781	0,625	7.974	1,117	1,379	1,234
10.350	0,884	1,036	1,172	165.520	0,934	0,881	0,943	37.347	1,001	0,965	0,964	7.792	1,177	0,266	0,226	25.734	0,862	0,800	0,928
2.960	1,016	0,662	0,652	92.514	1,054	1,080	1,024	16.030	0,937	1,077	1,149	14.396	1,000	0,769	0,768	11.397	1,092	1,256	1,150
9.096	0,658	0,409	0,622	134.394	0,963	0,932	0,968	23.766	0,830	0,917	1,105	8.768	0,799	0,279	0,370	15.932	1,003	0,905	0,902
4.412	1,015	1,308	1,288	91.123	1,026	1,075	1,049	8.191	1,073	1,483	1,382	3.687	1,082	0,778	0,719	8.537	0,864	1,132	1,309
5.978	1,261	1,318	1,043	361.872	0,987	1,046	1,060	66.177	1,007	1,028	1,021	14.282	1,025	0,315	0,307	29.956	0,931	1,103	1,184
2.267	0,874	0,632	0,722	60.533	0,950	0,975	1,026	5.737	1,000	1,294	1,293	1.459	0,784	0,901	1,149	3.951	1,050	1,107	1,054
2.520	1,143	0,159	0,139	75.384	0,954	0,985	0,986	5.511	1,001	0,958	0,958	4.140	0,920	0,071	0,078	6.057	0,991	0,115	0,116
1.206	0,779	0,867	1,113	38.806	1,067	0,794	0,744	4.100	0,901	0,701	0,675	2.489	1,003	0,477	0,476	3.444	1,002	0,824	0,822
176.551	0,942	0,552	0,586	9.413.529	0,987	0,990	1,004	1.718.217	1,005	1,049	1,043	679.026	0,955	0,628	0,651	1.250.890	0,970	0,915	0,943
0,836	0,167	0,578	0,649	1,494	0,046	0,207	0,201	1,305	0,086	0,264	0,247	1,458	0,175	0,380	0,397	1,578	0,098	0,287	0,228

TABLE 5 (cont)

Estimation for SUSI 1999 (based on 3 quarters): crosstabulation by ISIC and province (cont)

B. Number of workers

ISIC	all(2-4)				2				3				4			
	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC96 estimate	EC96 SUSI areas	SUSI simple expansion	SUSI ratio estimation
Prov																
11	938.884	0,859	1,104	1,139	4.797	0,209	0,277	0,359	148.716	0,924	1,006	1,089	259	0,149	0,000	0,000
12	1.078.247	1,003	0,965	0,962	3.388	0,793	0,794	0,016	197.295	0,991	0,996	1,005	1.260	1,057	0,092	0,087
13	589,505	0,979	1,157	1,182	4.019	0,546	0,490	0,598	121.275	1,063	1,361	1,257	321	0,370	0,000	0,000
14	392.477	1,038	1,046	1,012	2.344	1,161	0,930	0,533	51.475	1,025	1,015	0,990	1.004	0,670	0,128	0,190
15	248.755	1,074	1,287	1,189	5.190	1,068	0,946	0,905	42.545	0,814	0,833	1,023	498	0,763	0,188	0,223
16	730.701	1,002	1,256	1,254	6.467	0,680	1,071	1,576	124.128	1,070	1,230	1,150	2.177	0,677	0,058	0,083
17	132.698	1,018	1,190	1,169	1.243	1,473	0,171	0,116	22.653	0,887	0,989	1,148	120	0,587	0,000	0,000
18	724.060	0,952	1,044	1,085	4.844	1,330	0,483	0,363	157.254	1,082	1,249	1,155	317	0,784	0,201	0,263
31	1.423.767	0,987	1,022	1,037	37	1,240	0,000	0,000	185.040	0,971	1,118	1,149	105	0,690	0,000	0,000
32	4.021.872	0,988	1,113	1,127	25.217	0,808	0,874	1,081	492.339	0,969	1,069	1,103	1.657	1,211	0,040	0,033
33	4.608.647	0,990	1,135	1,145	64.919	1,061	0,926	0,873	1.540.332	0,963	1,117	1,125	1.603	1,228	0,008	0,036
34	577,144	0,902	1,078	1,086	16,448	0,987	1,037	1,050	164.506	1,004	1,172	1,167	32	1,517	0,000	0,000
35	5,168,904	0,977	1,148	1,175	69,190	0,797	0,825	1,035	1.149.622	0,872	1,150	1,218	2.410	1,285	0,060	0,046
51	520.942	0,970	1,096	1,131	6.527	0,756	0,205	0,378	173.063	0,801	1,106	1,258	30	0,000	0,000	
52	505.047	0,979	1,058	1,079	12.435	0,896	1,283	1,293	140.689	1,014	1,037	1,023	14	0,000	0,000	
53	298.475	1,013	1,200	1,245	4.150	1,020	1,402	1,426	128.925	0,807	1,184	1,187	222	1,840	0,054	0,516
61	382.498	1,007	1,215	1,207	16.437	0,831	1,089	1,284	77.714	0,985	1,244	1,262	877	1,168	0,118	0,101
62	247.105	1,044	1,206	1,154	22.817	0,580	1,415	1,650	38.895	1,090	1,385	1,325	105	1,575	0,000	0,000
63	492.391	0,981	1,080	1,110	10.653	0,840	2,343	2,753	118.376	1,119	1,121	1,002	180	0,186	0,152	0,774
64	250.709	1,032	1,349	1,301	2.245	0,718	3,054	4,249	27.636	0,945	1,169	1,252	330	1,023	0,170	0,105
71	438.617	0,959	0,972	1,014	6.269	0,668	0,545	0,791	128.244	0,589	1,057	1,049	101	1,047	0,000	0,000
72	244.763	0,801	1,173	1,187	2.740	0,380	0,212	0,841	53.789	0,992	0,999	1,008	359	1,327	0,257	0,194
73	889.272	1,007	1,256	1,242	2.740	1,228	2,925	2,275	192.923	1,011	1,164	1,152	745	0,963	0,299	0,310
74	163.572	0,978	1,107	1,134	1.825	1,330	1,251	0,902	41.020	0,984	0,970	0,985	429	1,038	0,305	0,290
81	208.697	0,968	0,999	0,100	3.269	0,777	0,000	0,000	30.081	1,141	0,149	0,130	168	0,476	0,000	0,000
82	117.062	0,893	0,819	0,925	2.151	1,330	1,025	0,741	21.464	0,981	0,832	0,870	104	2,808	0,000	0,000
Total	25.829.082	0,888	1,114	1,128	208.579	0,908	0,988	1,088	5.892.150	0,988	1,119	1,132	15.863	1,009	0,109	0,108
CV across provinces																
	1,445	0,027	0,269	0,203	1,509	0,278	0,606	0,556	1,606	0,073	0,223	0,212	1,136	0,022	1,030	1,358

TABLE 5 (cont)

Estimation for SUSI 1999 (based on 3 quarters): crosstabulation by ISIC and province

B. Number of workers (cont)

ISIC 5				6				7				8				9			
full EC98 estimate	EC98 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC98 estimate	EC98 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC98 estimate	EC98 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC98 estimate	EC98 SUSI areas	SUSI simple expansion	SUSI ratio estimation	full EC98 estimate	EC98 SUSI areas	SUSI simple expansion	SUSI ratio estimation
10.991	1,243	0,729	0,584	288.193	0,984	1,238	1,256	31.872	0,800	0,975	1,076	7.538	0,629	0,254	0,404	38.991	1,023	1,071	1,047
16.279	1,035	0,533	0,516	692.326	1,002	1,005	1,003	81.825	1,051	0,947	0,872	31.359	0,591	0,461	0,486	114.314	0,888	0,920	0,931
16.509	1,005	1,426	1,419	337.443	0,837	1,150	1,227	32.784	1,011	1,049	1,038	25.453	0,894	0,569	0,583	50.081	0,885	1,089	1,054
5.333	1,051	0,918	0,898	247.748	1,014	1,059	1,045	28.403	1,198	1,091	0,912	17.860	1,210	0,744	0,611	37.610	0,994	1,217	1,225
7.272	1,030	1,695	1,646	140.838	1,129	1,404	1,247	20.065	1,129	1,265	1,138	7.282	0,862	0,047	0,078	18.084	1,281	1,796	1,390
11.341	0,760	1,785	2,322	440.071	1,007	1,327	1,318	56.595	1,070	1,409	1,317	30.601	1,010	0,809	0,997	67.318	0,839	0,923	1,104
8.171	1,033	0,605	0,589	77.358	1,071	1,330	1,242	6.301	0,996	1,109	1,131	8.578	0,838	1,100	1,340	9.761	0,985	1,161	1,187
15.245	0,823	0,958	1,157	414.482	0,945	1,057	1,118	57.489	0,950	0,858	0,895	20.060	0,604	0,440	0,728	53.760	0,886	0,840	0,954
15.151	0,850	0,535	0,544	853.221	0,965	1,043	1,081	84.334	1,000	0,805	0,805	129.799	0,918	0,918	0,998	176.020	0,891	1,012	1,022
41.543	0,857	0,425	0,496	2.651.790	0,969	1,167	1,201	500.709	1,068	1,133	1,060	188.830	0,921	0,655	0,711	351.730	0,859	0,989	0,880
36.577	0,928	0,243	0,262	2.218.942	0,983	1,238	1,248	305.551	0,947	1,017	1,074	175.903	0,967	0,585	0,603	308.821	0,899	0,971	0,972
10.465	0,938	0,367	0,372	279.798	0,978	1,104	1,129	29.839	1,048	1,193	1,144	30.099	1,067	0,730	0,672	40.537	0,930	0,970	1,043
51.063	0,826	0,227	0,245	2.664.802	0,880	1,231	1,267	480.340	0,977	1,021	1,045	200.346	1,077	0,991	0,841	478.131	0,861	1,088	1,130
9.849	0,832	0,008	0,090	248.772	1,009	1,210	1,204	20.728	0,986	0,815	0,826	22.005	1,136	0,943	0,988	37.910	1,053	1,106	1,051
17.535	0,850	1,227	1,332	248.724	0,979	1,109	1,136	46.072	0,992	0,964	1,001	9.604	0,775	0,513	0,662	23.784	0,931	0,900	0,966
8.981	0,924	1,087	1,178	125.855	1,033	1,331	1,338	12.919	0,991	1,045	1,161	8.230	0,810	0,829	1,023	12.858	1,251	1,369	1,110
7.093	0,831	1,008	1,211	229.605	1,019	1,276	1,253	21.110	1,042	0,829	0,891	8.738	0,944	0,718	0,759	28.984	1,096	1,137	1,037
12.448	1,243	0,940	0,878	186.589	1,038	1,157	1,117	20.966	1,072	1,216	1,134	8.081	1,229	0,779	0,634	13.351	1,174	1,351	1,151
24.442	0,824	1,068	1,295	247.256	0,839	1,141	1,182	42.787	0,987	1,021	1,034	10.190	1,076	0,632	0,401	38.275	0,998	0,775	0,883
7.155	1,228	0,761	0,619	154.707	1,045	1,451	1,389	19.058	0,985	1,064	1,101	18.051	0,976	0,931	0,872	20.782	1,100	1,452	1,326
20.216	0,710	0,899	0,985	235.787	0,979	0,907	1,019	35.293	0,913	0,926	1,017	18.204	0,937	0,924	0,948	24.843	0,971	1,063	1,064
11.550	0,987	1,696	1,703	138.654	1,002	1,193	1,191	12.940	0,988	1,537	1,538	5.602	1,094	0,782	0,715	14.370	0,825	1,384	1,677
16.838	1,324	1,591	1,202	524.727	0,967	1,317	1,334	79.516	1,020	1,129	1,108	22.571	1,193	0,975	0,816	48.444	0,977	1,288	1,318
5.855	0,986	0,815	0,843	85.234	0,969	1,179	1,217	10.312	0,998	1,085	1,086	2.258	0,909	1,273	1,373	6.538	0,948	1,151	1,214
5.651	1,163	0,574	0,150	137.531	0,957	0,089	0,093	9.061	1,006	0,041	0,040	0.025	0,858	0,110	0,134	10.101	0,828	0,097	0,169
3.498	0,741	0,852	0,893	71.824	1,030	0,893	0,867	7.690	0,906	0,589	0,854	3.641	0,947	0,531	0,561	6.707	1,017	0,960	0,945
418.755	0,981	0,731	0,780	14.103.262	0,969	1,178	1,193	2.111.960	1,031	1,054	1,042	168.020	0,988	0,670	0,660	2.027.405	0,970	1,022	1,047
0.786	0,160	0,558	0,585	1.411	0,041	0,218	0,209	1.723	0,075	0,258	0,242	1,436	0,161	0,394	0,438	1,499	0,110	0,275	0,241

TABLE 6 (cont)

SUSI0 Estimates by subsample

B. Number of workers ISIC all(2-9)

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	Total	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	538.384	1,031	0,560	0,918	0,969	1,229	1,118	0,968	1,104	1,189	1,166	1,065	1,139
12	1.078.247	0,982	1,002	1,026	1,003	1,151	1,181	0,563	0,965	1,172	1,178	0,548	0,962
13	589.935	1,020	0,917	1,000	0,979	1,214	1,133	1,125	1,157	1,190	1,206	1,125	1,182
14	392.477	0,956	1,086	1,053	1,035	1,112	1,088	0,950	1,048	1,152	0,988	0,902	1,012
15	246.755	1,127	1,005	1,090	1,074	1,391	1,218	1,253	1,287	1,234	1,212	1,150	1,199
16	730.701	1,056	1,016	0,834	1,002	1,302	1,342	1,123	1,258	1,233	1,322	1,203	1,234
17	182.698	0,998	1,128	1,019	1,018	1,109	1,221	1,243	1,190	1,217	1,084	1,220	1,169
18	724.060	0,993	0,934	0,960	0,962	1,080	1,024	1,028	1,044	1,089	1,098	1,071	1,085
31	1.423.757	0,972	0,948	0,982	0,967	1,023	0,989	1,053	1,022	1,053	1,045	1,072	1,057
32	4.921.872	1,071	0,974	0,920	0,988	1,109	1,141	1,030	1,113	1,092	1,171	1,119	1,127
33	4.606.847	0,999	0,997	0,975	0,990	1,119	1,160	1,105	1,135	1,120	1,183	1,133	1,146
34	377.144	0,981	1,003	0,964	0,992	1,038	1,131	1,064	1,078	1,058	1,095	1,104	1,086
35	5.106.934	0,988	0,961	0,992	0,977	1,146	1,145	1,151	1,148	1,160	1,204	1,161	1,176
51	520.042	0,983	1,016	0,930	0,970	1,141	1,176	0,977	1,098	1,184	1,155	1,050	1,131
52	505.047	0,920	1,030	0,988	0,979	1,103	1,022	1,044	1,056	1,199	0,993	1,067	1,079
53	239.673	1,024	1,000	1,016	1,013	1,218	1,193	1,350	1,260	1,301	1,198	1,350	1,245
61	382.698	1,119	0,968	0,938	1,007	1,512	1,253	0,879	1,215	1,392	1,301	0,937	1,297
62	247.105	0,940	1,196	0,998	1,044	1,226	1,307	1,086	1,205	1,305	1,098	1,084	1,154
63	492.391	0,937	0,991	1,040	0,981	1,099	0,923	1,246	1,089	1,173	0,932	1,226	1,110
64	250.708	1,171	1,009	0,947	1,032	1,462	1,357	1,210	1,343	1,240	1,345	1,319	1,301
71	438.817	0,945	0,949	0,982	0,959	0,994	0,972	0,951	0,972	1,081	1,024	0,960	1,014
72	244.793	0,940	1,083	0,949	0,981	1,116	1,241	1,164	1,173	1,187	1,146	1,288	1,197
73	389.274	1,002	1,041	0,978	1,007	1,275	1,341	1,137	1,251	1,273	1,257	1,163	1,242
74	163.572	1,009	0,964	0,956	0,976	1,114	1,093	1,113	1,107	1,104	1,134	1,185	1,134
81	208.097	0,878	0,948	1,149	0,988	0,298	0,000	0,000	0,099	0,339	0,000	0,000	0,100
82	117.062	0,951	0,920	1,108	0,993	1,327	0,918	0,411	0,819	1,388	0,998	0,371	0,825
Total	25 829.032	1,008	0,983	0,972	0,988	1,147	1,138	1,056	1,114	1,130	1,158	1,080	1,128
CV across provinces													
	1,645	0,066	0,065	0,059	0,027	0,103	0,238	0,279	0,209	0,155	0,222	0,283	0,203

TABLE 7

The direction of change from one quarter to the next - SUSI99 subsamples

A. Number of establishments

ISIC all(2-9)	Estimates from SUSI 1999 (3 quarters) Linear (simple expansion)			whether same direction	decline or increase?	Estimates from SUSI 1999 (3 quarters) Ratio estimation			whether same direction	decline or increase?
	Subsample					Subsample				
Prov	1	2	3			1	2	3		
11	359,625	345,506	273,171	1	-2	381,437	352,187	324,738	1	-2
12	646,871	600,490	310,797	-1	0	681,984	682,520	299,535	-1	0
13	358,540	316,201	338,877	-1	0	355,762	341,882	336,657	1	-2
14	215,740	199,290	165,779	1	-2	229,890	171,035	154,702	1	-2
15	180,720	141,627	162,417	-1	0	149,195	149,750	150,165	1	2
16	473,754	461,196	408,519	1	-2	437,307	444,934	466,784	1	2
17	76,894	92,490	90,290	-1	0	95,554	82,764	85,902	-1	0
18	407,505	300,400	337,997	-1	0	416,658	438,241	441,697	1	2
31	752,003	736,004	742,411	-1	0	802,820	797,552	758,313	1	-2
32	3,316,527	3,270,066	2,999,522	1	-2	3,151,143	3,230,447	3,245,180	1	2
33	2,634,920	2,920,614	2,007,126	-1	0	2,950,698	2,888,912	2,880,038	1	-2
34	316,264	389,950	352,137	-1	0	342,292	388,121	368,703	1	2
35	3,254,580	3,155,648	3,270,062	-1	0	3,405,038	3,429,825	3,373,563	-1	0
51	290,699	327,233	255,651	-1	0	308,197	308,289	285,812	-1	0
52	318,874	307,582	309,402	-1	0	326,377	330,779	301,868	-1	0
53	183,286	181,392	183,196	-1	0	186,314	181,373	176,537	1	-2
61	280,067	217,716	167,002	1	-2	226,757	237,330	178,542	-1	0
62	149,036	167,818	134,862	-1	0	149,762	137,318	146,339	-1	0
63	293,293	270,537	310,845	-1	0	303,150	285,588	308,898	-1	0
64	166,463	163,656	150,725	1	-2	154,924	154,781	157,208	-1	0
71	244,466	220,345	230,995	-1	0	250,720	251,544	239,694	-1	0
72	150,030	167,164	147,623	-1	0	151,532	158,356	153,851	-1	0
73	590,168	633,764	555,552	-1	0	598,213	615,383	583,564	-1	0
74	100,892	81,818	93,583	-1	0	92,566	95,798	100,680	1	2
81	81,549	0	0	0	-1	33,750	0	0	0	-1
82	63,673	49,817	21,288	1	-2	66,410	53,464	17,828	1	-2
Total	16,079,013	15,888,372	14,889,910	1	-2	16,206,706	16,092,948	15,481,375	1	-2

direction of change		if change in same direction		direction of change		if change in same direction	
same (>=0)	6	decline (<0)	6	same (>=0)	14	decline (<0)	6
different (-1)	16	increase (>0)	0	different (-1)	12	increase (>0)	6
		no record	0			no record	0

TABLE 7 (cont)
The direction of change from one quarter to the next - SUSI99 subsamples

B. Number of workers

ISIC all(2-9)	Estimates from SUSI 1999 (3 quarters) Linear (simple expansion)			whether same direction	decline or increase?	Estimates from SUSI 1999 (3 quarters) Ratio estimation			whether same direction	decline or increase?
	Subsample					Subsample				
Prov	1	2	3			1	2	3		
11	660 027	601.963	521.196	1	-2	640.112	627.664	587.820	1	-2
12	1.240.997	1.273.268	606.531	-1	0	1.263.681	1.270.619	591.407	-1	0
13	716.353	688.221	663.633	1	-2	702.203	729.921	683.834	-1	0
14	436.340	421.919	372.974	1	-2	455.341	385.954	364.174	1	-2
15	243.141	300.566	309.140	-1	0	304.575	299.173	283.671	1	-2
16	951.366	900.849	820.878	-1	0	900.756	965.852	879.114	-1	0
17	146.598	162.047	164.943	1	2	161.493	148.899	151.909	-1	0
18	731.902	741.412	744.308	-1	0	787.767	793.755	775.215	-1	0
31	1.457.085	1.408.900	1.493.504	-1	0	1.490.677	1.489.718	1.528.347	-1	0
32	5.758.112	5.614.299	5.070.139	1	-2	6.375.287	6.765.251	5.508.195	-1	0
33	5.154.285	5.436.511	5.088.802	-1	0	5.160.054	5.460.349	6.218.548	-1	0
34	599.046	652.528	614.091	-1	0	610.758	631.872	637.104	1	2
35	5.852.986	5.618.238	5.880.088	-1	0	5.932.396	6.150.119	5.920.027	-1	0
51	593.168	611.768	507.849	-1	0	615.968	600.808	546.278	1	-2
52	657.143	516.404	527.165	-1	0	805.608	901.389	583.786	-1	0
53	395.156	257.225	410.011	-1	0	356.744	357.141	404.412	1	2
61	578.364	479.320	336.069	1	-2	516.958	497.839	358.394	1	-2
62	303.039	323.040	267.181	-1	0	322.362	270.193	267.830	1	-2
63	541.215	454.384	613.431	-1	0	677.702	458.725	603.608	-1	0
64	388.548	340.192	308.949	1	-2	312.097	337.324	330.748	-1	0
71	438.165	428.577	417.227	1	-2	451.232	449.334	424.092	1	-2
72	273.001	303.780	284.977	-1	0	290.523	280.417	310.017	-1	0
73	1.133.402	1.192.248	1.011.258	-1	0	1.131.659	1.144.792	1.034.006	-1	0
74	182.294	178.770	182.106	-1	0	180.598	185.479	190.487	1	2
81	82.122	0	0	0	-1	70.749	0	0	0	-1
82	131.972	107.473	48.163	1	-2	139.834	116.806	43.407	1	-2
Total	29.618.615	29.404.806	27.265.148	1	-2	29.388.894	29.907.425	23.060.317	-1	0

direction of change		If change in same direction	direction of change		if change in same direction
same (>=0)	10	decline (<0) 9	same (>=0)	12	decline (<0) 9
different (-1)	16	increase (>0) 1	different (-1)	14	increase (>0) 3
		no record 0			no record 0

ANNEX TABLE 9

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of establishments ISIC 2

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c2	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	2.690	1,210	1,153	0,900	0,954	0,352	0,190	0,231	0,258	0,291	0,166	0,452	0,270
12	1,769	0,479	0,822	1,732	0,844	0,970	0,558	0,109	0,546	2,026	1,733	0,063	0,646
13	2.823	0,469	0,565	1,331	0,785	0,240	0,414	1.057	0,570	0,522	0,783	0,794	0,728
14	1,240	1,295	1,001	0,832	0,876	1,133	0,569	0,195	0,632	0,875	0,563	0,567	0,722
15	4.089	1,643	0,460	0,969	1,031	0,215	0,207	1,534	0,652	0,131	0,432	1,584	0,633
16	4,754	0,492	0,976	0,415	0,628	0,295	0,258	0,505	0,353	0,589	0,264	1,217	0,462
17	574	1,586	2,324	0,451	1,454	0,144	0,577	0,224	0,315	0,091	0,248	0,497	0,217
18	2.911	3,160	0,504	0,722	1,452	0,284	0,395	0,148	0,266	0,090	0,725	0,205	0,162
31	49	1,080	1,849	0,000	1,012	0,000	0,000	0,000	0,000	0,000	0,000		0,000
32	14.702	0,522	0,582	0,893	0,698	0,471	0,438	0,584	0,481	0,603	0,752	0,880	0,723
33	48.267	1,246	1,081	0,867	0,972	1,411	0,794	0,288	0,623	1,132	0,734	0,451	0,847
34	9.042	1,149	0,921	0,780	0,934	1,406	1,024	0,555	0,995	1,223	1,112	0,760	1,046
35	47.371	0,538	0,537	0,975	0,717	0,591	0,490	0,838	0,606	0,988	0,703	0,856	0,845
51	3.947	1,396	0,667	0,108	0,724	0,513	0,207	0,207	0,342	0,439	0,310	1,825	0,473
52	7.043	0,339	1,169	0,840	0,763	0,849	1,772	0,841	1,154	2,503	1,515	1,001	1,474
53	2 113	0,046	1,709	2,162	1,306	2,091	1,280	0,084	1,122	43,399	0,749	0,039	0,239
61	3.622	0,869	0,549	1,152	0,900	1,315	0,459	1,066	0,949	1,316	0,835	0,924	1,051
62	10 090	0,553	0,789	0,985	0,770	0,715	0,860	0,856	0,044	1,292	1,091	0,970	1,058
63	5.614	0,537	1,297	0,161	0,798	1,314	0,779	0,973	1,022	1,403	0,501	6,044	1,261
64	1.449	1,240	0,739	0,054	0,578	3,129	3,021	0,260	2,319	2,532	4,764	5,459	3,622
71	4.540	0,735	0,278	0,503	0,805	0,876	0,294	0,173	0,448	1,192	1,050	0,216	0,740
72	1.488	0,000	0,904	0,101	0,235	0,134	0,585	0,144	0,281		0,638	1,422	0,889
73	1.770	0,487	2,145	0,743	1.128	0,774	3,384	1,907	2,221	1,558	1,857	2,565	1,969
74	1,081	0,916	1,780	1,695	1,403	1,662	1,541	0,461	1,218	1,815	0,866	0,266	0,832
81	2.699	1,032	0,729	0,639	0,800	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
82	1.153	2,146	1,845	0,503	1,498	1,116	1,795	0,123	1,043	0,520	0,873	0,255	0,676
Total	189.780	0,882	0,866	0,774	0,841	0,894	0,702	0,393	0,730	1,014	0,811	0,767	0,888
CV across provinces													
	1,667	0,667	0,545	0,704	0,817	0,833	1,088	0,937	0,743	3,121	0,997	1,273	0,784

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of workers ISIC 9

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c0	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	38.991	0,919	1,151	0,899	1,023	1,056	1,267	0,889	1,071	1,148	1,101	0,890	1,047
12	114.314	0,951	0,974	1,089	0,998	1,086	1,233	0,485	0,829	1,113	1,266	0,453	0,931
13	50.091	1,073	0,964	0,888	0,985	1,030	1,096	0,981	1,029	0,968	1,105	1,105	1,034
14	37.810	0,985	1,051	0,964	0,994	0,943	1,447	1,261	1,217	0,876	1,377	1,308	1,225
15	16.064	1,400	1,213	1,223	1,281	2,278	0,931	2,094	1,768	1,818	0,768	1,712	1,380
16	57.318	0,739	0,935	0,837	0,836	0,816	0,966	0,806	0,923	1,109	1,034	1,178	1,104
17	9.701	0,076	1,064	1,016	0,995	1,138	1,369	1,038	1,181	1,300	1,251	1,020	1,187
18	53.780	0,901	0,863	0,868	0,888	0,813	0,824	0,880	0,846	0,880	1,054	1,124	0,954
31	176.020	0,800	0,837	1,056	0,991	1,023	0,918	1,101	1,012	1,044	0,974	1,043	1,022
32	351.730	0,998	1,040	0,839	0,959	0,877	1,182	0,769	0,939	0,879	1,136	0,904	0,989
33	309.621	1,018	0,962	1,016	0,999	0,922	0,957	1,034	0,971	0,905	0,985	1,017	0,972
34	46.537	0,911	1,010	0,870	0,930	0,883	1,231	0,787	0,970	0,848	1,338	0,917	1,043
35	476.131	0,981	0,890	1,014	0,981	1,142	1,008	1,108	1,088	1,184	1,133	1,093	1,130
51	87.010	0,916	1,175	1,088	1,053	1,143	1,218	0,937	1,166	1,240	1,037	0,896	1,091
52	23.784	0,917	0,974	0,910	0,901	1,240	0,456	0,703	0,800	1,517	0,468	0,772	0,888
53	12.058	1,406	1,210	1,037	1,251	1,260	1,101	1,807	1,389	0,896	0,840	1,742	1,110
61	28.384	1,409	0,902	0,977	1,066	1,560	0,818	0,644	1,137	1,383	0,909	0,659	1,037
62	18.551	0,940	1,289	1,293	1,174	1,248	1,572	1,234	1,351	1,326	1,219	0,854	1,161
63	38.275	0,825	0,911	0,960	0,889	0,489	0,764	1,068	0,775	0,605	0,838	1,107	0,883
64	20.782	1,310	0,969	0,993	1,100	1,893	1,433	1,029	1,452	1,465	1,435	1,040	1,320
71	24.843	0,924	0,951	1,039	0,971	0,560	1,245	0,964	1,053	1,029	1,309	0,928	1,084
72	14.370	0,917	0,747	0,812	0,825	1,818	1,068	1,471	1,354	1,763	1,427	1,811	1,677
73	49.464	1,129	0,893	0,910	0,977	1,455	1,485	0,922	1,288	1,289	1,663	1,014	1,316
74	6.538	1,094	0,849	0,901	0,948	1,463	0,807	1,164	1,151	1,356	0,950	1,292	1,214
81	10.181	0,934	0,741	0,809	0,838	0,291	0,000	0,000	0,097	0,312	0,000	0,000	0,109
82	6.707	1,149	1,047	0,859	1,017	1,269	1,344	0,240	0,960	1,124	1,283	0,293	0,945
Total	2.027.405	0,980	0,966	0,972	0,976	1,095	1,092	0,969	1,022	1,046	1,099	0,897	1,047
CV across provinces	1,499	0,177	0,142	0,113	0,110	0,373	0,310	0,416	0,275	0,278	0,302	0,387	0,241

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of establishments ISIC 3

Prov	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas (incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
	c3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	88.150	0,930	1,028	0,785	0,908	1,047	0,939	0,681	0,889	1,125	0,914	0,690	0,879
12	58.278	0,936	1,161	0,976	1,024	1,126	1,224	0,447	0,932	1,201	1,054	0,458	0,910
13	39.411	1,025	1,032	1,106	1,058	1,015	1,107	1,115	1,079	0,981	1,073	1,006	1,020
14	20.881	0,994	1,283	1,021	1,099	0,917	1,162	0,805	0,888	0,923	0,890	0,593	0,808
15	17.660	0,746	0,858	0,950	0,851	0,984	0,950	0,685	0,768	0,889	1,107	0,721	0,800
16	54.905	1,226	0,972	0,849	1,046	1,098	0,071	0,651	0,883	0,898	0,897	0,801	0,870
17	9.806	0,828	0,879	0,894	0,867	1,037	0,844	0,737	0,873	1,283	0,861	0,824	1,007
18	89.492	1,081	1,179	0,823	1,028	1,070	1,184	0,928	1,064	0,989	1,013	1,127	1,035
31	33.218	1,071	0,930	0,936	0,979	0,879	0,973	1,061	0,971	0,821	1,045	1,134	0,862
32	384.572	0,925	0,931	0,949	0,935	0,855	0,913	0,878	0,815	0,924	0,981	1,031	0,979
33	668.767	0,975	0,961	0,960	0,965	1,072	1,029	1,044	1,049	1,100	1,071	1,058	1,066
34	81.418	0,930	1,089	1,012	1,010	1,011	1,225	1,185	1,141	1,087	1,124	1,175	1,130
35	327.113	0,881	0,927	0,905	0,908	0,999	1,029	0,936	0,964	1,039	1,110	1,035	1,062
51	88.038	0,875	0,876	0,771	0,874	1,105	1,069	0,861	1,011	1,133	1,220	1,118	1,157
52	73.174	0,988	1,285	0,950	1,068	1,048	1,372	0,952	1,123	1,081	1,067	1,002	1,052
53	77.684	0,904	1,084	0,857	0,985	0,937	1,089	0,899	0,979	1,039	0,998	0,939	0,990
61	39.243	1,157	0,985	0,927	1,009	1,111	1,214	0,817	1,047	0,961	1,258	0,901	1,036
62	16.996	1,072	1,285	0,753	1,137	1,464	1,134	0,910	1,169	1,087	0,882	1,210	1,028
63	62.094	1,341	1,088	1,051	1,159	1,275	0,876	0,859	1,016	0,851	0,807	0,855	0,877
64	12.416	0,870	0,953	1,037	0,956	1,057	0,938	1,025	1,007	1,090	0,984	0,989	1,020
71	54.362	0,929	0,964	1,144	1,009	1,070	1,011	1,328	1,135	1,152	1,060	1,161	1,129
72	29.420	1,010	0,890	1,085	0,995	0,906	0,743	1,028	0,892	0,890	0,833	0,948	0,894
73	106.293	0,859	0,944	1,145	0,982	0,786	0,857	1,301	0,915	0,915	0,908	0,962	0,931
74	20.079	1,034	0,709	1,070	0,937	0,809	0,781	0,941	0,844	0,783	1,102	0,880	0,900
81	19.490	0,971	1,356	0,938	1,083	0,385	0,000	0,000	0,122	0,375	0,000	0,000	0,112
82	9.652	1,046	0,621	1,226	0,964	0,040	0,373	0,212	0,478	0,311	0,601	0,173	0,498
Total	2.729.222	0,962	0,961	0,947	0,964	0,987	1,012	0,964	0,984	1,025	1,032	1,007	1,022
CV across provinces													
	1,091	0,138	0,170	0,122	0,076	0,204	0,285	0,337	0,229	0,174	0,244	0,327	0,223

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of establishments — ISIC 4

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	04	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	190	0,351	0,000	0,000	0,117	0,000	0,000	0,000	0,000	0,000			0,000
12	824	1,119	0,961	1,275	1,118	0,041	0,326	0,064	0,140	0,037	0,339	0,042	0,125
13	229	0,000	0,594	0,000	0,298	0,000	0,000	0,000	0,000		0,000		0,000
14	1.099	0,497	1,061	0,991	0,846	0,077	0,148	0,235	0,187	0,157	0,140	0,339	0,221
15	335	1,056	0,688	1,257	0,967	0,271	0,214	0,159	0,248	0,352	0,364	0,120	0,257
16	1,249	0,885	0,517	0,463	0,615	0,137	0,094	0,064	0,099	0,150	0,182	0,139	0,160
17	75	1,130	0,000	0,328	0,485	0,000	0,000	0,000	0,000	0,000		0,000	0,000
18	198	0,000	1,014	1,137	0,717	0,074	0,550	0,342	0,322		0,542	0,301	0,449
31	60	0,000	0,727	1,894	0,874	0,000	0,000	0,000	0,000		0,000	0,000	0,000
32	894	1,258	1,797	1,238	1,431	0,066	0,045	0,093	0,076	0,068	0,025	0,075	0,052
33	786	1,400	1,205	0,617	1,077	0,046	0,304	0,069	0,139	0,033	0,232	0,111	0,130
34	23	0,000	6,332	0,000	2,111	0,000	0,000	0,000	0,000		0,000		0,000
35	1.497	0,927	0,507	2,122	1,185	0,195	0,031	0,062	0,096	0,210	0,061	0,029	0,081
51	19	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000				
52	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000				
53	116	4,312	0,000	0,995	1,769	4,312	0,853	0,312	1,825	1,000		0,314	1,032
61	460	1,142	1,786	0,106	1,014	0,185	0,228	0,106	0,173	0,162	0,127	1,000	0,171
62	72	0,347	3,889	1,071	1,752	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
63	123	0,695	0,000	0,000	0,232	0,258	0,186	0,249	0,234	0,366			1,011
64	198	1,708	1,744	1,129	1,586	0,187	0,181	0,480	0,289	0,115	0,109	0,428	0,190
71	62	1,244	1,290	0,000	0,845	0,000	0,000	0,000	0,000	0,000	0,000		0,000
72	214	0,090	1,047	2,273	1,337	0,462	0,352	0,478	0,431	0,550	0,336	0,211	0,332
73	457	0,605	0,515	1,010	0,777	0,275	0,515	0,670	0,487	0,349	1,000	0,663	0,627
74	275	0,387	0,968	1,771	1,042	0,112	0,129	1,192	0,478	0,280	0,134	0,673	0,456
81	72	0,000	0,706	1,130	0,609	0,000	0,000	0,000	0,000		0,000	0,000	0,000
82	48	0,000	0,347	14,185	4,844	0,000	0,000	0,000	0,000		0,000	0,000	0,000
Total	8.058	0,833	0,948	1,540	1,002	0,182	0,174	0,167	0,181	0,188	0,185	0,184	0,181
CV across provinces	1,124	1,148	1,289	1,077	0,873	3,115	1,324	1,622	1,791	1,150	1,338	1,174	1,336

ANNEX TABLE 9 (cont)
SUSI98 Estimates by subsample - and ISIC (details on Table 6)

Number of establishments ISIC 3

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	85.150	0,930	1,025	0,765	0,908	1,047	0,809	0,681	0,888	1,125	0,914	0,830	0,979
12	58.878	0,933	1,161	0,978	1,024	1,125	1,224	0,447	0,932	1,201	1,054	0,458	0,910
13	59.411	1,035	1,032	1,106	1,086	1,015	1,107	1,115	1,078	0,981	1,073	1,005	1,020
14	20.381	0,994	1,383	1,021	1,099	0,917	1,142	0,805	0,888	0,928	0,890	0,593	0,808
15	17.590	0,746	0,050	0,830	0,634	0,654	0,350	0,585	0,766	0,339	1,107	0,721	0,900
16	54.906	1,225	0,972	0,840	1,045	1,098	0,871	0,681	0,883	0,896	0,897	0,801	0,870
17	8.806	0,828	0,078	0,084	0,607	1,037	0,044	0,737	0,073	1,253	0,961	0,824	1,007
18	66.492	1,081	1,179	0,823	1,028	1,070	1,194	0,928	1,064	0,980	1,013	1,127	1,035
31	33.216	1,071	0,930	0,836	0,978	0,879	0,973	1,061	0,971	0,821	1,046	1,134	0,992
32	384.572	0,925	0,931	0,949	0,935	0,855	0,913	0,978	0,915	0,924	0,981	1,031	0,979
33	605.767	0,975	0,961	0,960	0,963	1,072	1,029	1,044	1,048	1,100	1,071	1,066	1,066
34	81.418	0,930	1,080	1,012	1,010	1,011	1,225	1,168	1,141	1,087	1,124	1,175	1,130
35	977.113	0,891	0,927	0,905	0,906	0,825	1,029	0,836	0,964	1,039	1,110	1,035	1,062
51	88.038	0,975	0,878	0,771	0,874	1,105	1,069	0,881	1,011	1,138	1,220	1,118	1,167
52	73.174	0,968	1,265	0,950	1,066	1,046	1,372	0,832	1,123	1,051	1,067	1,002	1,052
53	77.684	0,904	1,094	0,557	0,985	0,937	1,089	0,809	0,975	1,036	0,999	0,939	0,990
61	39.213	1,157	0,965	0,907	1,009	1,111	1,214	0,847	1,017	0,981	1,258	0,901	1,036
62	18.996	1,372	1,285	0,768	1,187	1,484	1,134	0,910	1,169	1,067	0,882	1,210	1,029
63	62.091	1,341	1,066	1,051	1,159	1,275	0,876	0,888	1,016	0,951	0,807	0,855	0,877
64	12.416	0,970	0,953	1,087	0,986	1,057	0,938	1,025	1,007	1,090	0,984	0,989	1,020
71	54.352	0,929	0,954	1,111	1,009	1,070	1,011	1,326	1,136	1,152	1,060	1,161	1,126
72	29.420	1,048	0,803	1,085	0,998	0,908	0,743	1,028	0,892	0,890	0,833	0,948	0,894
73	106.283	0,859	0,944	1,115	0,952	0,786	0,857	1,101	0,915	0,915	0,908	0,962	0,931
74	20.079	1,084	0,700	1,070	0,937	0,800	0,781	0,941	0,844	0,783	1,102	0,880	0,900
81	19.490	0,971	1,356	0,938	1,033	0,363	0,000	0,000	0,122	0,375	0,000	0,000	0,112
82	9.652	1,048	0,821	1,228	0,984	0,648	0,373	0,212	0,478	0,811	0,801	0,179	0,498
Total	2.729.222	0,962	0,801	0,947	0,964	0,887	1,012	0,554	0,984	1,035	1,032	1,007	1,022
CV across provinces													
	1,591	0,189	0,170	0,122	0,078	0,204	0,285	0,337	0,229	0,174	0,244	0,327	0,223

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of establishments ISIC 4

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 (cases cross tot. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	ed	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	190	0,351	0,000	0,000	0,117	0,000	0,000	0,000	0,000	0,000			0,000
12	824	1,119	0,061	1,275	1,118	0,041	0,326	0,064	0,140	0,037	0,339	0,042	0,125
13	229	0,000	0,894	0,000	0,280	0,000	0,000	0,000	0,000		0,000		0,000
14	1.000	0,487	1,061	0,991	0,849	0,077	0,146	0,336	0,187	0,157	0,140	0,339	0,221
15	335	1,056	0,583	1,267	0,967	0,271	0,214	0,159	0,248	0,252	0,264	0,126	0,237
16	1.248	0,865	0,517	0,463	0,815	0,137	0,094	0,064	0,099	0,159	0,162	0,129	0,160
17	75	1,130	0,000	0,320	0,486	0,000	0,000	0,000	0,000	0,000		0,000	0,000
18	198	0,000	1,014	1,137	0,717	0,074	0,550	0,342	0,322		0,542	0,301	0,449
31	50	0,000	0,727	1,864	0,874	0,000	0,000	0,000	0,000		0,000	0,000	0,000
32	804	1,258	1,797	1,238	1,431	0,086	0,045	0,093	0,075	0,066	0,025	0,075	0,052
33	786	1,408	1,205	0,917	1,077	0,045	0,304	0,065	0,139	0,083	0,252	0,111	0,130
34	28	0,000	8,332	0,000	2,111	0,000	0,000	0,000	0,000		0,000		0,000
35	1.497	0,927	0,507	2,122	1,105	0,165	0,031	0,062	0,096	0,210	0,061	0,029	0,081
51	10	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000				
52	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000				
53	116	4,812	0,000	0,996	1,769	4,812	0,853	0,312	1,828	1,000		0,314	1,032
61	480	1,142	1,786	0,106	1,014	0,165	0,228	0,106	0,173	0,162	0,127	1,000	0,171
62	72	0,247	3,839	1,071	1,752	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
63	123	0,695	0,000	0,000	0,232	0,260	0,186	0,249	0,234	0,386			1,011
64	198	1,708	1,744	1,126	1,528	0,197	0,191	0,180	0,289	0,115	0,109	0,426	0,190
71	82	1,244	1,250	0,000	0,845	0,000	0,000	0,000	0,000	0,000	0,000		0,000
72	214	0,590	1,047	2,273	1,337	0,462	0,352	0,479	0,431	0,669	0,336	0,211	0,322
73	457	0,806	0,515	1,010	0,777	0,276	0,515	0,670	0,487	0,248	1,000	0,663	0,627
74	275	0,887	0,068	1,771	1,043	0,112	0,129	1,192	0,478	0,289	0,134	0,573	0,458
81	72	0,000	0,705	1,120	0,609	0,000	0,000	0,000	0,000		0,000	0,000	0,000
82	48	0,000	0,347	14,185	4,844	0,000	0,000	0,000	0,000		0,000	0,000	0,000
Total	8.660	0,923	0,943	1,140	1,002	0,182	0,174	0,187	0,181	0,198	0,185	0,164	0,181
CV across provinces	1,324	1,148	1,289	1,977	0,873	3,115	1,324	1,522	1,791	1,159	1,335	1,174	1,336

ANNEX TABLE B (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 5)

Number of establishments ISIC 5

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas (not found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c5	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	7.615	1,419	0,649	1,239	1,102	0,777	0,399	0,792	0,645	0,547	0,594	0,580	0,585
12	6.451	0,748	1,130	1,050	0,985	0,429	0,844	0,183	0,485	0,374	0,733	0,173	0,402
13	8.975	0,574	0,992	1,344	0,990	1,119	0,693	0,772	0,858	1,960	0,710	0,575	0,884
14	2.720	1,112	0,373	1,847	1,101	1,059	0,770	0,309	0,713	0,993	2,062	0,170	0,647
15	3.601	1,097	0,630	0,989	0,859	1,436	0,893	1,012	1,037	1,809	1,052	1,079	1,187
16	5.000	0,658	0,949	0,310	0,619	1,335	1,629	0,990	1,318	2,234	1,717	3,169	2,130
17	2.875	0,987	1,002	1,159	1,049	0,264	0,869	0,438	0,493	0,268	0,885	0,876	0,441
18	8.528	0,791	1,415	0,376	0,860	0,607	0,034	0,199	0,547	0,767	0,590	0,530	0,635
31	6.701	1,219	1,500	0,704	1,141	0,181	0,503	0,331	0,305	0,149	0,989	0,470	0,180
32	17.825	0,862	1,062	0,678	0,834	0,425	0,664	0,191	0,423	0,493	0,616	0,330	0,508
33	14.830	1,150	0,543	1,136	0,943	0,461	0,081	0,209	0,244	0,401	0,113	0,184	0,259
34	2.688	0,608	1,340	0,857	0,833	0,113	0,361	0,265	0,246	0,224	0,269	0,403	0,295
35	30.167	1,069	0,623	0,860	0,924	0,193	0,111	0,077	0,117	0,150	0,135	0,090	0,127
51	2.831	1,203	0,880	1,009	1,060	0,125	0,100	0,131	0,148	0,104	0,216	0,123	0,141
52	5.618	0,932	0,708	1,154	0,931	1,270	0,737	0,720	0,942	1,372	1,042	0,624	0,978
53	2.602	1,137	0,669	0,866	0,950	0,884	1,301	1,843	1,344	0,778	2,287	2,130	1,997
61	3.965	0,997	0,672	0,704	0,791	0,851	0,461	0,848	0,587	0,653	0,686	0,921	0,742
62	7.351	0,955	1,373	1,470	1,267	0,591	0,552	0,611	0,548	0,617	0,402	0,279	0,409
63	10.350	0,610	1,078	0,987	0,884	0,832	0,991	1,285	1,035	1,355	0,921	1,329	1,272
64	2.960	0,995	1,148	0,904	1,016	0,887	0,430	0,669	0,582	0,891	0,374	0,740	0,652
71	9.095	0,619	0,505	0,852	0,658	0,280	0,513	0,434	0,400	0,455	1,015	0,510	0,822
72	4.412	1,291	0,847	0,807	1,015	1,757	1,277	0,889	1,306	1,361	1,348	1,101	1,266
73	5.978	1,625	1,145	1,013	1,261	1,938	0,983	1,035	1,315	1,187	0,859	1,022	1,043
74	2.267	0,987	1,033	0,602	0,874	0,706	0,604	0,604	0,632	0,715	0,585	0,970	0,722
81	2.520	1,459	0,829	1,139	1,143	0,477	0,000	0,000	0,159	0,327	0,000	0,000	0,138
82	1.309	0,685	0,768	0,882	0,779	1,028	0,926	0,140	0,867	2,231	1,206	0,186	1,113
Total	175.551	0,978	0,919	0,929	0,942	0,842	0,542	0,472	0,552	0,858	0,590	0,509	0,538
CV across provinces		0,885	0,291	0,316	0,341	0,167	0,850	0,611	0,782	0,578	0,711	0,738	0,980	0,849

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of establishments ISIC B

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found amply in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c6	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	164.136	1,054	0,964	0,892	0,970	1,176	1,083	0,916	1,059	1,118	1,134	1,028	1,091
12	428.004	0,936	0,938	1,030	1,004	0,996	1,005	0,483	0,830	1,012	1,007	0,474	0,826
13	297.756	1,052	0,901	0,977	0,976	1,009	0,891	0,947	0,963	1,017	0,990	0,969	0,992
14	141.518	0,910	1,131	0,958	1,011	0,907	0,945	0,625	0,826	0,898	0,836	0,630	0,817
15	88.675	1,257	1,028	1,087	1,126	1,284	1,058	1,125	1,195	1,022	1,030	1,038	1,029
16	279.520	1,071	1,053	0,915	1,017	1,050	1,078	0,859	1,028	0,860	1,012	1,047	1,012
17	49.849	0,951	1,160	1,090	1,067	1,023	1,202	1,197	1,141	1,076	1,038	1,097	1,069
18	278.381	0,950	0,811	0,967	0,909	0,935	0,813	0,952	0,900	0,884	1,002	0,984	0,880
31	500.947	0,959	0,880	0,976	0,930	0,969	0,883	0,927	0,929	1,010	1,003	0,950	0,987
32	1.090.692	1,074	1,034	0,902	1,003	1,077	1,045	0,921	1,014	1,003	1,011	1,020	1,011
33	1.559.539	0,982	1,039	0,993	1,006	1,050	1,092	1,005	1,049	1,070	1,051	1,013	1,045
34	178.428	0,898	1,076	0,941	0,972	0,911	1,168	1,033	1,037	1,015	1,084	1,093	1,057
35	1.900.275	0,954	0,940	0,971	0,955	0,984	0,909	1,004	0,986	1,031	1,031	1,034	1,032
61	153.548	0,931	1,135	0,937	1,001	0,970	1,195	0,940	1,038	1,043	1,053	1,003	1,034
62	168.324	1,005	0,997	1,029	1,010	1,054	0,866	1,027	1,013	1,049	0,959	0,998	1,002
63	74.065	1,095	0,850	1,114	1,030	1,211	0,988	1,225	1,141	1,106	1,120	1,099	1,108
64	128.292	1,145	1,027	0,908	1,027	1,288	1,100	0,754	1,057	1,125	1,072	0,853	1,030
72	78.711	0,837	1,281	0,869	1,029	0,993	1,288	0,928	1,063	1,059	0,990	1,069	1,033
73	165.520	0,884	0,898	1,033	0,934	0,813	0,957	0,974	0,881	0,976	0,918	0,944	0,943
74	92.514	1,115	1,086	0,862	1,054	1,182	1,088	1,020	1,080	1,015	1,002	1,061	1,024
71	134.894	1,077	0,895	0,917	0,963	1,073	0,074	0,849	0,832	0,997	0,978	0,925	0,986
72	91.123	0,861	1,141	0,853	1,025	1,018	1,212	0,999	1,075	1,033	1,062	1,049	1,049
73	381.672	1,084	1,099	0,857	0,987	1,078	1,107	0,973	1,045	1,042	1,110	1,018	1,080
74	60.533	1,056	0,882	0,911	0,950	1,056	0,882	0,938	0,973	1,000	1,001	1,052	1,026
81	75.884	0,918	0,937	1,088	0,984	0,254	0,000	0,000	0,085	0,277	0,000	0,000	0,088
82	38.856	0,983	0,960	1,257	1,087	1,064	0,883	0,435	0,794	1,052	0,919	0,346	0,744
Total	9.413.529	1,003	0,995	0,969	0,997	1,027	1,015	0,929	0,990	1,023	1,017	0,969	1,004
CV across provinces	1,484	0,088	0,106	- 0,090	0,045	0,183	0,237	0,282	0,207	0,150	0,210	0,283	0,201

ANNEX TABLE 8 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of establishments ISIC 7

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 (census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c7	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	22.772	0,752	1,170	0,720	0,881	0,373	1,596	0,809	1,079	1,161	1,330	1,123	1,225
12	66.156	1,044	1,120	1,063	1,076	1,046	1,339	0,521	1,002	1,002	1,195	0,585	0,931
13	22.221	0,806	0,817	1,134	0,946	1,120	0,895	1,256	1,024	1,263	0,050	1,109	1,083
14	22.767	1,106	1,299	1,214	1,206	1,476	0,674	0,948	1,024	1,337	0,519	0,757	0,849
15	15.409	1,262	0,806	1,156	1,087	1,588	0,927	1,142	1,219	1,258	1,151	0,555	1,121
16	45.584	1,321	1,025	0,323	1,056	1,740	1,309	1,144	1,400	1,323	1,277	1,369	1,326
17	6.873	0,777	1,023	1,166	0,989	0,998	1,170	1,367	1,158	1,206	1,144	1,172	1,171
18	43.588	1,027	0,838	0,818	0,894	0,661	0,809	0,560	0,817	0,955	0,966	0,806	0,913
31	72.077	0,030	1,068	1,065	0,990	0,814	0,998	0,905	0,906	0,972	0,935	0,850	0,915
32	507.157	1,143	1,038	1,042	1,076	1,157	1,110	1,106	1,123	1,043	1,070	1,061	1,043
33	243.377	0,092	0,889	0,548	0,943	0,919	0,948	1,108	0,992	1,031	0,958	1,360	1,052
34	20.337	0,969	0,911	1,124	1,001	0,914	1,043	1,101	1,019	0,943	1,145	0,979	1,016
35	377.552	1,030	0,847	1,013	0,953	1,108	0,908	1,099	1,037	1,076	1,072	1,081	1,077
51	16.380	0,629	1,198	0,934	0,920	0,536	0,952	0,628	0,705	0,853	0,795	0,872	0,766
52	33.191	1,116	0,756	1,240	1,041	1,064	0,887	1,259	0,984	0,963	0,819	1,010	0,945
53	8.518	1,010	0,962	0,767	0,910	1,260	1,563	0,864	1,209	1,267	1,841	1,257	1,394
61	15.263	0,935	0,832	1,127	0,965	0,892	0,745	0,925	0,854	0,954	0,896	0,821	0,885
62	15.739	1,150	1,041	0,892	1,031	1,516	1,199	1,179	1,239	1,309	1,152	1,322	1,260
63	37.347	0,811	1,233	0,867	1,001	0,908	0,807	1,184	0,985	1,113	0,727	1,237	0,994
64	16.030	0,879	0,856	1,078	0,937	0,839	1,134	1,266	1,077	0,954	1,386	1,388	1,149
71	29.766	0,724	0,811	0,866	0,830	0,750	0,944	1,057	0,917	1,037	1,163	1,106	1,105
72	8.191	0,820	1,461	0,940	1,073	1,065	2,029	1,365	1,493	1,288	1,389	1,453	1,332
73	66.177	0,903	1,000	1,109	1,007	0,803	0,992	1,198	1,028	0,869	0,984	1,080	1,021
74	5.737	1,514	0,861	0,825	1,000	1,481	0,904	1,486	1,294	0,985	1,049	2,377	1,293
81	3.511	0,635	1,288	1,080	1,001	0,373	0,000	0,000	0,058	0,272	0,000	0,000	0,058
82	4.100	0,797	0,764	0,840	0,801	1,028	0,897	0,577	0,791	1,289	1,173	0,211	0,876
Total	1.718.217	1,021	0,981	1,015	1,006	1,072	1,013	1,061	1,049	1,049	1,033	1,045	1,043
CV across provinces													
	1,805	0,218	0,183	0,158	0,088	0,316	0,361	0,342	0,264	0,202	0,299	0,411	0,247

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample and ISIC (details on Table 6)

Number of establishments ISIC 5

	full EC98 estimate	Relative estimates of EC98 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	05	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	5.443	0,636	0,855	0,548	0,680	0,282	0,361	0,272	0,298	0,412	0,410	0,497	0,434
12	22.089	0,990	0,940	1,074	1,001	0,638	0,459	0,137	0,424	0,644	0,530	0,127	0,424
13	18.985	0,921	0,808	0,994	0,938	0,410	0,640	0,674	0,575	0,446	0,713	0,676	0,613
14	14.808	0,828	1,436	1,557	1,274	0,646	0,223	1,303	0,724	0,779	0,155	0,837	0,583
15	5.541	1,177	0,705	0,837	0,906	0,494	0,347	0,519	0,620	0,419	1,201	0,620	0,654
16	23.329	1,182	0,965	0,853	1,000	0,523	0,602	0,464	0,581	0,447	0,624	0,544	0,532
17	5.215	0,559	1,170	1,036	0,927	0,609	1,029	1,277	0,972	1,076	0,873	1,233	1,048
18	13.892	0,587	0,527	0,587	0,607	0,210	0,263	0,537	0,270	0,358	0,689	0,914	0,652
31	109.821	0,810	0,982	0,957	0,918	0,872	1,067	0,872	0,967	1,200	1,077	0,911	1,055
32	143.394	0,980	0,872	0,853	0,895	0,546	0,513	0,590	0,550	0,589	0,588	0,892	0,614
33	70.998	0,748	1,082	0,945	0,924	0,438	0,761	0,426	0,540	0,584	0,704	0,449	0,585
34	25.895	1,046	1,050	0,886	0,966	0,526	0,767	0,735	0,706	0,598	0,720	0,828	0,709
35	184.164	1,104	1,004	1,092	1,074	0,755	0,816	0,567	0,646	0,649	0,614	0,539	0,602
51	18.852	1,034	1,450	1,167	1,217	0,582	0,674	0,409	0,608	0,559	0,465	0,402	0,500
52	5.712	0,580	0,838	0,892	0,722	0,562	0,298	0,464	0,410	0,814	0,507	0,520	0,610
53	5.264	0,638	0,828	0,862	0,841	0,493	0,804	0,809	0,702	0,590	0,973	0,938	0,834
61	4.786	1,397	0,861	1,352	1,037	0,798	0,407	0,305	0,503	0,570	1,127	0,226	0,485
62	2.625	1,097	1,084	1,566	1,249	0,449	0,862	0,802	0,781	0,409	0,916	0,576	0,626
63	7.792	0,370	1,346	1,815	1,177	0,148	0,390	0,262	0,268	0,383	0,289	0,144	0,226
64	14.385	0,998	1,253	0,752	1,000	0,756	0,808	0,614	0,769	0,759	0,724	0,816	0,769
71	8.758	0,893	0,664	0,713	0,756	0,144	0,378	0,318	0,279	0,180	0,577	0,443	0,370
72	3.667	1,805	0,583	1,059	1,052	0,887	0,664	0,573	0,778	0,621	1,140	0,635	0,719
73	14.282	0,954	0,852	1,268	1,025	0,268	0,275	0,404	0,315	0,278	0,322	0,318	0,307
74	1.459	1,582	0,339	0,432	0,784	1,586	0,522	0,395	0,901	1,065	2,429	0,452	1,149
81	4.140	0,282	1,058	0,949	0,929	0,213	0,000	0,000	0,071	0,272	0,000	0,000	0,076
82	2.489	0,721	0,884	1,303	1,003	0,564	0,329	0,338	0,477	1,337	0,394	0,108	0,475
Total	679.086	0,950	0,973	0,972	0,965	0,832	0,655	0,596	0,626	0,665	0,674	0,613	0,661
CV across provinces	1,498	0,306	0,316	0,311	0,175	0,564	0,461	0,580	0,390	0,400	0,682	0,526	0,297

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of establishments ISIC 9

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	cD	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	23.987	0,940	0,891	0,915	0,916	0,929	0,846	0,832	0,902	0,988	1,050	0,909	0,985
12	67.450	0,957	1,044	1,082	1,028	1,023	1,080	0,446	0,850	1,069	1,034	0,414	0,827
13	32.394	0,858	1,000	0,816	0,935	1,006	0,840	0,778	0,909	1,020	0,940	0,952	0,972
14	20.611	1,009	1,063	1,105	1,059	1,016	0,998	1,126	1,047	1,007	0,939	1,020	0,989
15	11.317	1,322	1,030	1,293	1,201	1,822	0,774	1,776	1,491	1,263	0,751	1,374	1,163
16	35.908	0,754	1,028	0,836	0,873	0,538	0,803	0,902	0,748	0,714	0,781	1,079	0,897
17	5.704	0,887	1,144	0,879	0,970	0,854	1,537	1,087	1,193	1,075	1,344	1,287	1,230
18	37.448	0,960	0,834	0,868	0,887	0,684	0,742	0,892	0,773	0,713	0,890	1,028	0,871
31	78.146	0,968	0,944	0,891	0,968	0,885	1,030	1,088	1,034	1,018	1,090	1,098	1,069
32	215.392	0,968	1,042	0,874	0,982	0,843	0,976	0,725	0,848	0,870	0,937	0,830	0,882
33	201.465	1,019	0,979	1,013	1,004	0,891	0,914	1,014	0,939	0,876	0,993	1,001	0,936
34	28.782	0,894	1,058	0,928	0,981	0,819	1,062	0,830	0,910	0,917	1,015	0,859	0,948
35	321.597	0,957	0,892	1,020	0,956	0,832	0,838	0,931	0,901	0,976	0,940	0,918	0,942
51	18.996	0,967	0,976	0,982	0,975	1,137	1,127	0,852	1,039	1,176	1,135	0,868	1,063
52	14.389	0,835	0,844	0,857	0,878	0,831	0,460	0,765	0,688	0,996	0,533	0,820	0,784
53	5.981	0,946	1,406	1,119	1,157	1,001	1,222	1,575	1,288	1,059	0,869	1,403	1,094
61	17.091	1,324	0,830	0,948	1,034	1,553	0,700	0,727	0,884	1,172	0,848	0,767	0,962
62	7.974	1,012	1,289	1,051	1,117	1,043	1,636	1,457	1,379	1,031	1,269	1,386	1,234
63	25.734	0,858	0,870	0,857	0,862	0,580	0,813	1,007	0,800	0,876	0,935	1,176	0,929
64	11.897	1,224	0,977	1,074	1,092	1,503	1,198	1,066	1,258	1,227	1,227	0,993	1,150
71	15.832	0,980	1,046	1,004	1,003	0,866	0,660	0,728	0,805	1,037	0,947	0,725	0,902
72	8.587	0,958	0,836	0,769	0,884	1,332	1,078	0,987	1,132	1,348	1,287	1,284	1,309
73	28.898	0,969	0,913	0,913	0,931	1,360	1,018	0,930	1,103	1,404	1,115	1,019	1,184
74	3.951	1,125	0,986	1,039	1,050	1,461	0,826	1,031	1,107	1,299	0,840	0,993	1,054
81	6.057	1,133	0,834	1,007	0,991	0,344	0,000	0,000	0,115	0,304	0,000	0,000	0,116
82	3.444	1,046	1,070	0,891	1,002	1,270	0,926	0,272	0,824	1,214	0,060	0,305	0,822
Total	1.250.690	0,976	0,965	0,970	0,970	0,932	0,923	0,859	0,915	0,956	0,958	0,917	0,943
CV across provinces	1,578	0,162	0,136	0,115	0,098	0,325	0,321	0,386	0,287	0,229	0,274	0,337	0,228

ANNEX TABLE 9 (cont)

SUSI98 Estimates by subsample - and ISIC (details on Table 6)

Number of workers ISIC 2

	full EC95 estimate	Relative estimates of EC95 variables, from SUSI 1999 census areas incl. found empty in SUSI				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	e2	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	4.797	1,333	0,905	0,449	0,898	0,360	0,213	0,259	0,277	0,270	0,235	0,578	0,309
12	3.368	0,917	0,552	1,208	0,793	1,067	1,006	0,128	0,734	1,729	1,821	0,106	0,928
13	4.049	0,810	0,471	1,258	0,848	0,173	0,505	0,763	0,460	0,213	1,023	0,607	0,668
14	2.314	1,777	1,271	0,498	1,101	1,382	0,301	0,206	0,530	0,777	0,237	0,419	0,539
15	5.158	1,473	0,584	1,150	1,089	0,169	0,280	1,480	0,848	0,115	0,480	1,295	0,605
16	6.167	0,517	1,080	0,443	0,600	0,217	1,010	1,988	1,071	0,419	0,935	4,489	1,575
17	1.243	1,485	2,583	0,244	1,473	0,078	0,313	0,121	0,171	0,053	0,116	0,497	0,110
18	4.844	2.052	0,696	1,242	1,330	0,355	0,404	0,089	0,483	0,465	0,581	0,072	0,383
31	87	0,613	3,105	0,000	1,240	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
32	25.217	0,910	0,628	0,838	0,000	1,054	0,597	0,970	0,874	1,158	0,951	1,096	1,081
33	64.918	1,409	1,171	0,602	1,081	1,579	0,914	0,287	0,928	1,120	0,780	0,477	0,873
34	10.448	1,391	0,865	0,706	0,987	1,518	1,060	0,643	1,037	1,091	1,214	0,768	1,050
35	69.160	0,629	0,834	0,586	0,797	0,858	0,894	0,924	0,825	1,360	0,842	0,987	1,035
51	6.527	1,239	0,920	0,109	0,756	0,419	0,260	0,183	0,285	0,338	0,272	1,719	0,378
52	12.465	0,551	1,131	1,208	0,998	0,752	2,104	1,024	1,293	1,368	1,690	0,783	1,298
53	4.150	0,043	1,477	1,557	1,026	2,985	1,271	0,123	1,452	69,526	0,851	0,082	1,425
61	16.437	0,887	0,700	0,906	0,831	1,481	0,589	1,129	1,068	1,670	0,641	1,247	1,264
62	22.617	0,646	1,037	0,922	0,869	1,171	1,417	1,659	1,415	1,812	1,365	1,798	1,830
63	10.880	1,039	1,251	0,220	0,840	1,308	1,218	4,420	2,313	1,254	0,972	19,176	2,753
64	2.245	1,279	0,804	0,073	0,719	3,243	3,658	0,251	3,054	4,099	4,550	3,675	4,249
71	6.268	0,574	0,441	1,062	0,685	1,062	0,350	0,223	0,545	1,851	0,793	0,212	0,781
72	2.748	0,000	0,887	0,109	0,330	0,073	0,404	0,078	0,212	#DIV/0!	0,545	0,756	0,841
73	2.740	1,360	1,490	0,976	1,288	3,395	2,957	2,418	2,923	2,429	1,985	2,478	2,270
74	1.925	0,569	1,556	1,838	1,388	1,522	1,649	0,683	1,251	1,570	1,060	0,356	0,902
81	3.259	0,984	0,838	0,508	0,777	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
82	2.151	1,287	2,601	0,371	1,380	0,803	1,097	0,275	1,025	0,693	0,799	0,741	0,743
Total	296.570	0,961	0,954	0,803	0,908	1,155	0,888	0,905	0,908	1,202	0,941	1,126	1,088
CV across provinces													
	1.509	0,187	0,575	0,830	0,276	1,021	0,908	1,252	0,805	#DIV/0!	0,917	2,143	0,856

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of workers — ISIC 3

	full EC98 estimate	Relative estimates of EC98 variables, from SUSI 1999 census areas incl. found empty in SUSI				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	146.716	0,948	0,938	0,887	0,924	1,089	1,052	0,877	1,008	1,148	1,124	0,950	1,088
12	137.286	0,954	1,098	0,920	0,991	1,267	1,255	0,484	0,886	1,327	1,144	0,805	1,005
13	121.275	1,107	1,027	1,114	1,083	1,170	1,527	1,365	1,361	1,057	1,486	1,344	1,257
14	51.476	0,931	1,096	1,049	1,025	0,985	1,162	0,897	1,016	1,058	1,060	0,896	0,990
15	42.545	0,830	0,869	0,923	0,814	0,680	1,073	0,739	0,833	1,080	1,213	0,800	1,023
16	124.123	1,299	0,889	0,921	1,070	1,569	1,189	0,832	1,230	1,284	1,203	0,903	1,160
17	22.653	0,812	0,864	0,926	0,867	1,244	0,952	0,771	0,969	1,531	1,101	0,833	1,140
18	157.854	1,177	1,206	0,853	1,082	1,213	1,536	1,003	1,249	1,031	1,273	1,158	1,165
31	165.040	1,074	0,947	0,893	0,971	0,863	1,276	1,209	1,116	0,803	1,347	1,355	1,149
32	882.336	1,040	0,893	0,975	0,969	1,007	1,091	1,109	1,069	0,969	1,221	1,137	1,103
33	1.540.332	1,006	0,980	0,991	0,993	1,143	1,150	1,058	1,117	1,134	1,174	1,088	1,125
34	164.906	1,006	1,012	0,933	1,004	1,176	1,151	1,158	1,172	1,160	1,107	1,166	1,167
35	1.149.622	0,951	0,986	0,977	0,972	1,182	1,223	1,158	1,180	1,221	1,237	1,184	1,214
51	173.053	0,851	0,850	0,034	0,851	1,164	1,214	0,846	1,108	1,226	1,416	1,136	1,296
52	146.869	0,891	1,309	0,843	1,014	1,086	1,257	0,788	1,037	1,222	0,962	0,908	1,023
53	126.925	0,830	1,116	0,847	0,997	1,067	1,164	1,326	1,184	1,147	1,044	1,395	1,187
61	77.714	1,095	0,862	1,002	0,998	1,851	1,125	0,998	1,244	1,470	1,306	0,998	1,262
62	35.095	1,115	1,214	0,761	1,030	1,723	1,201	1,171	1,365	1,545	0,959	1,838	1,325
63	118.378	1,300	1,048	1,009	1,119	1,358	0,939	1,069	1,121	1,044	0,898	1,059	1,002
64	27.036	0,851	0,887	1,002	0,846	1,068	0,883	1,502	1,185	1,124	1,109	1,600	1,232
71	128.244	0,884	0,978	1,104	0,990	0,995	1,011	1,168	1,057	1,121	1,034	1,068	1,069
72	58.769	1,012	1,012	0,851	0,992	0,947	0,936	1,116	0,999	0,906	0,925	1,172	1,008
73	192.923	0,900	0,981	1,152	1,011	1,133	1,092	1,267	1,164	1,260	1,114	1,100	1,152
74	41.020	1,071	0,814	1,068	0,984	0,860	1,128	0,883	0,970	0,764	1,355	0,873	0,886
81	36.081	1,096	1,243	1,085	1,141	0,446	0,000	0,000	0,149	0,407	0,000	0,000	0,150
82	21.454	1,165	0,658	1,080	0,931	0,542	0,407	0,487	0,632	0,816	0,838	0,451	0,679
Total	5.882.159	1,006	0,985	0,975	0,988	1,129	1,161	1,067	1,119	1,122	1,178	1,094	1,132
CV across provinces													
	1,006	0,139	0,164	0,097	0,073	0,243	0,270	0,219	0,223	0,215	0,246	0,219	0,212

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of workers ISIC 4

	full EC98 estimate	Relative estimates of EC98 variables, from SUSI 1999 census areas incl. (could empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	01	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	298	0,447	0,000	0,000	0,149	0,000	0,000	0,000	0,000	0,000			0,000
12	1,280	0,974	1,089	1,113	1,067	0,027	0,213	0,035	0,092	0,028	0,195	0,032	0,087
13	321	0,000	1,128	0,000	0,376	0,000	0,000	0,000	0,000		0,000		0,000
14	1,691	0,425	0,798	0,808	0,676	0,052	0,102	0,230	0,128	-0,123	0,128	0,286	0,190
16	496	0,772	0,630	0,957	0,753	0,251	0,145	0,107	0,166	0,325	0,273	0,112	0,223
16	2,177	0,885	0,937	0,431	0,877	0,078	0,094	0,037	0,055	-0,118	0,058	0,085	0,083
17	120	0,941	0,000	0,821	0,587	0,000	0,000	0,000	0,000	0,000		0,000	0,000
18	317	0,000	1,154	1,138	0,764	0,046	0,343	0,214	0,201		0,298	0,188	0,283
31	105	0,000	0,593	1,378	0,650	0,000	0,000	0,000	0,000		0,000	0,000	0,000
32	1,657	0,906	1,204	1,521	1,211	0,046	0,024	0,050	0,040	0,051	0,020	0,033	0,033
33	1,603	1,704	1,427	0,548	1,226	0,023	0,149	0,034	0,068	0,013	0,104	0,061	0,056
34	32	0,000	4,551	0,000	1,517	0,000	0,000	0,000	0,000		0,000		0,000
35	2,410	1,534	0,469	1,892	1,295	0,121	0,019	0,030	0,060	0,079	0,042	0,020	0,046
51	30	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000				
52	14	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000				
53	222	4,508	0,000	1,040	1,849	2,253	0,446	0,163	0,364	0,300		0,157	0,516
61	677	0,839	2,586	0,072	1,166	0,126	0,166	0,072	0,118	0,150	0,060	1,000	0,101
62	108	0,540	8,472	0,714	1,575	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
63	190	0,586	0,000	0,000	0,196	0,173	0,120	0,161	0,152	0,205			0,174
64	336	2,231	1,591	1,447	1,623	0,116	0,112	0,283	0,170	0,052	0,084	0,195	0,105
71	101	0,764	2,376	0,000	1,047	0,000	0,000	0,000	0,000	0,000	0,000		0,000
72	359	0,653	1,394	1,034	1,327	0,275	0,210	0,286	0,257	0,422	0,160	0,149	0,184
73	745	0,869	0,737	1,284	0,953	0,169	0,316	0,411	0,299	0,190	0,429	0,325	0,310
74	429	0,395	0,915	1,693	1,056	0,072	0,053	0,764	0,306	0,197	0,090	0,404	0,280
81	145	0,000	0,607	0,741	0,475	0,000	0,000	0,000	0,000		0,000	0,000	0,000
82	101	0,000	0,921	7,488	2,606	0,000	0,000	0,000	0,000		0,000	0,000	0,000
Total	15,953	0,979	0,985	1,052	1,009	0,110	0,105	0,113	0,109	0,112	0,107	0,106	0,108
CV across provinces	1,138	1,239	1,023	1,377	0,822	2,909	1,266	1,550	1,535	1,056	1,202	1,435	1,358

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of workers ISIC 5

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c5	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	18.991	1,463	0,863	1,617	1,246	0,854	0,349	0,983	0,729	0,583	0,405	0,895	0,534
12	16.279	0,890	1,302	1,142	1,095	0,539	0,723	0,179	0,588	0,243	0,710	0,157	0,518
13	18.509	0,693	1,104	1,237	1,005	1,464	1,481	1,333	1,428	2,487	1,251	1,078	1,419
14	5.333	0,965	0,842	1,805	1,061	1,353	1,250	0,351	0,918	1,195	2,042	0,219	0,868
15	7.272	1,340	0,404	1,286	1,030	2,976	0,667	1,144	1,596	2,221	1,377	0,904	1,549
16	11.341	0,907	1,031	0,342	0,760	1,521	2,208	1,465	1,765	1,787	2,141	4,287	2,322
17	5.171	0,775	1,176	1,140	1,033	0,246	0,752	0,506	0,505	0,318	0,648	0,441	0,439
18	15.245	0,809	1,475	0,601	0,828	1,180	1,038	0,655	0,958	1,539	0,704	1,534	1,167
31	16.151	0,941	1,284	0,706	0,860	0,773	0,384	0,443	0,533	0,821	0,297	0,827	0,544
32	11.543	1,054	0,943	0,575	0,857	0,501	0,560	0,214	0,425	0,475	0,594	0,372	0,496
33	38.677	0,963	0,988	1,236	0,828	0,522	0,047	0,182	0,243	0,542	0,080	0,131	0,282
34	10.485	0,955	1,821	0,588	0,988	0,106	0,558	0,443	0,367	0,111	0,418	0,046	0,372
35	61.063	1,002	0,718	1,061	0,828	0,351	0,191	0,140	0,227	0,350	0,288	0,132	0,245
51	9.940	1,228	0,728	0,990	0,982	0,071	0,107	0,085	0,088	0,068	0,147	0,006	0,090
52	17.635	0,922	0,919	0,823	0,888	1,390	0,728	1,066	1,227	2,060	0,790	1,295	1,282
53	8.061	0,580	0,976	1,108	0,924	0,968	0,951	1,342	1,087	1,404	0,974	1,211	1,176
61	7.933	1,154	0,589	0,738	0,831	1,050	1,042	0,927	1,006	0,940	1,739	1,254	1,211
62	12.446	0,975	1,219	1,484	1,243	0,851	0,380	0,588	0,840	0,873	0,743	0,481	0,676
63	24.442	0,807	1,069	0,977	0,881	0,307	0,952	1,436	1,098	1,330	0,891	1,472	1,205
64	7.188	1,500	0,986	1,190	1,228	1,063	0,535	0,584	0,761	0,705	0,643	0,576	0,619
71	20.216	0,537	0,830	0,964	0,710	0,355	0,832	0,912	0,699	0,661	1,321	0,948	0,985
72	11.550	1,121	0,898	0,794	0,937	2,069	1,900	0,821	1,596	1,845	2,116	1,035	1,703
73	16.806	1,897	1,281	0,843	1,324	2,434	1,102	1,237	1,591	1,283	0,874	1,523	1,202
74	5.855	0,938	1,007	0,953	0,966	0,874	0,680	0,382	0,815	0,581	0,584	0,920	0,843
81	6.561	1,400	0,879	1,160	1,183	0,533	0,000	0,000	0,178	0,358	0,000	0,000	0,150
82	3.463	0,392	0,637	1,193	0,741	1,288	0,845	0,102	0,652	3,157	1,013	0,085	0,893
Total	418.795	0,984	0,927	0,874	0,961	0,879	0,603	0,830	0,731	0,894	0,737	0,847	0,760
CV across provinces	0,796	0,340	0,279	0,306	0,160	0,682	0,843	0,652	0,558	0,692	0,583	0,584	0,595

ANNEX TABLE 8 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of workers ISIC 6

	full EC98 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. (and empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c6	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	288.183	1,087	0,945	0,920	0,984	1,482	1,202	1,074	1,236	1,347	1,271	1,168	1,266
12	692.926	0,991	0,982	1,091	1,002	1,200	1,197	0,818	1,005	1,210	1,219	0,599	1,003
13	937.443	1,016	0,990	0,986	0,937	1,224	1,058	1,069	1,160	1,303	1,231	1,142	1,227
14	247.746	0,947	1,072	1,021	1,014	1,172	1,107	0,897	1,039	1,236	1,033	0,876	1,045
15	145.833	1,202	1,091	1,004	1,129	1,492	1,358	1,362	1,404	1,241	1,255	1,245	1,247
16	440.071	1,016	1,041	0,963	1,007	1,264	1,463	1,233	1,327	1,244	1,423	1,280	1,316
17	77.656	0,951	1,189	1,083	1,071	1,190	1,358	1,448	1,330	1,201	1,142	1,395	1,242
10	414.482	0,969	0,845	1,031	0,945	1,149	0,904	1,118	1,057	1,198	1,069	1,084	1,115
31	853.221	0,990	0,915	0,960	0,955	1,095	0,958	1,068	1,043	1,105	1,058	1,078	1,081
32	2.651.790	1,088	0,985	0,894	0,989	1,302	1,210	1,051	1,187	1,198	1,228	1,175	1,201
33	2.219.942	0,990	1,015	0,973	0,993	1,184	1,315	1,218	1,239	1,155	1,295	1,201	1,248
34	279.798	0,928	1,032	0,978	0,978	1,054	1,134	1,125	1,104	1,139	1,099	1,151	1,129
35	2.684.602	0,968	0,876	0,996	0,980	1,196	1,253	1,244	1,231	1,226	1,284	1,260	1,257
51	249.772	0,969	1,086	0,970	1,009	1,262	1,276	1,107	1,215	1,302	1,172	1,142	1,204
52	248.724	0,867	0,838	1,043	0,879	1,083	1,006	1,221	1,103	1,131	1,073	1,170	1,126
53	125.355	1,138	0,886	1,093	1,033	1,383	1,247	1,512	1,381	1,215	1,437	1,383	1,338
61	223.005	1,102	1,057	0,859	1,049	1,492	1,474	0,866	1,278	1,354	1,394	0,664	1,253
62	138.589	0,905	1,214	0,990	1,038	1,145	1,343	0,984	1,157	1,268	1,105	0,994	1,117
63	247.258	0,859	0,909	1,050	0,939	1,144	0,847	1,242	1,111	1,332	1,041	1,183	1,182
64	154.707	1,222	1,037	0,887	1,045	1,596	1,483	1,274	1,451	1,306	1,445	1,438	1,389
71	216.787	1,041	0,806	0,910	0,978	1,126	0,806	0,879	0,897	1,081	1,001	0,666	1,019
72	138.554	0,907	1,171	0,928	1,002	1,059	1,306	1,213	1,183	1,167	1,315	1,310	1,191
73	524.727	0,998	1,076	0,887	0,987	1,349	1,481	1,111	1,317	1,351	1,385	1,253	1,334
74	95.231	0,939	1,045	0,922	0,969	1,158	1,144	1,236	1,179	1,233	1,094	1,341	1,217
81	187.531	0,902	0,870	1,199	0,957	0,869	0,000	0,000	0,089	0,334	0,000	0,000	0,093
82	71.844	0,930	0,987	1,172	1,030	1,193	1,023	0,463	0,893	1,253	1,036	0,395	0,867
Total	14.103,882	1,008	0,989	0,967	0,988	1,220	1,216	1,109	1,179	1,211	1,224	1,144	1,193
CV across provinces													
	1,411	0,098	0,097	0,052	0,041	0,184	0,249	0,291	0,218	0,156	0,232	0,290	0,209

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of workers — ISIC 7

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas (not found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c7	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	31,872	0,830	1,097	0,791	0,909	0,814	1,308	0,804	0,976	0,580	1,193	1,018	1,078
12	64,025	1,014	1,055	1,045	1,051	0,940	1,209	0,593	0,917	0,935	1,104	0,667	0,872
13	32,784	0,934	0,827	1,271	1,011	1,112	0,727	1,310	1,049	1,190	0,879	1,030	1,033
14	20,403	1,104	1,350	1,133	1,196	1,348	0,848	1,077	1,091	1,221	0,628	0,855	0,912
15	20,085	1,213	0,866	1,308	1,129	1,463	1,304	1,067	1,255	1,222	1,508	0,818	1,133
16	58,098	1,261	1,000	0,848	1,070	1,729	1,314	1,185	1,409	1,371	1,314	1,200	1,317
17	8,301	0,753	1,056	1,134	0,981	0,934	1,052	1,342	1,109	1,240	0,997	1,183	1,131
18	57,489	0,970	0,874	1,023	0,958	0,952	0,858	0,783	0,858	0,973	0,851	0,746	0,895
31	84,334	0,843	1,121	1,059	1,000	0,826	0,924	0,935	0,895	1,008	0,824	0,883	0,893
32	660,769	1,137	1,043	1,028	1,069	1,170	1,117	1,113	1,133	1,028	1,071	1,082	1,060
33	306,551	0,945	0,976	0,920	0,947	1,003	0,927	1,120	1,017	1,062	0,950	1,218	1,074
34	26,039	1,029	1,021	1,007	1,046	1,041	1,330	1,207	1,193	1,012	1,303	1,310	1,141
35	460,940	1,168	0,803	0,983	0,977	1,110	0,890	1,064	1,021	0,962	1,108	1,105	1,043
51	20,726	0,807	1,172	0,800	0,956	0,591	0,977	0,878	0,815	0,732	0,834	0,884	0,826
52	46,072	1,033	0,782	1,132	0,952	1,088	0,825	1,238	0,954	1,052	0,799	1,095	1,001
53	12,919	1,030	0,806	0,767	0,901	0,844	1,405	0,790	1,046	0,947	1,560	1,000	1,161
61	21,510	1,089	0,917	1,110	1,042	1,065	0,813	0,908	0,928	0,978	0,886	0,811	0,891
62	20,968	1,182	1,058	0,999	1,072	1,311	1,246	1,090	1,218	1,128	1,179	1,081	1,134
63	42,787	0,785	1,242	0,935	0,987	0,997	0,907	1,161	1,021	1,270	0,780	1,241	1,034
64	19,563	0,918	0,893	1,146	0,985	0,864	1,100	1,238	1,034	0,941	1,232	1,126	1,101
71	80,293	0,872	0,886	1,010	0,913	0,888	0,948	0,972	0,925	0,893	1,107	0,962	1,017
72	12,943	0,899	1,299	1,037	0,995	1,310	1,922	1,378	1,527	1,378	1,528	1,328	1,539
73	79,518	0,982	0,975	1,109	1,020	0,965	1,148	1,276	1,129	0,583	1,177	1,157	1,108
74	10,312	1,345	0,841	0,007	0,998	1,297	0,890	1,068	1,085	0,985	1,035	1,323	1,055
81	9,901	0,858	1,106	1,059	1,028	0,122	0,000	0,000	0,041	0,142	0,000	0,000	0,040
82	7,099	0,720	0,884	0,810	0,806	1,028	0,853	0,188	0,889	1,426	0,935	0,231	0,854
Total	2,111,960	1,057	0,971	1,004	1,011	1,006	1,012	1,064	1,054	1,027	1,042	1,059	1,042
CV across provinces													
	1,723	0,173	0,148	0,129	0,076	0,285	0,328	0,328	0,258	0,266	0,302	0,344	0,242

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of workers ISIC 3

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	q3	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	7.536	0,630	0,865	0,571	0,629	0,238	0,291	0,233	0,254	0,378	0,425	0,409	0,404
12	31,389	1,030	0,872	1,059	0,991	0,751	0,485	0,208	0,461	0,729	0,550	0,195	0,488
13	25,453	1,025	0,977	0,862	0,894	0,447	0,694	0,598	0,580	0,439	0,710	0,610	0,583
14	17,960	0,844	1,318	1,404	1,218	0,659	0,264	1,300	0,744	0,792	0,291	0,870	0,811
15	7,282	1,162	0,861	0,843	0,952	0,423	0,997	0,519	0,647	0,388	1,158	0,616	0,679
16	30,601	1,058	0,931	1,071	1,019	0,853	0,537	0,636	0,609	0,619	0,577	0,534	0,597
17	6,573	0,610	1,060	1,055	0,883	0,802	1,160	1,517	1,190	1,571	1,047	1,333	1,240
18	20,069	0,585	0,623	0,824	0,604	0,168	0,379	0,254	0,440	0,332	0,508	1,209	0,728
31	128,799	0,020	1,004	0,829	0,918	0,917	0,987	0,843	0,916	1,119	0,983	0,907	0,986
32	168,830	0,954	0,900	0,950	0,921	0,841	0,534	0,789	0,655	0,702	0,693	0,831	0,711
33	125,903	0,861	1,076	0,971	0,987	0,404	0,825	0,548	0,595	0,449	0,774	0,563	0,609
34	38,000	1,223	1,125	0,912	1,087	0,882	0,752	0,746	0,730	0,641	0,656	0,819	0,672
35	203,346	1,194	1,006	1,029	1,077	0,041	0,634	0,598	0,691	0,705	0,630	0,579	0,641
51	22,055	1,008	1,282	1,168	1,139	0,789	0,632	0,510	0,643	0,779	0,613	0,437	0,585
52	9,864	0,802	0,882	1,140	0,775	0,615	0,322	0,501	0,518	1,021	0,553	0,527	0,882
53	8,285	0,879	0,787	0,815	0,810	0,789	0,842	0,897	0,829	0,897	1,141	1,052	1,023
61	8,735	1,372	0,321	1,138	0,944	1,342	0,584	0,222	0,716	0,978	1,820	0,195	0,759
62	3,931	0,939	1,242	1,505	1,229	0,415	0,939	0,982	0,779	0,442	0,756	0,652	0,634
63	10,193	0,354	1,081	1,785	1,076	0,182	0,451	0,532	0,432	0,535	0,446	0,352	0,401
64	18,051	1,047	1,210	0,701	0,976	0,789	1,098	0,872	0,851	0,770	0,807	0,858	0,872
71	13,284	0,994	1,014	0,604	0,837	0,598	0,423	0,352	0,324	0,200	0,417	0,438	0,348
72	5,502	1,753	0,563	1,005	1,094	1,114	0,754	0,446	0,752	0,850	1,392	0,460	0,715
73	22,571	1,039	1,098	1,462	1,193	0,305	0,348	0,801	0,375	0,370	0,317	0,271	0,344
74	2,259	1,423	0,547	0,458	0,800	2,889	0,844	0,306	1,273	1,876	1,642	0,572	1,673
81	6,025	0,840	1,005	0,820	0,059	0,344	0,000	0,000	0,115	0,537	0,000	0,000	0,134
82	3,841	0,866	0,928	1,257	0,947	0,966	0,511	0,116	0,531	1,474	0,651	0,092	0,581
Total	963,029	0,977	0,934	0,997	0,995	0,578	0,873	0,661	0,670	0,694	0,654	0,663	0,680
CV across provinces													
	1,498	0,348	0,284	0,289	0,161	0,896	0,444	0,564	0,394	0,530	0,547	0,566	0,488

ANNEX TABLE 9 (cont)

SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of workers ISIC 9

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c9	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	38.991	0,919	1,151	0,999	1,023	1,058	1,287	0,889	1,071	1,148	1,101	0,890	1,047
12	114.314	0,951	0,974	1,059	0,988	1,058	1,233	0,495	0,929	1,113	1,268	0,463	0,931
13	50.091	1,073	0,934	0,933	0,985	1,039	1,096	0,981	1,039	0,965	1,103	1,105	1,054
14	37.610	0,965	1,051	0,964	0,994	0,943	1,447	1,281	1,217	0,978	1,377	1,308	1,225
15	18,064	1,498	1,219	1,223	1,281	2,276	0,994	2,094	1,760	1,618	0,758	1,712	1,380
16	57.318	0,736	0,835	0,537	0,835	0,816	0,565	0,989	0,923	1,109	1,094	1,178	1,104
17	9,781	0,879	1,094	1,016	0,995	1,138	1,389	1,036	1,101	1,300	1,251	1,020	1,187
18	53.760	0,901	0,860	0,888	0,885	0,513	0,924	0,998	0,845	0,690	1,064	1,124	0,954
31	178,020	0,980	0,997	1,058	0,991	1,023	0,913	1,101	1,012	1,044	0,974	1,043	1,022
32	351.730	0,998	1,040	0,839	0,969	0,877	1,182	0,799	0,939	0,879	1,138	0,904	0,980
33	309,821	1,018	0,982	1,016	0,999	0,922	0,957	1,034	0,971	0,935	0,895	1,017	0,972
34	46.337	0,911	1,010	0,870	0,920	0,853	1,251	0,797	0,970	0,948	1,238	0,917	1,043
35	478.131	0,991	0,890	1,014	0,961	1,142	1,008	1,108	1,086	1,164	1,133	1,093	1,130
51	37.910	0,916	1,175	1,068	1,053	1,143	1,218	0,957	1,106	1,248	1,037	0,898	1,091
52	23.784	0,817	0,974	0,910	0,901	1,240	0,456	0,703	0,800	1,517	0,460	0,772	0,888
53	12.658	1,406	1,310	1,037	1,251	1,260	1,101	1,807	1,380	0,896	0,840	1,742	1,110
61	28.334	1,409	0,902	0,977	1,056	1,950	0,818	0,844	1,137	1,383	0,906	0,659	1,037
62	13.551	0,940	1,289	1,293	1,174	1,246	1,672	1,234	1,354	1,326	1,219	0,954	1,151
63	38.276	0,825	0,911	0,960	0,899	0,409	0,764	1,063	0,779	0,605	0,830	1,107	0,803
64	20.792	1,310	0,999	0,990	1,100	1,093	1,423	1,029	1,452	1,445	1,435	1,040	1,320
71	24.643	0,924	0,851	1,039	0,971	0,950	1,245	0,964	1,053	1,028	1,309	0,928	1,084
72	14.370	0,917	0,747	0,812	0,825	1,616	1,055	1,471	1,384	1,763	1,427	1,811	1,677
73	49.444	1,129	0,893	0,910	0,977	1,455	1,188	0,922	1,288	1,289	1,663	1,014	1,318
74	6.538	1,094	0,849	0,901	0,946	1,493	0,807	1,164	1,151	1,358	0,950	1,292	1,214
81	10.191	0,934	0,741	0,989	0,888	0,291	0,000	0,000	0,097	0,312	0,000	0,000	0,108
82	6.707	1,146	1,047	0,836	1,017	1,289	1,244	0,248	0,960	1,124	1,283	0,290	0,945
Total	2.027.405	0,990	0,989	0,972	0,976	1,035	1,062	0,969	1,022	1,045	1,009	0,997	1,047
CV across provinces													
	1,499	0,177	0,142	0,113	0,110	0,373	0,310	0,416	0,275	0,278	0,302	0,387	0,241

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample – and ISIC (details on Table 6)

Number of establishments — ISIC 5

	full EC96 estimate	Relative estimates of EC96 variables, from SUSI 1999 census areas incl. found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC96 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c5	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	7.815	1,419	0,649	1,239	1,102	0,277	0,366	0,792	0,545	0,547	0,564	0,639	0,556
12	6.451	0,748	1,150	1,058	0,985	0,429	0,844	0,383	0,485	0,574	0,733	0,473	0,492
13	0.975	0,374	0,862	1,344	0,960	1,119	0,653	0,772	0,858	1,560	0,710	0,575	0,894
14	2.730	1,112	0,373	1,817	1,101	1,059	0,770	0,309	0,713	0,963	2,062	0,170	0,647
15	3.601	1,097	0,530	0,838	0,889	1,436	0,663	1,012	1,037	1,309	1,052	1,078	1,167
16	5.000	0,598	0,949	0,310	0,810	1,335	1,899	0,990	1,318	2,234	1,717	3,169	2,130
17	2.075	0,807	1,002	1,168	1,049	0,264	0,669	0,430	0,453	0,258	0,655	0,375	0,441
18	6.523	0,791	1,415	0,375	0,880	0,607	0,834	0,199	0,547	0,767	0,590	0,530	0,635
31	5.701	1,218	1,500	0,704	1,161	0,161	0,103	0,331	0,205	0,149	0,059	0,470	0,180
32	17.325	0,862	1,062	0,578	0,834	0,425	0,854	0,191	0,423	0,493	0,616	0,330	0,508
33	14.600	1,150	0,543	1,135	0,943	0,461	0,061	0,208	0,244	0,401	0,113	0,184	0,259
34	2.833	0,503	1,340	0,667	0,833	0,113	0,391	0,266	0,246	0,224	0,269	0,403	0,295
35	30.167	1,069	0,623	0,880	0,924	0,163	0,111	0,077	0,117	0,160	0,135	0,090	0,127
51	2.831	1,203	0,830	1,060	1,050	0,125	0,188	0,131	0,148	0,104	0,214	0,123	0,141
52	3.848	0,802	0,708	1,154	0,931	1,279	0,737	0,720	0,912	1,372	1,042	0,524	0,979
53	2.502	1,187	0,569	0,855	0,858	0,834	1,301	1,845	1,344	0,778	2,287	2,130	1,567
61	3.965	0,997	0,672	0,704	0,791	0,651	0,461	0,548	0,587	0,653	0,686	0,821	0,742
62	7.886	0,958	1,378	1,470	1,267	0,691	0,552	0,411	0,518	0,617	0,402	0,279	0,409
63	10.050	0,810	1,076	0,967	0,864	0,832	0,991	1,285	1,036	1,365	0,921	1,329	1,172
64	2.560	0,995	1,149	0,904	1,015	0,837	0,430	0,669	0,582	0,891	0,374	0,740	0,852
71	8.095	0,616	0,505	0,852	0,858	0,280	0,513	0,434	0,409	0,455	1,016	0,510	0,622
72	4.412	1,291	0,947	0,807	1,015	1,767	1,277	0,859	1,308	1,391	1,348	1,101	1,266
73	5.978	1,625	1,145	1,013	1,281	1,928	0,863	1,035	1,315	1,187	0,869	1,022	1,043
74	2.267	0,987	1,033	0,602	0,874	0,706	0,504	0,584	0,632	0,715	0,585	0,970	0,722
81	2.620	1,459	0,828	1,139	1,143	0,477	0,000	0,000	0,159	0,327	0,000	0,000	0,139
82	1.206	0,885	0,798	0,882	0,779	1,523	0,928	0,146	0,887	2,231	1,205	0,168	1,313
Total	176.661	0,976	0,919	0,829	0,842	0,642	0,542	0,472	0,552	0,595	0,590	0,509	0,585
CV across provinces													
	0,886	0,291	0,316	0,341	0,167	0,650	0,611	0,762	0,576	0,711	0,738	0,990	0,849

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample - and ISIC (details on Table 6)

Number of establishments ISIC 6

Prov	full EC98 estimate	Relative estimates of EC98 variables, from SUSI 1999 (census areas incl. found empty in SUSI) Subsample 1	2	3	1-3	Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion) Subsample 1	2	3	1-3	Ratio estimation Subsample 1	2	3	1-3
11	184.186	1,054	0,964	0,892	0,970	1,178	1,083	0,916	1,069	1,116	1,124	1,026	1,091
12	428.004	0,985	0,998	1,030	1,004	0,996	1,005	0,488	0,930	1,012	1,007	0,474	0,926
13	207.758	1,052	0,901	0,977	0,976	1,069	0,891	0,947	0,969	1,017	0,990	0,969	0,982
14	141.040	0,910	1,131	0,893	1,061	0,907	0,945	0,825	0,926	0,998	0,936	0,830	0,917
15	88.675	1,257	1,028	1,087	1,124	1,284	1,058	1,125	1,158	1,022	1,030	1,035	1,029
16	279.626	1,071	1,063	0,915	1,017	1,060	1,078	0,959	1,028	0,980	1,012	1,047	1,012
17	49.649	0,951	1,160	1,090	1,067	1,023	1,202	1,197	1,141	1,076	1,038	1,097	1,069
18	278.301	0,950	0,811	0,967	0,909	0,935	0,813	0,962	0,900	0,984	1,002	0,984	0,990
31	500.947	0,969	0,830	0,978	0,938	0,953	0,883	0,927	0,926	1,010	1,003	0,960	0,987
32	1.560.592	1,074	1,034	0,902	1,003	1,077	1,045	0,921	1,014	1,003	1,011	1,020	1,011
33	1.559.559	0,982	1,039	0,993	1,004	1,050	1,092	1,005	1,049	1,070	1,051	1,013	1,045
34	178.628	0,888	1,076	0,961	0,972	0,911	1,166	1,093	1,037	1,015	1,084	1,098	1,067
35	1.500.275	0,954	0,940	0,971	0,955	0,984	0,989	1,004	0,998	1,031	1,031	1,034	1,032
51	153.549	0,831	1,135	0,937	1,001	0,970	1,195	0,940	1,035	1,043	1,003	1,009	1,004
52	168.324	1,005	0,997	1,029	1,010	1,054	0,956	1,027	1,013	1,049	0,969	0,998	1,002
53	74.695	1,095	0,880	1,114	1,030	1,211	0,866	1,225	1,141	1,108	1,120	1,099	1,108
61	128.292	1,145	1,027	0,908	1,027	1,258	1,100	0,784	1,057	1,125	1,072	0,863	1,030
62	79.711	0,937	1,281	0,869	1,028	0,993	1,208	0,988	1,063	1,059	0,990	1,059	1,033
63	165.620	0,834	0,936	1,038	0,934	0,813	0,857	0,974	0,881	0,975	0,918	0,944	0,943
64	92.514	1,115	1,093	0,962	1,054	1,192	1,088	1,020	1,060	1,015	1,002	1,051	1,024
71	134.354	1,077	0,895	0,917	0,969	1,073	0,874	0,849	0,932	0,997	0,978	0,925	0,958
72	91.123	0,851	1,141	0,953	1,020	1,013	1,212	0,909	1,076	1,033	1,062	1,049	1,049
73	361.072	1,034	1,069	0,867	0,987	1,078	1,187	0,873	1,048	1,042	1,110	1,018	1,050
74	60.533	1,056	0,832	0,911	0,850	1,056	0,882	0,926	0,976	1,000	1,001	1,082	1,026
81	76.884	0,918	0,937	1,098	0,984	0,254	0,000	0,000	0,085	0,277	0,000	0,000	0,085
82	38.896	0,583	0,960	1,257	1,067	1,064	0,883	0,438	0,794	1,082	0,919	0,348	0,744
Total	9.418.579	1,003	0,998	0,959	0,987	1,027	1,015	0,929	0,990	1,023	1,017	0,958	1,004
CV across provinces	1,494	0,090	0,106	0,090	0,046	0,183	0,237	0,282	0,207	0,150	0,210	0,283	0,201

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of establishments ISIC 8

	full EC98 estimate	Relative estimates of EC98 variables, from SUSI 1999 census gross incl. (found empty in SUSI)				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c8	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	5.413	0.838	0.855	0.546	0.680	0.252	0.351	0.272	0.295	0.412	0.410	0.497	0.434
12	22.089	0.990	0.940	1.074	1.001	0.533	0.499	0.137	0.426	0.544	0.530	0.127	0.424
13	18.985	0.921	0.895	0.894	0.839	0.410	0.540	0.674	0.575	0.449	0.713	0.676	0.613
14	14.808	0.823	1.436	1.557	1.274	0.646	0.223	1.303	0.726	0.779	0.165	0.837	0.568
15	8.541	1.177	0.705	0.837	0.906	0.494	0.847	0.519	0.820	0.419	1.201	0.620	0.604
16	23.829	1.182	0.965	0.853	1.000	0.529	0.802	0.464	0.531	0.447	0.534	0.544	0.532
17	5.215	0.588	1.179	1.006	0.827	0.609	1.029	1.277	0.872	1.078	0.873	1.233	1.048
18	13.892	0.587	0.527	0.587	0.567	0.210	0.383	0.537	0.370	0.368	0.580	0.914	0.852
31	101.621	0.810	0.862	0.897	0.816	0.972	1.057	0.872	0.987	1.209	1.077	0.911	1.055
32	143.694	0.960	0.872	0.853	0.895	0.546	0.513	0.560	0.550	0.669	0.588	0.652	0.614
33	70.996	0.746	1.082	0.945	0.824	0.436	0.701	0.424	0.540	0.534	0.704	0.449	0.565
34	25.858	1.046	1.050	0.888	0.995	0.626	0.797	0.735	0.706	0.696	0.720	0.828	0.709
35	134.164	1.164	1.004	1.052	1.074	0.765	0.510	0.567	0.848	0.849	0.854	0.539	0.602
51	16.862	1.084	1.450	1.167	1.217	0.882	0.874	0.469	0.608	0.659	0.455	0.402	0.500
52	6.712	0.689	0.583	0.892	0.722	0.562	0.286	0.464	0.440	0.214	0.507	0.520	0.610
53	5.284	0.835	0.826	0.862	0.841	0.493	0.804	0.809	0.702	0.590	0.873	0.888	0.824
61	4.736	1.397	0.361	1.352	1.037	0.795	0.407	0.305	0.503	0.570	1.127	0.228	0.485
62	2.825	1.067	1.084	1.565	1.249	0.449	0.992	0.902	0.781	0.409	0.816	0.676	0.625
63	7.792	0.370	1.346	1.315	1.177	0.146	0.390	0.252	0.258	0.293	0.289	0.144	0.226
64	14.365	0.866	1.253	0.762	1.000	0.758	0.908	0.814	0.759	0.759	0.724	0.816	0.759
71	8.766	0.899	0.851	0.713	0.756	0.144	0.378	0.315	0.279	0.160	0.577	0.443	0.370
72	3.557	1.005	0.883	1.059	1.082	0.997	0.884	0.673	0.778	0.621	1.140	0.535	0.719
73	14.282	0.954	0.852	1.288	1.025	0.295	0.275	0.404	0.315	0.279	0.322	0.318	0.307
74	1.458	1.582	0.339	0.482	0.784	1.666	0.822	0.195	0.901	1.066	2.429	0.452	1.149
81	4.140	0.782	1.058	0.948	0.929	0.213	0.000	0.000	0.071	0.272	0.000	0.000	0.076
82	2.480	0.721	0.984	1.303	1.003	0.964	0.329	0.138	0.477	1.337	0.334	0.106	0.475
Total	679.028	0.960	0.973	0.972	0.965	0.632	0.655	0.596	0.628	0.685	0.674	0.813	0.661
CV across provinces	1.480	0.306	0.316	0.211	0.175	0.554	0.481	0.586	0.390	0.480	0.632	0.526	0.397

ANNEX TABLE 9 (cont)
SUSI99 Estimates by subsample - and ISIC (details on Table 8)

Number of establishments ISIC 8

	full EC98 estimate	Relative estimates of EC98 variables, from SUSI 1999 census gross incl. found empty in SUSI				Estimates from SUSI 1999 (3 quarters), compared to EC98 Linear (simple expansion)				Ratio estimation			
		Subsample				Subsample				Subsample			
Prov	c8	1	2	3	1-3	1	2	3	1-3	1	2	3	1-3
11	5.413	0,838	0,835	0,846	0,880	0,252	0,351	0,272	0,295	0,412	0,410	0,497	0,434
12	22.089	0,990	0,940	1,074	1,001	0,933	0,499	0,137	0,426	0,544	0,530	0,127	0,424
13	18.935	0,921	0,896	0,894	0,939	0,410	0,540	0,674	0,575	0,449	0,713	0,676	0,613
14	14.808	0,828	1,436	1,557	1,274	0,846	0,223	1,303	0,726	0,779	0,155	0,837	0,568
15	8.541	1,177	0,705	0,837	0,906	0,494	0,847	0,519	0,820	0,419	1,201	0,620	0,684
16	23.829	1,182	0,965	0,853	1,000	0,529	0,802	0,464	0,531	0,447	0,524	0,544	0,532
17	5.215	0,866	1,179	1,035	0,927	0,609	1,029	1,277	0,972	1,078	0,873	1,233	1,048
18	13.892	0,687	0,527	0,587	0,587	0,210	0,383	0,537	0,370	0,368	0,580	0,914	0,852
31	101.621	0,810	0,962	0,867	0,816	0,972	1,057	0,872	0,987	1,200	1,077	0,911	1,055
32	143.694	0,960	0,872	0,853	0,895	0,546	0,513	0,590	0,560	0,669	0,588	0,652	0,614
33	70.996	0,748	1,082	0,945	0,924	0,436	0,761	0,424	0,540	0,534	0,704	0,449	0,585
34	25.858	1,046	1,050	0,883	0,995	0,826	0,797	0,735	0,706	0,690	0,720	0,828	0,709
35	134.164	1,164	1,004	1,052	1,074	0,765	0,616	0,667	0,848	0,849	0,814	0,539	0,602
51	16.862	1,084	1,450	1,167	1,217	0,882	0,874	0,469	0,608	0,659	0,455	0,402	0,500
52	6.712	0,689	0,583	0,892	0,722	0,562	0,286	0,464	0,440	0,214	0,507	0,520	0,610
53	5.284	0,835	0,825	0,862	0,841	0,493	0,804	0,809	0,702	0,590	0,873	0,888	0,824
61	4.786	1,397	0,361	1,352	1,037	0,795	0,407	0,305	0,503	0,570	1,127	0,228	0,485
62	2.825	1,097	1,084	1,565	1,249	0,449	0,993	0,902	0,781	0,409	0,816	0,676	0,625
63	7.792	0,370	1,346	1,315	1,177	0,146	0,390	0,252	0,269	0,293	0,289	0,144	0,226
64	14.365	0,866	1,253	0,762	1,000	0,758	0,908	0,614	0,759	0,759	0,724	0,816	0,759
71	8.768	0,899	0,651	0,713	0,756	0,144	0,378	0,316	0,279	0,160	0,577	0,443	0,370
72	3.557	1,005	0,683	1,059	1,082	0,997	0,884	0,673	0,778	0,621	1,140	0,836	0,719
73	14.232	0,964	0,852	1,286	1,025	0,295	0,275	0,404	0,315	0,279	0,322	0,318	0,307
74	1.458	1,582	0,339	0,482	0,784	1,868	0,622	0,195	0,901	1,066	2,429	0,452	1,149
81	4.140	0,782	1,068	0,948	0,929	0,213	0,000	0,000	0,071	0,272	0,000	0,000	0,076
82	2.480	0,721	0,984	1,303	1,003	0,964	0,329	0,138	0,477	1,337	0,334	0,105	0,475
Total	678.028	0,990	0,973	0,972	0,985	0,632	0,555	0,596	0,628	0,685	0,674	0,813	0,661
CV across provinces	1,480	0,306	0,316	0,211	0,175	0,554	0,481	0,588	0,390	0,460	0,632	0,526	0,397

LECTURE NOTES :
SAMPLING OVER TIME

Report # 21

by
Vijay Verma

February, 2001

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. ISSUES 218

II. DIVERSE OBJECTIVES 218

III. COMBINATION OF OBJECTIVE; EXAMPLES 222

IV. DURATION OF THE SURVEY PERIOD 223

V. ROTATION BETWEEN ROUNDS 224

VI. ESTIMATION OF CHANGE WITH SAMPLE OVERLAP 225

VII. AGGREGATING OVER SAMPLES 227

VIII. CURRENT LEVELS WITH PARTIAL OVERLAPS 227

IX. QUESTIONS 229

I. ISSUES

There are two main issues to be considered in relation to sampling over time.

The first concerns representativeness of the sample in the time dimension. The use of survey results almost always involves generalizations beyond the limited period for which the information is actually collected. Hence, in addition to the selection of units from a population of units, a sample survey involves the selection of time segments from the entire interval to which the results may be applied. Proper representation of the time dimension can be very important in determining the sample design, especially in the presence of pronounced seasonal and other temporal variations. Only a representative selection of time periods (individual months, quarters, seasons etc.) can permit valid inferences to be made from the sample observations to the total period of interest (such as a whole year or a period of several years). This requirement raises important questions such as: When should the survey be conducted? How long should be the survey and reference periods? Should the survey be concentrated in a single period, or spread out over a number of periods? How should the sample units be distributed for data collection over the survey period? Should individual units be enumerated only once, or repeatedly during the period? Should there be explicit sampling of time segments, whether for defining the information to be collected or identifying the units to be included in the survey? Such considerations generally apply to any survey, though they are more important for certain types of surveys than others.

The second issue concerns the relationship of samples over time in a continuing survey, i.e. in a survey involving repeated rounds. The samples can be independent from round to round or be overlapping. The overlaps may be complete or partial. An overlap may involve following the same ultimate units such as households or individuals, or only retaining the same set of areas or other higher stage units. Different patterns of sample rotation may be adopted to achieve different degrees of overlaps in the samples, and so on.

II. DIVERSE OBJECTIVES

The choice of sample design over time is determined by the survey objectives in terms of the type of statistics to be measured. Several distinct though related objectives may be identified.

It is important to be clear about which of these objectives are relevant in the design of a particular survey.

(1) *Measuring structural features*. In many surveys the primary objective is to obtain information of a structural nature of relatively long-term interest, such as information on the pattern of fertility or the structure of economic activity. In so far as the characteristics of interest do not change rapidly, the survey may be a one-time operation or repeated only at relatively long intervals. In any case, the samples for repetitions will usually be designed independently. On the other hand, the timing and duration of the survey field work deserves a careful consideration. To be of long-term value, the timing of the survey should not coincide with periods of abnormal conditions. For the same reason, and to the extent seasonal and other transient effects are important, the survey period should be long enough to even out these effects (see 3 below).

(2) *Meeting special needs*. Some surveys are designed to meet special needs, such as assessing an abnormal condition after a natural calamity. Here the timing will be determined entirely by the special purpose of capturing the abnormal situation, and generally it will be desirable to make the survey period as short as possible.

(3) *Averaging over time*. The measurement of many characteristics, particularly economic indicators such as employment, income and household expenditure, are affected by seasonal, cyclic and haphazard changes which need to be averaged out to assess the underlying situation. This requires spreading out of the survey period, such as to a whole year, and distributing the sample to be enumerated into `sub-rounds' to average out the transient factors over the whole reference period such as a year.

(4) *Aggregating over time*. This refers to aggregating statistics over time to increase the sample size available for more detailed tabulation and analysis of the data. This objective is distinguished from previous one in that here the main constraint is the sample size which can be enumerated in a given time; by contrast, for the purpose of averaging out transient factors the sample may be spread out deliberately, even when a more concentrated enumeration is feasible. Aggregation is most efficient when the samples being aggregated are non-overlapping and drawn independently of each other.

(5) *Shortening the recall period by cumulating information for individual units*. When information has to be collected with a relatively long reference period, there may be a

choice between (i) obtaining that information through retrospective questioning in a single visit, or (ii) cumulating the information from repeated enumeration of the same units over a period of time, using a correspondingly shorter recall period in each enumeration. The objective of second alternative is to obtain more accurate information than that possible with the long recall period, and it necessarily requires the use of fully or partly overlapping samples in terms of individual units. Examples are the measurement of annual incomes from quarterly or monthly enumeration, or of the individual's usual labor force status for the past year from separate observations for several months or weeks during the year.

The objective of repeated enumeration for such purposes is primarily substantive, determined by the nature of the variables to be measured: for example measuring actual annual incomes is not the same thing as estimating the annual incomes simply as 12 times the reported incomes for a single month for each unit, since the two can give different pictures of the income distribution. This is quite different from cumulating events such as incidence of illness, injuries, births etc. simply to reduce magnitude of the sampling error; the latter objective does not necessitate following up of the same individuals, as it can also be achieved by increasing the number of units in the sample.

Repeated enumeration can be expensive and increase respondent burden. In situations where neither of the above two approaches is feasible or acceptable, it may be necessary to consider a third alternative: (iii) revising the type of information sought, such as accepting income or employment status measured with a short reference period during a single visit, even if the resulting measures are theoretically less satisfactory than those measured, for each unit, over a longer period.

(6) *Bounding the reference period.* A closely related objective is to increase accuracy of the information obtained by `bounding' the reference period between repeated enumeration of the same respondents, that is, by fixing the boundaries of the reference period more clearly in terms of certain identifiable events such as pairs of successive interviews. Like the previous case, this also involves repeated enumeration of the same individual units - but the objective here is different: the objective being to fix more precisely the boundaries of the reference period, rather than to aggregate several short recall periods to construct a longer reference period. Examples are follow-up surveys in which births, deaths and other events or changes occurring during the interval between successive

interviews are recorded generally by reference to and with the help of the information obtained earlier.

(7) *Measuring current levels.* Many surveys are repeated at regular intervals to obtain a time series of current estimates, such as monthly or quarterly estimates of employment and unemployment, production or other economic indicators. Such surveys usually involve strict requirements concerning the timing and duration of the survey rounds. Variance of a current estimate is the same when the samples for the rounds are either independent or completely overlap; however, for complex estimators special procedures can yield modest reduction in variance with partly overlapping samples.

(8) *Measuring net change.* Net change refers to change at the aggregate level, with individual level changes in opposite directions cancelled out. Examples are month-to-month changes in rates of unemployment and other economic indicators. Measuring net changes does not necessitate the use of overlapping samples, though the precision of the estimates can be markedly increased even with partial overlaps. The overlaps do not need to be at the level of ultimate units; overlapping higher stage units such as sample areas is often cheaper and more convenient, and can also give large increases in precision.

The increase in precision in measuring change depends on the degree of correlation over time, which of course tend to be larger for the same individuals and households, than with overlaps only at the level of higher stage units such as sample area.

The measurement of change generally requires larger samples than the measurement of levels with the same precision, though the associated design effects tend to be smaller when estimating change. Often estimates of net change also benefit from the cancelling of similar biases affecting the statistics being differenced.

(9) *Measuring gross change.* This refers to the changes at the level of individual units (such as households and persons), taking account of changes in either direction. Examples are changes in the status of economic activity, industry, occupation, place of residence, marital status etc. of individual persons. Measurement of gross change necessarily requires repeated and strictly comparable measurement on the same set of units of analysis. Even random errors in the repeated measurements would result in biased estimates of gross change. (By contrast, such errors only increase variances in the case of measures of net change.)

(10) *Measuring dynamics of change*. The measurement of dynamics of change also requires following-up of the same individuals. The distinction from the previous objective is mainly that of degree: usually it is necessary to follow the same `panel' of individuals over a relatively prolonged period of time and obtain inter-related information which can capture and illuminate the dynamics of change at the micro-level. Panel surveys present complex problems of sample implementation and analysis, as distinct from ordinary `cross-sectional' surveys including those involving sample overlaps.

(11) *Sampling of time segments for collecting information*. Apart from spatial sampling of units to be included in the survey, the reference period covered for any selected unit can also be subject to the sampling process: it is often not appreciated in survey practice that it is by no means essential to cover (to take a `census' of) the entire reference period for each unit in the sample. The objective of sampling the time segments include reducing the work-load, reducing respondent burden, and spreading out the sample to more units within the given resources. In an establishment survey for instance, information may be obtained from each selected unit for a sample of weeks during the year; the sampling of units and of time-segments may be combined in various ways to obtain the required overall selection probabilities. The same applies to confining intensive observation such as keeping of diaries to a sample of time periods in a household expenditure survey which uses extended reference period for measuring other variables.

(12) *Sampling of time segments for selecting units*. A related objective can be to use time-sampling to identify the units to be included in the sample. This procedure is useful for capturing mobile populations, making use of their propensity to visit identifiable fixed locations (the locations themselves may be appropriately sampled). Examples are surveys of clients visiting hospitals or other institutions, survey of tourists or international migrants at ports of entry and exit, and surveys of nomads or live stock at water points.

III. COMBINATION OF OBJECTIVE; EXAMPLES

Usually several of these objectives may be present in any given survey, in which case the survey design has to be worked out as a compromise between differing requirements.

A few examples may be mentioned.

In the Indian National Sample Survey, the survey programe basically consists of a continuing series of annual surveys. The main subject matter covered and the sample is generally changed from year to year. Fieldwork for each annual survey is divided into monthly sub-rounds, each covering a distinct representative subsample. The primary objectives of the design are measuring structural features, and averaging and aggregating over months of the survey.

The labor force survey in Indonesia is conducted on a quarterly basis. Each quarter a different sample is enumerated, but there is a large overlap between quarterly samples exactly one year apart. The main objectives are averaging and aggregating over time for the quarterly rounds; measuring current level quarterly and, more importantly, annually; and measuring year to year net change.

The "continuous" household budget survey of Spain is conducted on a quarterly basis. Each selected household is enumerated for 8 consecutive quarters before being dropped from the sample, thereby giving a 7/8 overlap between successive quarters. Information on various items of expenditure is obtained with differing reference periods, from three months (retrospective questioning) to continuous recording of daily expenditures in a diary. The objectives of this design include: averaging over quarters of the year [3]; shortening and bounding the recall period [5 and 6]; and sampling time segments for collecting detailed information on daily expenditures [11].

IV. DURATION OF THE SURVEY PERIOD

In many surveys, it is important to spread out the survey period sufficiently to average out the effect of seasonal and other short-term variations which may distort the result of main interest. Many survey of household income and expenditure are for example conducted over a period of one full year. Similar considerations apply to surveys of agriculture.

By contrast, there are many topics which can be investigated adequately without worrying too much about seasonal or other temporal variations. In such surveys the duration of field work is determined primarily by practical considerations such as the rate at which the work can be done effectively and efficiently. Examples are fertility and similar demographic surveys, in many of

which the whole field work has been completed within a short period of 1-3 months.

While proper representation of the time dimension is important, the above provides an example where representativeness in terms of space can take precedence over representativeness in time, "because correlations are greater for a given unit over time than for neighboring units in space" (Kish, Survey Sampling, Sec 12.5). For the same reason, greater ambiguity can often be tolerated in the definition of limits of time as opposed to the more clearly identifiable limits of space.

V. ROTATION BETWEEN ROUNDS

Sample rotation means that from one round to the next, some or all of the sample units are replaced by new units. In many surveys, the pattern of rotation is "symmetrical". This means that new sets of units are introduced into the sample at regular intervals, and once introduced, each set is retained or dropped from the sample following the same pattern. Once such a system is established, it provides a fixed degree of overlap from any one round to the next, and between any two rounds separated by a given time interval.

Many surveys use a straightforward pattern of rotation. The sample consists of "n" subsamples; at the beginning of each survey period, one new subsample is introduced; and each subsamples remains in the survey for n consecutive periods(rounds). The overlap between rounds decreases linearly as the interval separating them increases. For two samples introduced i interval apart the overlap is (n-i)/n, up to i=(n-1); after which (i=n) the overlap becomes zero.

More complicated rotation patterns can be used to vary the degree of sample overlap and how it changes with time.

Another point to note is that in many situations, the sample is rotated slowly (or not at all) at higher stage units, and more rapidly as we move to lower stage units. The objective is to reduce the cost and inconvenience of changing the PSUs and other higher stage units.

At the ultimate stage, the rotation pattern is often applied in terms of addresses or other structural units, rather than to the actual households or persons living in those. Because of mobility of the latter type of units, the overlaps actually achieved are somewhat reduced.

VI. ESTIMATION OF CHANGE WITH SAMPLE OVERLAP

The difference in estimates between two overlapping samples is given by

[1] $$Var(\bar{y}-\bar{x}) = Var(\bar{y})+Var(\bar{x})-2.Cov(\bar{y},\bar{x})$$

The two samples may refer to the same statistic measured at two points in time as in a repeated survey. The covariance term arises from the overlap or other relationship between the samples. Generally it is a positive quantity because of repeated measurement on the same units. Hence precision in estimating the change is increased with overlapping samples. In a multistage design the samples may be related though common higher stage units without necessarily involving overlaps in terms of the same ultimate stage units.

For two simple random samples each of size n and with proportion P of the samples overlapping, and for simplicity assuming that the element variances are the same, the above expression becomes

[2] $$Var(\bar{y}-\bar{x}) = \frac{S^2}{n}+\frac{S^2}{n}-2PR\frac{S^2}{n}; \quad Var_o = \frac{2S^2}{n}$$

or more concisely:

[3] $$Var(\bar{y}-\bar{x}) = Var_o.(1-PR)$$

Var_o is the variance of the difference when there is no overlap (i.e. with independent samples). With a complete overlap the variance is reduced by the factor (1-R), where R is the correlation coefficient.

The estimation of change can be made more precise by giving increased weight (W_P) to the overlapping part which contributes less to the variance, and giving a correspondingly smaller weight (W_Q) to the non-overlapping part. (This of course assumes that each of the two parts constitutes a representative sample in its own right.) In general terms we can write

[4] $$(\bar{y}-\bar{x})_W = W_P(\bar{y}-\bar{x})_P+W_Q(\bar{y}-\bar{x})_Q$$

giving for the variance of the weighted differences

[5] $$Var(\bar{y}-\bar{x})_W = Var_o.[(1-R)W_P^2/P + W_Q^2/Q]$$

The above follows from the observation that the first component corresponds to a complete overlap, but with the sample size reduced from n to n.P; while for the second

component, there is no overlap and the sample size is n.Q. The optimum solution for minimum variance of the difference is to weight the components in inverse proportion of their unit variances, i.e. reduce the relative weight of the non-overlapping component by the factor (1-R). This gives the normalized weights (adding to 1.0 for the two components) as

[6] $$W_P = P/(1-QR); \quad W_Q = (1-R).Q/(1-QR)$$

and hence

[7] $$Var(\bar{y}-\bar{x})_W = Var_o.[(1-R)/(1-QR)]$$

With reweighting as above, a modest overlap such as P=1/3 may give almost the full benefit (1-R) of complete overlap.

For the more general case when the two samples are of different sizes, we have

[8] $$Var(\bar{y}-\bar{x}) = \frac{S_x^2}{n_x} + \frac{S_y^2}{n_y} - 2\,R_{xy}\,S_x\,S_y.\frac{n_c}{n_x n_y}$$

giving, with the element variances assumed equal,

[9] $$Var_o = S^2.[\frac{1}{n_y} + \frac{1}{n_x}] = \frac{2S^2}{\tilde{n}}, \text{ with } \tilde{n} = \frac{2n_x n_y}{n_x+n_y}$$

and hence replacing [3] with

[10] $$Var(\bar{y}-\bar{x}) = Var_o.(1-\tilde{P}R); \text{ with } \tilde{P} = \frac{2n_c}{n_x+n_y}$$

In application, the above formulation is somewhat more general than assumed in its derivation. Differences in the element variances (usually not important in samples from the same population with a short time interval between them) can be absorbed in the definition of the sample sizes if necessary. Approximately the relationship should also hold for more complex sample designs, though with clustering where the overlap is at the level of higher stage units rather than at the level of the ultimate units, the gain in precision will be underestimated. In such situations, it may be possible to improve the approach by defining P as the proportion of units coming from the common PSUs, and R as the correlation between these units (which may be much lower than that for actually overlapping ultimate units).

VII. AGGREGATING OVER SAMPLES

For aggregation over samples, the reverse is the case. The variance of the sum of two sample estimates is:

[11] $$Var(\bar{y}+\bar{x}) = Var(\bar{y})+Var(\bar{x})+2.Cov(\bar{y},\bar{x})$$

so that the positive correlation due to the overlap increases the variance of the aggregate as follows:

[12] $$Var(\bar{y}+\bar{x}) = Var_o.(1+PR)$$

The efficiency can be improved by weighting up the non-overlapping component (which contributes less to the variance of the aggregate) by the factor (1+R), i.e. by assigning the normalized weights

[13] $$W_P = P/(1+QR); \quad W_Q = (1+R).Q/(1+QR)$$

giving

[14] $$Var(\bar{y}+\bar{x})_W = Var_o.[(1+R)/(1+QR)]$$

In comparison with the case of differences, the gain due to the reweighting is modest. Aggregation is served best by having no overlap.

VIII. CURRENT LEVELS WITH PARTIAL OVERLAPS

For an ordinary mean such as

[15] $$\bar{y} = Q.\bar{y}_Q + P.\bar{y}_P$$

the estimation of a current level is not affected by the degree of overlap. However with a partial overlap, a modest improvement can be achieved by adjusting the estimate for the overlapping part from regression over the full earlier sample, and then appropriately reweighting the two components to produce a composite estimate of the current level. The current estimate for the

overlapping part is adjusted from the results of the previous sample as

[16] $$\hat{y}_P = \bar{y}_P + R(\bar{x} - \bar{x}_P)$$

Here $\bar{x}$ is the estimate from the previous round for the whole sample of that round, and $\bar{x}_P$ the same for the part overlapping with the current round.

(Note. In the above, the regression coefficient has been replaced by the correlation coefficient R on the reasonable assumption that $S_x = S_y$.)

Next, the composite current estimate is computed as

[17] $$\bar{y} = W_Q.\bar{y}_Q + W_P.\hat{y}_P$$

with weights

[18] $$W_Q = Q - p.Q; \quad W_P = P + p.Q; \quad where\ p = PQ.\frac{R^2}{1 - Q^2 R^2}$$

giving

[19] $$Var(\hat{y}) = (1 - p).Var(\bar{y})$$

It can be shown that minimum variance

[20] $$Var(\hat{y})_{opt} = (1 - p_{opt}).Var(\bar{y})$$

is obtained with

[21] $$p_{opt} = \frac{1 - \sqrt{1 - R^2}}{2}$$

corresponding to the overlap

[22] $$P_{opt} = \frac{\sqrt{1 - R^2}}{1 + \sqrt{1 - R^2}}$$

The actual gain may be larger in a time series of survey rounds since the preceding rounds also benefit from the composite estimation.

Several variations of the basic procedure described above are possible. In some establishment surveys for example, information is obtained from each establishment in a round

for two or more periods preceding the survey, so that composite estimates can be produced even without an overlap between the rounds in terms of the actual establishments enumerated.

IX. QUESTIONS

Describe the terms with examples: survey period; reference period; moving and fixed reference period; recall period; bounded and unbounded recall period; gross and net change.

(1) In the monthly Current Population Survey of the US, a new subsample is introduced each month; it is enumerated for 4 months, then dropped for the next 8 months, and again introduced for the final time for 4 months. At any time in the survey, there are eight such subsamples - from the one newly introduced to the one being enumerated for the last (eighth) time. Show this pattern of rotation diagrammatically. How does the level of overlap vary as a function of the time interval between a pair of rounds? What is the objective of choosing such a rotation pattern? The overlaps actually achieved are lower than those suggested above. Can you think of a reason for that?

(2) In a quarterly labor force survey in a developing country, there is no overlap in the sample from one quarter to the next, but a 50% overlap in the samples for any pair of rounds exactly one year apart. Discuss the objectives of this pattern and practical aspects in its implementation. (This illustration is taken from Indonesia.)

(3) In relation to the effect of overlaps on estimating changes and cumulations, construct tables showing the factors

1-PR, (1-R)/(1-QR); 1+PR, (1+R)/(1+QR)

as functions of P and R with the following values of these parameters:

P = 0, 1/4, 1/3, 1/2, 2/3, and 1

R = 0, .25, .50, .75, .90, .95, and 1.0.

(4) Develop and discuss the following expressions:

[23] $$Var_o = S^2.[\frac{1}{n_y} + \frac{1}{n_x}] = \frac{2S^2}{\tilde{n}}, \text{ with } \tilde{n} = \frac{2n_x n_y}{n_x + n_y}$$

[24] $$Var(\bar{y} - \bar{x}) = Var_o.(1 - \tilde{P}R); \quad \tilde{P} = \frac{2n_c}{n_x + n_y}$$

for two samples with overlaps for the following special cases:

[25]

$$n_c = 0$$
$$n_x = n_y = n_c$$
$$n_x = n_y > n_c$$
$$n_x > n_y = n_c$$
$$n_y > n_x = n_c$$

In the above, y refers to the current sample, x to that for the preceding period, and c to the overlap between the two. (These illustrations come from *Survey Sampling* by L Kish.)

AREA SAMPLING
FOR AGRICULTURE CENSUS

Report # 28

Sri Budianti Sukmadi

March, 2001

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 233

II. AREA SAMPLING 234

III. BASIC TABLES FOR EVALUATION 235

IV. CONSTRUCTION OF MASTER FRAME AND MASTER SAMPLING FRAME 238

V. METHOD OF EVALUATION 239

- **A. Possible Methods** 239
- **B. Recommended Method** 240
  - **1. Complete Count Basis** 240
  - **2. Sample Basis** 240

APPENDIX A TABLES TO BE PRODUCED 241

I. INTRODUCTION

The scope of the Agriculture Census includes all agricultural households with or without land holding and all agricultural establishments. Census Blocks (CB) derived from the 2000 Population Census can be very effective 1) if used as a benchmark for listing agricultural households and 2) if used for updating the maps of census blocks before conducting field operations of the Agriculture Census (which takes place about one year before field operations).

Three alternative enumeration procedures are possible:

1. Complete count by listing all buildings and households in every census block. In this case, there is still some scope for adding more items to the non-Agricultural sector as long as they don't increase the burden on respondents.
2. Complete count for concentration areas only and on a sample basis for other areas (either urban or rural).
3. Sample: by selecting census blocks either in concentration areas or other areas.

If no budget constraint existed, the first alternative would provide the most complete results. Construction of a master frame for subsequent agriculture surveys will become more comprehensive and a more careful construction and selection of items in the listing questionnaire can be accomplished. Items in the listing questionnaire should allow two things: first the provision of basic agricultural data (such as the structure of the agricultural sector) for policy formulation and second the production of a master frame and a master sampling frame. Detailed characteristics to be obtained will not be designed as part of a complete count, but rather on a sample basis as was done in previous agriculture censuses because the operation would be otherwise too complicated and would not produce results before a long lag.

The possibility of using Population Census 2000 results as a master frame with employment characteristics can be considered as long as:

1. It is conducted on a sample basis because available budgets do not allow complete count.
2. The compilation of detailed characteristics after the complete count operation is still needed. Due to the time lag between the complete count and the sample census (one year), the master frame must be based on sources other than the agriculture census itself.

II. AREA SAMPLING

Processing of the Population Census 2000 will be finished around June 2001, but the evaluation needs to be started as soon as possible, by using data from Population Census 2000 of some provinces which have already been processed (i.e. cleaned) in March. The data need to be evaluated just for stratification purposes, at least for identifying concentration areas and other areas.

A complete count of all buildings and households in all census blocks will be very expensive since the listing activity has to be conducted in both concentration areas and other areas including those with no agricultural households. One should note that during the listing activity planning questionnaire design should also include variable selection, questionnaire format, field operation, database construction and use of data either for policy planning or for the construction of a master frame and a master sampling frame.

There are five alternative concepts which need to be considered for stratification purposes:

Concept 1: the agricultural household (a household having at least one farmer) as a sampling unit.

Concept 2: households having a member engaged in agricultural activities by sub-sector and whose employment status is:

- Self-employed
- Self-employed assisted by temporary employees
- Employer

Households having more than one farmer with different employment status would then be classified as farming households with the sub-sector corresponding to the farmer whose status is "employer". Or if this status is not there the farming households would be classified according to the sub-sector of the farmer whose status is "self-employed assisted by temporary workers". And if this status was also not there then it would be classified according to the sub-sector of the farmer whose status is "self-employed". If all farmers in the household have the same status, then they would be classified according to the main activity of farmers in the census block.

Concept 3: The same as concept 2 but with the priority of the head of household activity in the agricultural sector. If the head of household does not engage in an agricultural activity then the next member of household who is engaged in an agricultural activity would be considered as a criterion for the sub-sector.

Concept 4: Using "persons engaged as farmers" rather than agricultural households. One

household may have more than one farmer.

Concept 5: Using all household members engaged in agricultural activities including those whose status is "employee".

The basic data is used to determine whether the census block provides a good potential for agriculture and which concept is the best for defining potential agricultural areas. This is not only important for identifying concentration areas and for stratification, but also for calculating the budget needed for conducting the agricultural census.

III. BASIC TABLES FOR EVALUATION

The census blocks are divided into three categories: common, preparation and special census blocks. Evaluation will apply to common census blocks and will be conducted in several provinces by using the complete count results of the Population Census 2000. Since provinces have different characteristics and different patterns of agricultural activity, these differences need to be evaluated. Evaluation should include the following steps:

1. Basic data on agriculture activities: compare the total number of agricultural households (and their size) by using the five concepts mentioned above and differentiating between urban and rural areas.

2. Distribution of census blocks by group of agricultural households: that will provide an idea for how to divide them into concentration areas and other areas. Census blocks with no agricultural households will automatically be included in non concentration areas and will be grouped in one stratum.

3. Census blocks with no agricultural households but with employees engaged in agricultural activities: this domain is closely related to agricultural activities of either land holding or non land holding farmers.

To make sure which concepts are suitable for stratification, the evaluation of tables and tests of significance will be adopted by using Student's *t*-test statistics. If the statistics is not

significantly different among the concepts, then using the farmer concept is the best way for stratification since it is the easiest one to apply to the population census items (individual information). Using the total number of farmers is good for stratification because the farmer is a direct member of the household who engages in agricultural activities as a manager or as self-employed. Compared with the concept used in the previous agricultural census, using the agricultural household is preferable. In applying the agriculture household concept one has no problem identifying individual farmer activities if the questionnaire is adequately designed. The tabulation plan will be prepared to consider the stratification possibility.

Tables designed for stratification consideration include the following:

Table 1 : Total number of common census blocks and agricultural households by concept.

Table 2 : Total number of common census blocks and farmers by sub-sector.

Table 3 : Total number of common census blocks and household members engaged in agricultural activities including employees.

Table 4 : Total number of common census blocks with agricultural households divided into those with and without employees in agriculture.

Table 5 : Total number of census blocks by total number of agricultural households.

Table 6 : Total number of census blocks by total number of farmers.

Table 7 : Total number of agricultural households by number of activities.

Table 8 : Total number of common census blocks by type of industry.

Table 9 : Total number of census blocks by agricultural household activity and group of households.

Table 10 : Total number of census blocks by group of total number of non-agricultural households activity.

Table 11 : Total number of census blocks by group of agricultural households and kabupaten/kota.

The study using Population Census 2000 needs to be realized as soon as possible, especially if the agriculture census will be conducted in 2003.

The proposed integrated designs are divided into:

a. Agriculture with land holding divided by:
   - Food crops
   - Horticulture
   - Animal husbandry
   - Plantation/estate crops

b. Agriculture with no holding
   - Fishery

For some sub-sectors, it is still not clear where they fit, e.g. land fishery, forestry and other rare cases of agricultural activities. Design integration mentioned above applies not only to implementation of the listing activity but also to that of the agriculture census itself and the subsequent agricultural surveys.

Single purpose surveys collecting data for every commodity are not practical because:

- First, they do not use an integrated questionnaire, and therefore would not allow production of data covering the agricultural sector as a whole or at least by group of sub-sectors.

- implementation takes a long time and field operations are complex. For example if in the plantation/ estate crops sub-sector we have to conduct single commodity surveys (such as on rubber, coconut, palm oil, cocoa, tobacco etc), implementation will be difficult because there would be too many surveys required.

- they are very costly since each survey will need a separate budget (for recruiting enumerators, training etc.). Even listing will be very costly.

Grouping surveys through integrated designs will be more practical if a good frame is available with adequate information on the type, structure, and characteristics of commodities.

The Population Census 2000 data provide an adequate base for making stratification because the observation and construction of sample selection obtained from agriculture census data itself will take much time and is not possible according to:

- Processing time, due to the last agriculture census processing were conducted in around one year.

- The time of observation for stratification and sample selection.

- Preparation of the whole planning should be based on already fixed design so the construction of activity proposal including the budget has to be submitted to the government long before.

Beside those problems, we have also to consider the policy of regional autonomy to give guidance to the BPS regional offices. The simple method determined should be able to be adopted easily either by subject matter unit or regional offices and more flexible if they want to get more additional information to meet their regional policy needs depending on the regional priority.

IV. CONSTRUCTION OF MASTER FRAME AND MASTER SAMPLING FRAME

A maximum of six master sampling frames can be constructed: food crops, horticulture, animal husbandry, plantation/estate crops, fishery, and other activities. However, one can minimize the number of master sampling frames by merging some sub-sectors where field operations are similar and can therefore be conducted at the same time. Food crops and horticulture, for example, are closely related: one wet land area sometimes has mixed crops or can be planted with paddy in one season and with vegetable in another. Food crop activities and animal husbandry are also highly correlated. Past experience suggests that many farmers are engaged in both activities. In this case, a correlation test can be attempted to determine whether the assumed correlation of variable integration in the stratification, listing, questionnaire, and field operation, is valid.

If the Agriculture Census 2003 is conducted as a complete count, the master frame and master sampling frame would be constructed from the Agriculture Census 2003 as a whole. Construction of a master frame from Population Census 2000 would still be needed: it would be very useful at least for 2003 and 2004. For 2003, it will be used for the sample census (after the implementation of the complete count) for detailed information and for any agricultural survey during that time since no other source would be available. For 2004, it could be used as a frame for sample selection before the results of the Agriculture Census are available. Starting in 2005, the master frame and master sampling frame derived from the Agriculture Census 2003 will be used for subsequent agricultural surveys.

To make sure that the results of Population Census 2000 are suitable for constructing a master frame either for stratification (concentration areas or non concentration area) or to decide which integrated designs to adopt, other data sources such as Podes can be used. In Podes there are data on land used for agriculture (classified by wet land, dry land, plantation/estate crops area) which can be used for evaluation.

V. METHOD OF EVALUATION

A. Possible Methods

Four types of evaluation can be conducted:

a. Evaluation of basic data by using Tables 1 to 11 to determine the distribution of census blocks and the structure of data.

b. The significance test: i.e. whether there exists a high correlation between activities. The simple test of significance will suggest which agricultural activities can be grouped in one master sampling frame. Simple sample designs are recommended: in every stratum/domain of study, for example, by systematic sampling or PPS sampling using the total number of agricultural households as sampling units. The selection of the ultimate sampling units depends on the type of agriculture survey, e.g. food crops, horticulture etc.

c. Measuring the level of relative variation of the total number of agricultural households of each sub-sector in the census block.

d. Evaluation of how big the census blocks to be selected in each type of master sampling frame from the master frame.

The construction of a master sampling frame should be done with sampling and with a maximum of five sampling frames. The types of master sampling frame, stratification, and sampling unit are described in Table 12:

Table 12
Types of Sampling Frame, Stratification and Sampling Units

Master sampling frame	Stratification	Sampling unit
Food crops Horticulture Animal husbandry Plantation/estate crops Fishery	Census block concentration Census block non concentration Census block with no agricultural household Empty census block which is subject to change to common census block	Census block Agricultural household as size Agricultural household in specific sub-sector as an ultimate sampling unit

B. Recommended Method

1. Complete Count Basis

If the Agriculture Census 2003 is conducted on a complete count basis, the evaluation should be based on the results of this census. The construction of a master frame and a master sampling frame will be more reliable especially if the characteristics in the complete count census covers the needed data. Construction of the questionnaire will have to be integrated with data needed for the master frame. The overall planning of agriculture survey activities will have to be linked with the planning and preparation of census activities including concepts and definitions, classification and level of estimation (province, kabupaten/kota, etc). The level of estimation will affect the construction of the master sampling frame.

2. Sample Basis

The sample selection as mentioned before is used as a basic master sampling frame (Population Census 2000 results as a benchmark) and after that the data from Population Census 2000 will be revised with the data from Agriculture Census 2003 for constructing the master sampling frame. Evaluation of basic data and tests of significance by using the Agriculture Census 2003 results will be more comprehensive and reliable, but will also depend on the data (response bias, non response, effect of seasonal factors, and interviewer qualification during field operations).

APPENDIX A
TABLES TO BE PRODUCED

Table A.1
Total Number of Common CB's and Agriculture Households by Type o f Concept

Province: ..... Urban/Rural/Urban + Rural

Type of concept	Number of CB's		Number of agriculture households	Average per CB Col (4) : (3)
	With no agriculture households	With agriculture households		
(1)	(2)	(3)	(4)	(5)
1. Concept 1 Agriculture households				
2. Concept 2 a. Food crops b. Horticulture c. Fishery d. Animal husbandry e. Plantation Total	x	x		x
3. Concept 3 a. Food crops b. Horticulture c. Fishery d. Animal husbandry e. Plantation Total	x	x		x

Total number of common CB's: Urban:..... Total number of households: Urban: ....
Rural: ..... Rural: .....
Urban + rural: ..... Urban+rural: .....

Household data: one household is classified only in one sub-sector

Column (4) : total 2a to e = total 3a to e = total agriculture households (concept 1)

Table A.2
Total Number of Common CB's and Farmers by Sub-sector
(Concept 4)

Province :.....

Sub-sector	Number of CB's		Number of farmers	Average farmers per CB Col (4):(3)
	With no farmer	With farmers		
(1)	(2)	(3)	(4)	(5)
1. Urban				
a. Food crops				
b. Horticulture				
c. Fishery				
d. Animal Husbandry				
e. Plantation				
f. Total				
2. Rural				
a. Food crops				
b. Horticulture				
c. Fishery				
d. Animal Husbandry				
e. Plantation				
f. Total				
3. Urban+Rural				
a. Food crops				
b. Horticulture				
c. Fishery				
d. Animal Husbandry				
e. Plantation				
f. Total				

Column (4) : f =a+b+c+d+e+f

Table A.3
Total Number of Common CB's and Household Members Engaged in Agriculture Activities Including Employees
(Concept 5)

Province : .....

Sub-sector	Number of CB's		Number of persons in agriculture	Average persons per CB Col. (4) : (3)
	With no person in agriculture	With person in agriculture		
(1)	(2)	(3)	(4)	(5)
1. Urban				
a. Food crops				
b. Horticulture				
c. Fishery				
d. Animal Husbandry				
e. Plantation				
f. Total				
2. Rural				
a. Food crops				
b. Horticulture				
c. Fishery				
d. Animal Husbandry				
e. Plantation				
f. Total				
3. Urban + Rural				
a. Food crops				
b. Horticulture				
c. Fishery				
d. Animal Husbandry				
e. Plantation				
f. Total				

Column (4) : f =a+b+c+d+e+f

Table A.4
Total Number of Common CB's with No Agriculture Household by With Employee / With No Employee in Agriculture

Province : .....

Item	Total CB's	With employee in agriculture	With no employee in agriculture
(1)	(2)	(3)	(4)
Urban			
1. Number of CB's			
2. Number of employees	x		
3. Average employees per CB	x		
Rural			
1. Number of CB's			
2. Number of employees			
3. Average employees per CB	x		
	x		
Urban + Rural			
1. Number of CB's			
2. Number of employees	x		
3. Average employees per CB	x		

Table A.5
Total Number of CB's by Total Number of Agriculture Households

Province : .....

Group of agriculture households	Urban		Rural		Urban+Rural	
	CB	Average per CB	CB	Average per CB	CB	Average per CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0						
1 - 4						
5 - 9						
10 - 19						
20 - 39						
40 - 59						
60 - 99						
100 - 149						
150						
Total						

Table A.6
Total Number of CB's by Total Number of Farmers

Province : .....

Group of farmers	Urban		Rural		Urban+Rural	
	CB	Average per CB	CB	Average per CB	CB	Average per CB
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0 1 - 4 5 - 9 10 - 19 20 - 39 40 - 59 60 - 99 100 - 149 150						
Total						

Table A.7
Total Number of Agricultural Households by Number of Activities

Province : .....

Number of activities	Urban		Rural		Urban + Rural	
	Number	Percentage	Number	Percentage	Number	Percentage
(1)	(2)	(3)	(4)	(5)	(6)	(7)
One Two Three or more						
Total		100.0		100.0		100.0

Table A.8
Total Number of Common CB's by Type of Industry

Province : .....

Type of activity *)	Urban		Rural		Urban + Rural	
	Number	Percent	Number	Percent	Number	Percent
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Total CB's		100.0		100.0		100.0
2. Agriculture Households						
3. Farmers						
a. Food crops						
b. Horticulture						
c. Fishery						
d. Animal husbandry						
e. Plantation						
4. Other establishments						
a. Manufacturing						
b. Trade						
c. Services						
d. Transportation						
e. Others						

*) Each line percentage to the total is calculated separately

Table A.9
Total Number / Percentage of CB's by Agriculture Households Activity and Group of Agriculture Household

Province : ..... Urban / Rural / Urban + Rural

Group of agriculture households in CB	Food crops	Horticulture	Fishery	Animal husbandry	Plantation
(1)	(2)	(3)	(4)	(5)	(6)
0					
1 - 4					
5 - 9					
10 - 19					
20 - 39					
40 - 59					
60 - 99					
100 - 149					
150					
Total	100.0	100.0	100.0	100.0	100.0

Table A.10
Total Number / percentage of CB's by Group of Total Number of Non Agriculture Households

Province :..... Urban / Rural / Urban + Rural

Number of households in CB	Manufacturing	Trade	Services	Transportation	Others
(1)	(2)	(3)	(4)	(5)	(6)
0 1 - 4 5 - 9 10 - 19 20 - 39 40 - 59 60 - 99 100 - 149 150					
Total	100.0	100.0	100.0	100.0	100.0

Table A.11
Total Number / Percentage of CB's by Group of Agriculture Households and Kabupaten / Kota

Province : ..... Urban / Rural / Urban + Rural

Kab / Kota	Group of agriculture households									Total
	0	1-4	5-9	10-19	20-39	40-59	60-99	100-149	150	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
										100.0 100.0
Total										100.0

Table A.12
Coefficient of Variation (CV) of Total Number of Agriculture Households by Kabupaten / Kota

Province : ..... Urban / Rural / Urban + Rural

Kab / Kota	Agriculture households	Sub-sector				
		Food crops	Horti-culture	Fishery	Animal husbandry	Plantation
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Province						

Guidance for Tabulation

1. Table 1, 2, 3:

Concept 1 : Agriculture household (a household having at least one farmer) as a sampling unit.

Concept 2 : The data used for stratification purposes should be households having members engaged in agricultural activities by sub-sector and whose employment status is:
- Self-employed
- Self-employed assisted by temporary employees
- Employer

Households having more than one farmer with different employment status would then be classified as farming households with the sub-sector corresponding to the farmers whose status is "employer". Or if this status is not there the farming households would be classified according to the sub-sector of the farmer whose status is "self-employed assisted by temporary workers". And if this status were also not there then it would be classified according to the sub-sector of the farmer whose status is "self-employed". If all farmers in the household have the same status, then they would be classified according to the main activity of farmers in the census block.

Concept 3 : Same as in concept 2 but with the priority of the head of household activity in the agriculture sector. If the head of household does not engage in any agriculture activity then the next member of the household who is in charge of an agricultural activity will be considered as a criterion for the sub-sector.

Concept 4 : Using "persons engaged as farmers" rather than agricultural households. In one household there might be more than one farmer.

Concept 5 : Using all household members engaged in agricultural activities including those whose status is employee.

2. Table 4 : CB with no agricultural households (main industry code 1 to 5 and employment status code 4 and 5 or main industry code 6 to 9)

3. Table 5: Agriculture household: at least one person as farmer

4. Table 6 : See Table 1 (concept 1)

5. Table 7 : Household with main industry code 1 to 5 and employment status (combination, one or more combination) code 1 to 3

6. Table 8 : CB with member of household in employment status code 1 to 3, and type of main industry code 1 to 9 and 0.

7. Table 9 : See Table 8, each column is calculated separately

8. Table 10 : See Table 8, each column is calculated separately

9. Table 11: See Table 1 (concept 1)

CHOICE OF SAMPLING UNIT IN THE AGRICULTURE DAMPLE CENSUS 2003

Report # 32

by
Sri Budianti Sukmadi

May, 2001

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 253

II. PLANNED PILOTS AND FRAME CONTENT 254

III. VARIATION OF COMMODITY TYPE AND PRODUCTION . . 256
- **A. HORTICULTURE** 256
- **B. PLANTATION** 258
- **C. LIVESTOCK AND POULTRY** 259

IV. SAMPLING FRAME EVALUATION 261
- **A. Type of Sampling Unit** 261
- **B. Agriculture Household By Commodity as a Sampling Unit** 261
- **C. Agriculture Household By Sub-Sector As a Sampling Unit** . 262

V. CONCLUSIONS 264

APPENDIX A PRE-CODED COMMODITIES IN ST2003-L2 266

I. INTRODUCTION

According to the present plan, the Agriculture Census 2003 will produce three things: basic agricultural data (such as the structure of the agricultural sector) for policy formulation, a master frame and a master sampling frame. Detailed characteristics, which will be obtained from a detailed questionnaire, will not be based on a complete count, but rather on a sample basis as was done in previous agriculture censuses, since the operation would be otherwise too complicated and would not allow production of results before a long lag. Two activities will also be undertaken in pilot studies:

1. A pilot study on basic data by the complete count method by using questionnaire called ST2003-L1 and ST2003-L2 as a listing questionnaire and a questionnaire for individual agricultural households designed to provide basic structural data.

2. A pilot study to obtain detailed agricultural data by sub-sectors and commodities divided into

   a. Agriculture with land holding:
      - Paddy and main food crops
      - Horticulture
      - Animal husbandry
      - Plantation/estate crops
      - Fish culture
      - Wood forestry culture

   b. Agriculture with no holding:
      - Fishery
      - Forestry (collecting forest products such as wood and hunting animals)

The pilot studies will include also an evaluation of the master frame derived from the Population Census 2000 which is currently in the finalization and evaluation phase. As mentioned in an earlier report[1], grouping surveys through integrated designs relying on stratifications based on the Population Census 2000 will be more practical.

The possible combinations of sampling frame, stratification, and sampling units are provided in Table 1.

---

[1] "Use of Population Census 2000 to Improve Household Survey Frame," STAT Project Report #18, December, 2000.

Table 1
Combination of Master Frame, Stratification and Sampling Unit

Master sampling frame	Stratification	Sampling unit
Food crops Horticulture Animal husbandry Plantation/estate crops Fishery	Census block concentration Census block non concentration Census block with no agriculture household Empty census block which is subject to change to common census block	Census block Agriculture households as size Agriculture households in specific sub-sector as an ultimate sampling unit.

II. PLANNED PILOTS AND FRAME CONTENT

When evaluating the planned pilot studies against data available in the frame, two main problems arise:

1. Sectoral classification in the frame and that of the sample census are not fully consistent: the sample census has eight sub-sectors while the frame has only five. Three sub-sectors (fish culture, wood forestry culture and forestry –collecting forest products such as wood and hunting animals-- are not available in the frame because they were grouped with other activities in the Population Census 2000. Because these three sub-sectors cover specific locations, it may be more practical to give them special treatment after processing of ST2003-L1 and ST2003-L2 is finished. In any case, these sub-sectors are more useful for regional planning than for national planning, particularly in "potential areas" as defined by the autonomy policy.

2. Second, if the sample census was not based on an integrated design, data at the provincial and national levels will be very difficult, if not impossible, to produce. Although from the ST2003-L1 and ST2003-L2 forms basic structural data can be produced for national and small area administrative units, such data are very limited: it will include, for example, the total number of agricultural households by commodity, or the number of plantation areas, number of trees, number of livestock/poultry but will not cover finer geographic distributions.

The present plan also leans towards adopting a uni-purpose design by commodity directly as a sampling unit, thus making it possible to produce data only at the kabupaten level. Some of the drawbacks of this method are:

1. Every kabupaten has different main "potential" commodities. The number may be less than five commodities or more than ten. It is not clear who determines the main commodities in the Census Block area, village, sub-district or kabupaten. Also, what should the threshold number of commodities to consider as main "potential" commodities be: 5 or a different number? These questions need to be answered clearly if such methodology is adopted.
   If the agriculture household by each type of commodity is considered as a sampling unit, every kabupaten will have a different sampling unit classification depending on the main relevant commodities in that area. It will be possible, for example, that an agriculture household will be included in more than one sampling unit (e.g. a horticulture agriculture household can be also considered as an orange farming sampling unit and a mango farming sampling unit). For the estimation at the kabupaten level there will be no problem because ST2003-L2 can be used as a basis of estimation for each commodity. But it will be a constraint for producing estimations at the provincial and national levels (since it will be impossible to directly sum up over kabupatens because different kabupatens will have different sets of commodities).

2. What frame will be used for that purpose? It is will not be possible to use the Population Census 2000 results since no detailed information by commodity is available. If the Agriculture Census 2003 itself (forms ST2003-L1 and ST2003-L2) is to be used, then the time consumed for processing by commodity would be more than one year, because of the following reasons:

   a. According to the present plan, there are already 20 fixed pre-coded commodities and several open spaces in the questionnaire for additional commodities. In some kabupatens the main "potential" commodities may not be one of the fixed 20 commodities (e.g. carrot, watermelon, salacia). The items to be entered manually will have to be coded by using the commodity code which should be prepared before processing activities begin. The ST2003-L2 complete count and processing activities will have to be finished at least two months before sample census field operations start because the data will be needed for sample allocation by commodity and to determine the main "potential" commodities.

b. If the field operation of the sample census is to be conducted one year after the full census, a new listing will be required because the ST2003-L1 and ST2003-L2 would be out of date. In this case, a new listing will be needed for estimation by using ratio estimators or other relevant methods.

III. VARIATION OF COMMODITY TYPE AND PRODUCTION

A. HORTICULTURE

A large number of agricultural commodities exist, especially for horticulture (which has more than fifty types), and they vary from one area to another. The present Agriculture Census 2003 draft questionnaire contains already 20 pre-coded horticultural commodities:

1. Guava
2. Orange
3. Mango
4. Pineapple
5. Papaya
6. Banana
7. Rambutan
8. Mustard greens
9. Long nourishing bean
10. Chili
11. Solanum melongena
12. Tomato
13. Cucumber
14. Spinach
15. Shallots
16. Rose
17. Orchid
18. Jasmine
19. Ginger
20. Saffron/curcuma

and more space is available for additional commodities.

According to the vegetable production figures in 1999 (see Table 2, where only four commodities are used for illustration), the production varies widely from one area to another. In Sumatera Utara, for example, the four commodities are the major crops produced in that province as well as in Jawa Barat and Jawa Tengah. But for Jambi the major commodity is only potatoes. DI Yogyakarta has only spring onions (production of while potatoes is very limited) and Kalimantan Tengah has almost nothing of the four commodities.

Table 2
Production of Four Types of Vegetables By Province (tons)

Province	Spring Onion	Potatoes	Cabbage	Mustard Greens
1. DI. Aceh	5143	3759	2753	627
2. Sumatera Utara	80593	204851	212370	46608
3.Sumatera Barat	16323	13929	99431	726
4. Riau	574	-	-	-
5. Jambi	885	30931	6874	-
6. Sumatera Selatan	486	55	1655	440
7. Bengkulu	2595	865	74992	13024
8. Lampung	2370	297	12004	392
9. DKI Jakarta	-	-	-	3
10. Jawa Barat.	107865	326661	447326	97900
11. Jawa Tengah	255973	148704	274339	34232
12. DI. Yogyakarta	37809	229	1951	76
13. Jawa Timur	196982	66610	164369	63138
14. Bali	15995	4537	58952	3691
15. NTB	47065	854	2913	23
16. NTT	11678	1052	807	286
17. Kal. Barat	47	2	61	-
18. Kal. Tengah	0	-	35	-
19. Kal. Selatan	56	-	8	-
20. Kal. Timur	91	-	343	-
21. Sulawesi Utara	1936	4588	1620	915
22. Sulawesi Tengah	7844	115	595	135
23. Sulawesi Selatan	9184	18127	33510	2428
24. Sul. Tenggara	1095	-	265	1
25. Maluku	505	2	102	-
26. Irian Jaya	1888	70	3430	291
Indonesia	804982	826238	1400705	264936

Source: Statistik Indonesia 1999

In such cases it is difficult to use each commodity per agricultural household as a sampling unit, because each province does not have the same commodities, so we have to define sampling units commodity by commodity before doing the sample selection. Calculating the sampling fraction will also be very complicated (one by one for each commodity in each kabupaten).

Fruit production also varies from one commodity to another and from one province to another. Table 3 provides an illustrative example of the variety in fruit production for only four commodities: mango, orange, banana and durian. The four commodities are the major commodities either at the provincial or the national level. Other commodities such as pineapple and salacia may or may not be major commodities in some provinces. In Jawa Timur three out of the four commodities in the table belongs to "potential" crops (mango, orange, and bananas). But in Sumatera Selatan bananas are the major crop with its production around 140 thousand tons, a significant amount compared to other provinces. Production of durian (10270 tons) may not be so important for Jawa Timur, but for Jambi (16442 tons) it is very important (it is the major crop).

Table 3
Production of Four Types of Fruits By Province (tons)

Province	Mangos	Oranges	Bananas	Durian
1.DI Aceh	11328	10854	30500	11410
2. Sumatera Utara	8451	54876	122008	21749
3. Sumatera Barat	3436	26823	86733	6600
4. Riau	4109	7884	41011	6547
5. Jambi	3454	2460	23655	16442
6. Sumatera Selatan	9243	4062	140186	9252
7. Bengkulu	1447	1629	10484	3065
8. Lampung	6471	1480	141117	6380
9. DKI Jakarta	10572	50	2817	379
10. Jawa Barat	95342	18068	898199	26450
11. Jawa Tengah	101303	9200	479956	13547
12. DI Yogyakarta	14920	1448	35687	1490
13 Jawa Timur	298249	121163	534458	10270
14. Bali	19225	41360	88138	2271
15. N TB	10707	2121	49428	750
16.NTT	44757	14372	112628	19
17. Kal. Barat	3202	108016	30448	9952
18. Kal. Tengah	8810	8497	11826	9840
19. Kal Selatan	4476	16988	22149	3757
20. Kal. Timur	9480	4295	22722	2368
21. Sulut	17724	1403	25437	2576
22. Sulteng	8152	2878	45690	3007
23. Sulsel	58490	177260	171216	6976
24. Sultenggara	20474	5426	20846	502
25. Maluku	21354	2073	10701	3747
26. Irian Jaya	588	717	2009	124
Indonesia	795764	645403	3160049	173470

Source: Statistik Indonesia 1999

The degree of heterogeneity of production (even when limited to four commodities) at the provincial level is not similar to that of vegetables. Vegetable production is more heterogeneous than fruit production, both between and within provinces. This affects the choice of sampling unit: should these units be limited only to major crops in each kabupaten, should estimates be made the provincial and/or national level?

B. PLANTATION

The same problem applies also to plantation commodities since here also we have many commodities with very heterogenous production between and within provinces. The ST2003-L2 questionnaire has the following fixed pre-coded commodities:

1. Oil Palm
2. Coconut
3. Rubber
4. Cane Sugar
5. Cashew
6. Coffee
7. Pepper
8. Cocoa
9. Tobacco
10. Tea
11. Clove

Other crops will be entered in the open spaces provided and will be coded one by one. Table 4 shows the planted area and production of small holders estate crops for some major crops.

Table 4
Planted Area and Production in 1999

Crops	Planted area (000 Ha)	Production (000 tons)
Oil palm	972,7	1326,6
Coconut	3558,8	2632,5
Rubber	2888,1	1182,4
Tea	65,0	34,1
Pepper	120,5	52,1

Some plantation crops are more concentrated in some provinces. For example, the major crop of pepper comes primaly from Lampung province and tea comes from Jawa Barat. The five commodities are destined mainly for export, so it is very important to have figures both at the national and regional levels.

C. LIVESTOCK AND POULTRY

The same problem of defining the sampling unit applies to the livestock and poultry sector as well, but the situation in this case is not as complicated as in the above cases because the number of varieties in livestock and poultry is limited. Only seven types of livestock exist: milk cows, cows, buffalos, horses, pigs, goats and sheep; and only four types of poultry: domestic hens, laying hens, broilers and ducks/manila ducks. The population of livestock by province is provided in Table 5 and that of poultry in Table 6.

Table 5
Livestock Population By Province in 1999 (000)

Province	Milk Cow	Cow	Buffalo	Horse	Goat	Sheep	Pig
1. DI Aceh	0,2	704,8	433,6	3,4	682,1	140,7	0.6
2.Sumatera Utara	7,8	257,0	265,6	5,6	725,2	165,9	913,0
3. Sumatera Barat	0,6	426,2	228,3	6,3	306,4	1,9	47,2
4. Riau	-	148,9	57,1	0,0	489,9	-	1092,2
5. Jambi	0,0	162,7	85,9	0,5	125,7	53,0	16,0
6. Sumsel	0,1	548,2	155,0	1,9	605,0	109,2	76,5
7. Bengkulu	-	69,8	27,9	0,1	65,6	3,2	1,2
8. Lampung	0,1	448,6	53,9	0,2	746,2	45,6	86,1
9. DKI Jakarta	4,3	-	1,0	0,3	8,3	2,9	-
10. Jawa Barat	79,5	151,5	357,4	11,5	1825,2	3464,7	19,1
11. Jawa Tengah	107,9	1231,2	196,7	17,0	3111,4	1838,2	103,4
12. Yogyakarta	3,8	201,2	7,3	0,9	270,0	76,0	5,8
13. Jawa Timur	129,2	3408,8	149,4	30,3	2255,9	1355,5	27,9
14. Bali	0,1	535,1	9,2	1,1	113,1	0,1	983,9
15. NTB	-	440,6	208,4	73,7	282,7	29,5	22,5
16. NTT	-	813,2	164,8	151,5	654,9	151,3	2287,3
17. Kal. Barat	0,1	177,8	8,1	-	122,2	0,0	958,7
18. Kal. Tengah	-	56,3	10,5	0,0	22,8	3,0	175,8
19. Kal. Selatan	0,1	147,0	37,3	1,5	64,6	3,5	6,5
20. Kal. Timur	0,1	41,5	14,7	0,1	62,3	1,4	92,5
21. Sulut	0,0	304,6	0,2	48,9	108,8	-	333,3
22. Sulteng	-	286,1	7,4	5,4	219,8	9,5	270,3
23. Sulsel	0,0	864,9	257,4	164,6	504,1	2,0	610,2
24. Sultenggara	-	295,8	11,5	6,1	124,4	0,4	24,3
25. Maluku	-	118,1	24,4	10,5	331,8	7,5	124,9
26. Irian Jaya	0,1	80,5	2,1	2,8	53,0	2,9	565,1
Indonesia	334,0	11920,4	2775,1	544,2	13881,4	7467,9	8848.3

Source: Statistik Indonesia 1999

Table 6
Poultry Population By Province in 1999 (000)

Province	Domestic Hen	Layer	Broiler	Duck
1. DI Aceh	24518,1	245,2	1064,2	3438,7
2. Sumatera Utara	20911,4	3804,8	41450,5	2254,5
3. Sumatera Barat	7487,7	1110,0	11381,4	1694,7
4. Riau	4401,1	925,7	13933,6	278,6
5. Jambi	4391,4	327,0	2300,0	723,8
6. Sumatera Selatan	16393,0	2319,0	5207,0	1302,0
7. Bengkulu	1927,9	21,9	2387,5	80,2
8. Lampung	15106,7	1907,5	3222,3	439,2
9. DKI Jakarta	133,8	-	579,9	70,8
10. Jawa Barat	29568,6	7525,7	106530,2	2938,0
11. Jawa Tengah	32213,6	5726,2	71860,9	3507,8
12. DI Yogyakarta	4960,9	1000,0	8750,0	210,0
13. Jawa Timur	37147,0	6791,3	100327,7	2286,3
14. Bali	5800,5	1017,1	1275,6	539,5
15. NTB	6368,8	316,0	-	388,3
16. NTT	9154,0	134,1	885,8	191,7
17. Kal. Barat	4200,5	2691,6	14997,2	420,8
18. Kal. Tengah	2770,4	19,3	1298,0	154,9
19. Kal. Selatan	4015,7	625,9	2812,1	1610,0
20. Kal. Timur	3249,3	387,0	4463,8	230,4
21. Sulut	2395,6	914,1	2267,2	426,0
22. Sulteng	2656,9	107,3	6105,7	151,8
23. Sulsel	15065,8	3555,4	11945,9	2384,9
24. Sultenggara	6628,4	35,1	980,0	279,1
25. Maluku	2341,0	99,3	935,3	135,7
26. Irian Jaya	1538,4	320,1	1433,7	116,5
Indonesia	265346,5	41926,6	418395,5	26254,4

Source: Statistik Indonesia 1999

IV. SAMPLING FRAME EVALUATION

A. Type of Sampling Unit

There are two type of sampling units :

1. **By type of commodity**: this is defined as a situation where one agricultural household plants more than one commodity. In this case the household belongs to more than one agricultural household (e.g. a household cultivating three crops --chilli, tomato and carrot– would belong to three households: one in chilli agriculture, one in tomato agriculture and one in carrot agriculture; and each crop is defined as a sampling unit. This method is called a uni-purpose design and it requires great care in sampling selection, particularly with regards to defining the type of commodity to be tabulated.

2. **By sub-sector**: this would cover, for example, food crops, horticulture, plantation, etc but not commodities. When the sample census method uses an integrated design by sub-sector, the agriculture household of the relevant sub-sector becomes the sampling unit. An alternative sampling selection method would use a systematic arrangement based on major crops. With this method it is easy to make estimation of data by crop and produce automatic tabulations. One advantage of systematic arrangements is the possibility of producing data at the regional level depending on "potential" crops in that area.

B. Agriculture Household By Commodity as a Sampling Unit

What steps need to be followed if agricultural household by commodity are chosen as a sampling unit?

1. First, cultivated crops by agriculture household must be used as a basis to define independently the sampling unit for each crop. The domains of study are agricultural households for each crop and the number of domains depends on the total number of commodities. The number of crops needs to be defined **before** sample selection in order to construct domains of study.

2. Secondly, the level of estimation needs to be defined. Are the same crops produced in each kabupaten or are there different crops depending on the "potential" crops in the kabupaten? What level of detail is needed for publication? What are the priority levels for

tabulations; is it by "potential" crops in each kabupaten (may be different from one kabupaten to another) or by "potential" crops in each province?

3. One drawback of this method is that it will not be possible to sum up the data from low levels to higher levels of aggregation. The data for the structure of agricultural production by sub-sector will be difficult to obtain. The only way to do it will be to ensure that every kabupaten covers the same list of basic commodities.

4. Another drawback is that this method will require a special listing questionnaire, especially for sampling selection, because the sample for each commodity will need to be selected independently. It will be possible for one agricultural household to belong to the sample of more than one crop (uni-purpose survey for each commodity).

C. Agriculture Household By Sub-Sector As a Sampling Unit

If the agricultural household by sub-sector is selected as a sampling unit, then these steps need to be followed:

1. First, agricultural households must be defined: e.g. food crop agriculture household, horticulture agriculture household, plantation agriculture household etc.) as a sampling unit. An integrated sampling design could be adopted and the sampling unit selection can be arranged systematically according to major crops cultivated for design efficiency. A combination design can also be made by using stratification (census blocks as a basis for stratification by category of concentration area and non-concentration area).

2. The level of estimation can then be derived directly from the kabupaten level to the provincial and national levels because agricultural households in the sub-sector are used as the sampling unit, thus making it possible to have only one inflation factor per sub-sector. Estimation for every crop can be done automatically during the tabulation phase and the relative standard error can also be produced by commodity. By using the arrangement by major crops in selecting the sample, combined with stratification, the relative standard error would be reduced.

3. The listing questionnaire will only be needed to categorize the sub-sector (agriculture household by major crop) and for sampling selection . The detailed questionnaire has to be prepared for all

agricultural household activities (not only for one crop). The structure of household activities for cross-sectional analysis will be easy to do with this method, as will the analysis of characteristics of agricultural household in general.

A general sampling procedure can be applied through stratification. Estimators, for example, will be calculated using ratio estimates:

$$\hat{Y}_{ks} = \sum_{h=1}^{H} \frac{y_{hks}}{x_{hks}} X_{hks} \qquad R_{hks} = \frac{y_{hks}}{x_{hks}}$$

where H = number of strata of kabupaten. The kabupaten is used as a domain of study or a sub-population.

X_h is the total number of agriculture households in sub sector s, such as horticulture agriculture households,

x_{hs} = total sample size of agriculture households in sub-sector s

y_{hs} = characteristics of agriculture households in sub-sector s from the sample

Estimation for the provincial and national levels using the simple method of summing up from each stratum:

$$\hat{Y}_{ps} = \sum_{k=1}^{K} \hat{Y}_{ks} \rightarrow$$ provincial level by crop automatically prepared in tabulation phase

$$\hat{Y}_{s} = \sum_{p=1}^{P} \hat{Y}_{ps} \rightarrow$$ national level by crop automatically prepared in tabulation phase

If we use every commodity cultivated by agricultural households as a sampling unit and "potential" crops depend on the kabupaten level, the provincial and national levels cannot be calculated by summing up from the kabupaten level.

$$\hat{Y}_{ksi} = \sum_{h=1}^{H} \frac{y_{hksi}}{x_{hksi}} X_{hksi}$$

X_{hsi} = total number of agriculture households of commodity i, in stratum h, sub-sector s. i depends on the potential crop in stratum h (e.g. kabupaten)

x_{hsi} = total sample size of agriculture households in sub-sector s

y_{hsi} = characteristics of agriculture households in sub-sector s and commodity i

Note that $\hat{Y}_p \neq \sum_{k=1}^{K} \hat{Y}_k$, i is different for every kabupaten

Tables 7 and 8 summarize the main implications of using the uni-purpose vs the integrated design.

Table7
Sampling Unit and Selection for Uni-Purpose Survey
(agriculture household by commodity as the sampling unit)

Name of head of household	.....	Agriculture household by commodity				
		.....*)	.....*)		.....*)	.....*)
A B C D		- -	-		- - -	- - -
		x_{hs1}	x_{hs2}			x_{hsm}

$\hat{Y}_{hs1}$ $\hat{Y}_{hs2}$ etc

Table 8
Sampling Unit and Selection for Integrated Design
(agriculture household by sub-sector as the sampling unit)

Name of head of household	.....	Main potential crop **)				
		.....	.....		.....	.....
A B C D		- -	- - -		- - - -	- - -
		x_{hs1}	x_{hs2}			x_{hsm}

*) Agriculture household can be belong to more than one column
**) Agriculture household can be belong to one column only
= mark used for sample selection arrangement

V. CONCLUSIONS

Statistical Law Number 16 of 1997 states that, in support of national development, BPS is responsible for producing a reliable, effective and efficient national statistics system covering

national, regional and cross-sectoral data. As a national program, the Agriculture Census needs to ensure that reliable national (and provincial) figures on the agricultural sector are produced. The above discussion highlighted the pros and cons of two methodologies that can be used in the census. The main conclusions are:

1. Adoption of an integrated design by sub-sector will make it easier to create reliable data at the national, regional and kabupaten levels. It will also make it easier to produce cross-commodity tabulations for every sub-sector. The implementation of sample selection in the field and the estimation procedures will be simpler, will reduce the bias in sample selection and the inflation factor to be used. Estimation of each sub-sector will use only the number of agriculture households of the relevant sub-sector.

2. Uni-purpose design by type of commodity in each sub-sector and with different treatment of potential crops in every kabupaten will produce only data per crop and it will not be possible to sum up the data at provincial or national levels. The sample selection will become more complicated, as each commodity will use a different sample selection procedure. The inflation factor will equal the total number of commodities.

3. Alternatively, one can use a combination of the two methods, for example, 1) to define the major crops needed to be published at the regional and national levels and then agricultural households cultivating those crops would be defined as the sampling unit at the national level, 2) to define the major crops at provincial level and 3) to define the main crops at the kabupaten level. In this case the same procedure for sampling selection can be used as mentioned before. The types of "potential" crops at the national, provincial, and kabupaten levels in some areas may or may not be the same.

4. Which method to choose will depend on the objective and the purpose of the census itself and the level of estimation.

APPENDIX A

PRE-CODED COMMODITIES IN ST2003-L2

HORTICULTURE

1. Guava	8. Mustard Greens	15.Shallots
2. Orange	9. Chili	16. Rose
3. Mango	10. Long nourishing bean	17. Orchid
4. Pine apple	11. Solanum Melongena	18. Jasmine
5. Papaya	12. Tomato	19. Ginger
6. Banana	13. Cucumber	20. Saffron
7. Rambutan	14. Spinach	

PLANTATION

1. Oil Palm	5.Cashew	9. Tobacco
2. Coconut	6. Coffee	10. Tea
3. Rubber	7. Pepper	11. Clove
4. Cane Sugar	8. Cocoa	

LIVE STOCK AND POULTRY

1. Milk Cow	5. Pig	1. Domestic Hens
2. Cow	6. Goat	2. Layer
3. Buffalo	7. Sheep	3. Broiler
4. Horse		4. Duck

Remark:

- ST2003-L1: List of Agriculture Households by sub sector
- ST2003-L2: Individual Questionnaire for each Agriculture Household

SAMPLE SELECTION METHOD FOR AGRICULTURE SAMPLE CENSUS 2003

Report # 37

by
Sri Budianti Sukmadi

July, 2001

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 269

II. TYPE OF HOUSEHOLD 269

III. CHOICE OF DESIGN 270
- **A. Uni-Purpose Design by Type of Commodity** 270
- **B. Integrated Design by Sub-sector** 272

IV SAMPLING UNIT AND LEVEL OF ESTIMATION 273
- **A. Sampling Unit** 275
- **B. Level of Estimation** 279
  - **1. Two Stage Sampling Design** 279
  - **2. Double Sampling** 279

IV CONCLUSIONS 280

I. INTRODUCTION

An earlier report[1] suggested that there were two types of sampling unit: one relying on commodities and one on sub-sectors. The sampling unit by type of commodity is used in a situation where an agricultural household plants one or more commodities. In this case, the same household may belong to more than one sampling unit. The sampling unit by sub-sector would cover households, for example, by food crop, horticulture, plantation etc., but not by commodity.

Three alternative designs for sample selection were provided in that report:

1. An integrated design by sub-sector. This design will make it easier to create reliable data at the national, provincial and kabupaten levels.

2. A uni-purpose design by type of commodity in each sub-sector and with different treatment of potential crops in every kabupaten: this will produce only data per crop and it will not be possible to sum up the data at the provincial or the national levels.

3. Or a combination of the above two methods, but only limited to a certain number of crops. This design will produce estimates until the provincial and national levels. The type of crops need to be defined first before sample selection is done in order to be able to construct domains of study.

The basic data used for frame construction and sample selection are forms ST2003-L1 and ST2003-L2 which contain information on agricultural households by type of commodity. These forms will produce data for households by type, agricultural households in general, and agricultural households by type of commodity.

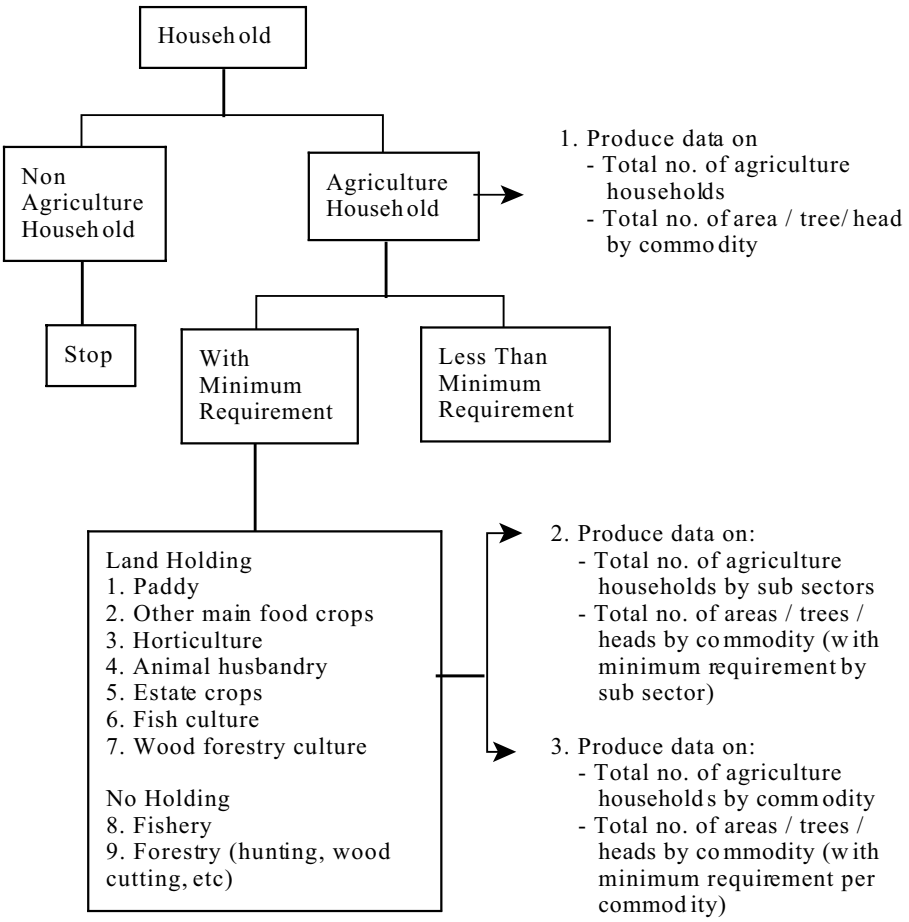
II. TYPE OF HOUSEHOLD

From the listing questionnaire (ST 2003-L1), benchmark data can be obtained: that is, data to determine agricultural structure and for the construction of the master frame used for sample selection. It is suggested to produce the general structure of household conditions (see Chart 1) in the tabulation plan as clearly as possible and to publish it.

---

[1] *Choice of Sampling Unit in the Agriculture Sample Census 2003*, STAT Project report #32, May 2001.

Chart 1: Structure of Agriculture Households



Note:
1. Population figure (Population)
2. Sub population 1 (Domain 1)
3. Sub population 2 (Domain 2)

III. CHOICE OF DESIGN

Any sampling unit could be produced from ST2003-L1 and ST2003-L2. The choice of unit will depend on the objective of the sample census/surveys, sample size requirement and estimation procedure.

A. Uni-Purpose Design by Type of Commodity

The uni-purpose design by type of commodity practically is very difficult to apply for the following reasons:

a. It is impossible to sum up data by commodity to the national and provincial levels.

b. Field operations become very complicated, especially allocation of enumerators for distributing the respondent evenly either by location or by commodity.

c. Listing costs are too high (each commodity needs a different listing and a different sample selection). Although it is possible to have one census block selected for more than one commodity, every commodity will still need a different form (listing questionnaire).

d. Processing of ST2003-L2 with no pre-coded commodities will be time consuming and may not be finished before the sample census period.

If this design is chosen then the projected timetable for sample selection will be:

- for paddy: first week of September 2003, i.e. after the listing activities (ST2003-L1 and ST2003-L2) are completed in August
- for non-paddy: April 2004, before the sample census activities (planned in May - June).

The projected cost of the sample census will be:

$$C = C_0 + \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{t=1}^{T} n_{skt} c_1 + \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{t=1}^{T} n_{skt} \overline{m}_{skt} c_2$$

where:
- C = total cost
- C_0 = overhead cost
- c_1 = unit cost of first stage sampling unit (e.g. transportation and listing)
- c_2 = unit cost of secondary sampling unit (e.g. interview)
- n_{skt} = total samples of first stage sampling unit (census block) in the sub sector s, kabupaten k, and commodity t
- $\overline{m}_{skt}$ = average total number of agriculture households per census block of commodity t, in kabupaten k, and sub sector s

The selection of n_{skt} has to wait for the results of ST2003-L2 by commodity and the total number of sample lists (which equals the total number of commodities that will be included in the sample census). The listing has to be conducted before the sample census period because otherwise the ST2003-L2 would already be out of date (the time lag is around one year) and there may be changes in commodity type produced by agricultural households. Particular attention has to be paid to the time reference for both the listing, the detailed questionnaire itself and the minimum requirement of data on agricultural household activities.

B. Integrated Design by Sub-sector

In the integrated design by sub-sector using one type of first stage sampling unit in each sub-sector (census block as integrated sample for each sub-sector), and then for selection of ultimate sampling units, two alternatives are available:

a. Two-stage sampling design by using the same ultimate sampling unit (agriculture household) for each sub-sector and for sampling selection or using main commodity/potential commodity as mentioned in the earlier report. This approach is referred to as a "multipurpose survey" and will produce detailed data for units selected from Domain 1 (as in Chart 1). The questionnaires have to be integrated for all commodities in each sub-sector. The benefits from this design include: 1) the possibility to sum up the data at the provincial and national levels, 2) field operations will be easier to conduct, 3) one listing is used for all sub-sectors, 4) the frame will be available a short time after complete count of the agricultural census and 5) a reduced cost.
The cost of the sample census can be expressed as:

$$C = C_0 + \sum_{s=1}^{S}\sum_{k=1}^{K} n_{sk} c_1 + \sum_{s=1}^{S}\sum_{k=1}^{K} n_{sk} \overline{m}_{sk} c_2$$

Compared with the uni-purpose design,

$$\sum_{s=1}^{S}\sum_{k=1}^{K} n_{sk} < \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{t=1}^{T} n_{skt}$$

That is, the total number of sampled census blocks using the integrated design (with only one listing for each sub-sector) will be smaller than the total number of sampled blocks using the uni-purpose design (with one listing for one commodity).

b. Two-phase sampling design (double sampling) using agricultural households by commodity as the ultimate sampling unit. The listing questionnaire used for sample selection is more or less the same as in the two-stage sampling design but the target sample and the sampling selection (which is a combination of uni-purpose and two-stage sampling designs) is determined by commodity. As mentioned before, the weakness of this approach is that only some potential crops can be used for estimates until the provincial and national levels. Monitoring of sample selection and matching results of interviews with the type of commodity selected in the sample will also become more complicated. The degree of matching is very important because it will affect the inflation factor for each commodity.

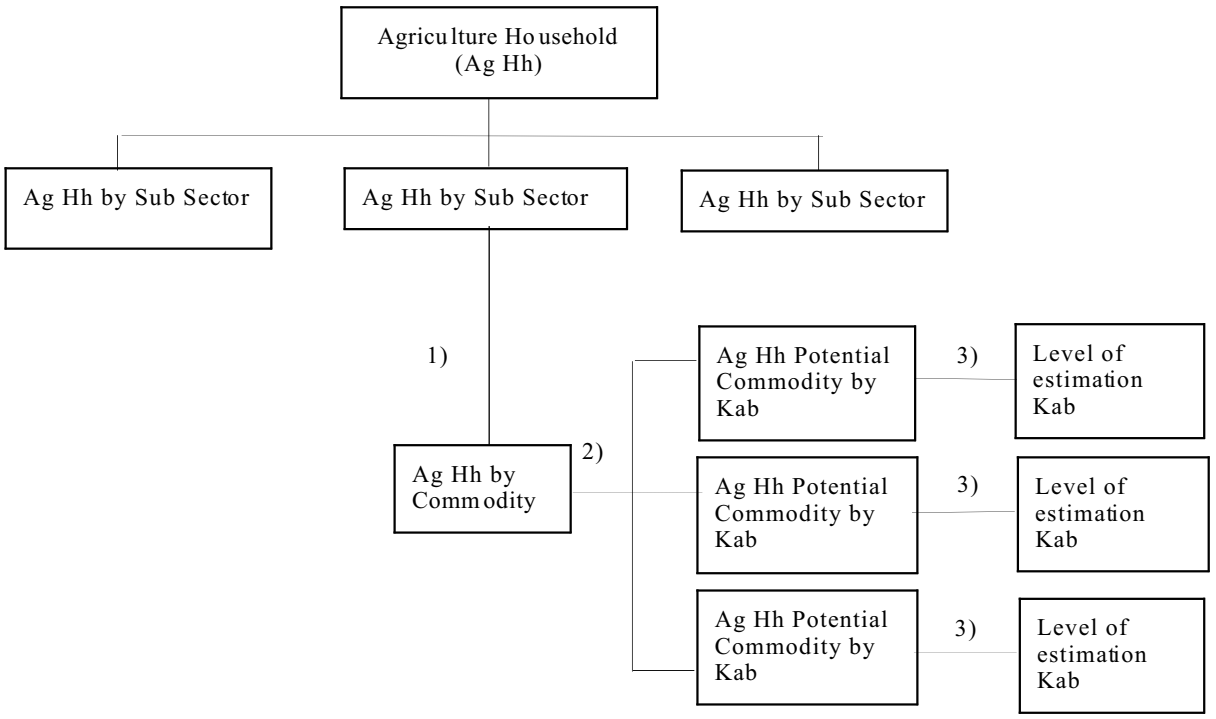
IV SAMPLING UNIT AND LEVEL OF ESTIMATION

The ultimate sampling unit has to be defined consistently with the unit of analysis and the format of the questionnaire. Based on interdepartmental discussions to date, the level of estimation target are:

- National potential commodities will be tabulated at the national, provincial and kabupaten levels.
- Special potential commodities for kabupatens will be tabulated at the kabupaten level only.

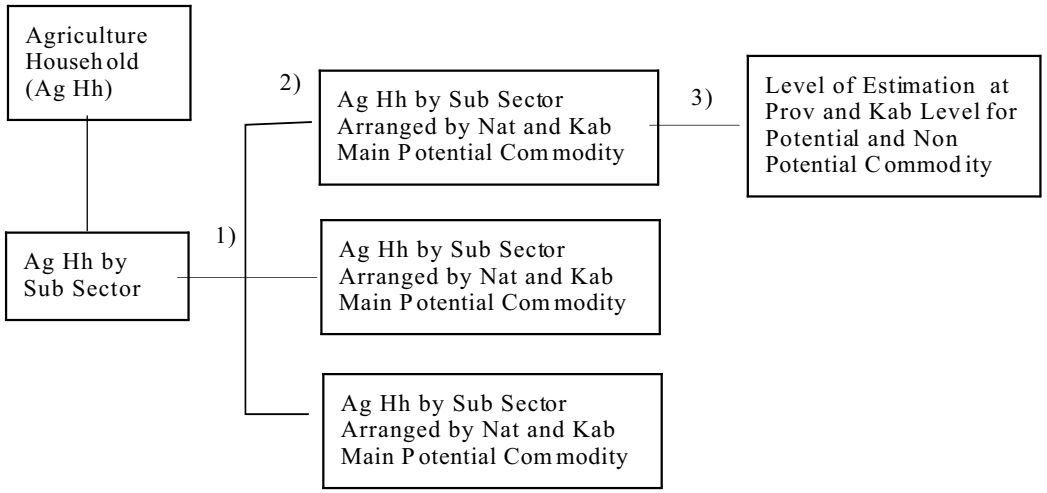
The type of potential commodity will be determined using area planted, the share in gross domestic product or other criteria. Besides the level of estimation, it is also necessary to evaluate the structure of each sub-sector, especially in horticulture and plantation crops since there are multiple crops within one area. With regards to the level of estimation, the multi-purpose design with two-stage sampling is the most appropriate to use. A comparison of the designs and their implications are shown in Charts 2 through 4.

Chart 2: Uni-Purpose Design



1) Selection of census block at the kabupaten level from ST2003-L2 (after processing ST2003-L2 is finished) by commodity
2) Fixed commodities at the kabupaten level vary from one kabupaten to another.
3) Produce data at the kabupaten level

Chart 3: Multi-Purpose Two-Stage Sampling Design

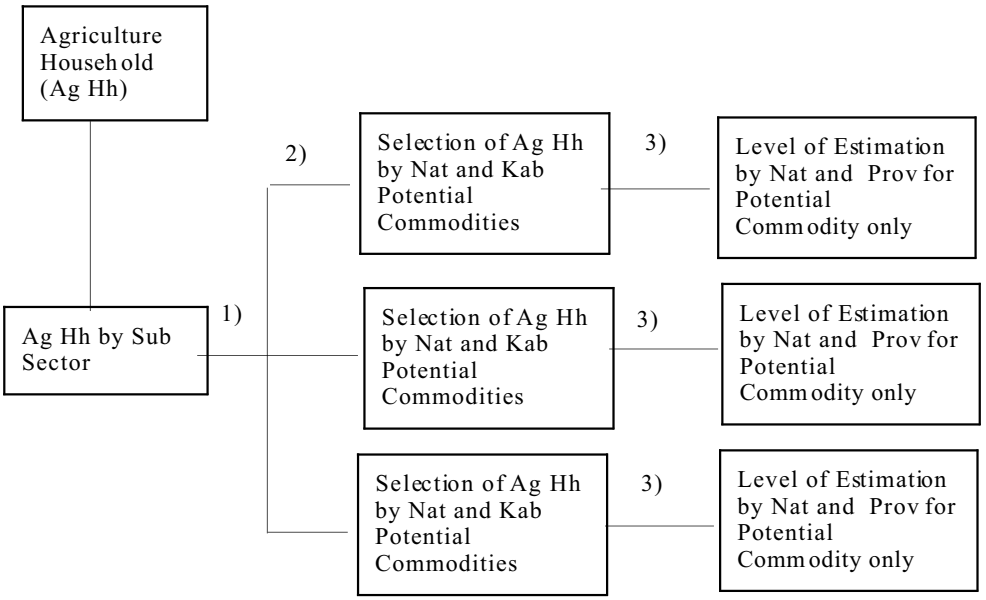


1) Sample selection of census blocks by sub-sector at the kabupaten level

2) Sample selection of agriculture household at the kabupaten level by sub-sector and systematic arrangement by main potential commodity (main potential commodity for each agriculture household)

3) - Possibility to produce data for all commodities (potential and non-potential) by calculating the sample size and standard error at national, provincial and kabupaten levels and designing the questionnaire for the agriculture household sector as a whole.
- The questionnaire is divided into a core and a module. The core questionnaire is used for collecting general information on sub-sector with various commodities while the module questionnaire is used for collecting more detailed information on potential commodities. In this case, the population will be all agriculture households by sub-sector at the national, provincial and kabupaten levels.
- Possibility to evaluate activities within sub-sectors, especially for mixed crops and livestock/poultry. In this case, it might be difficult to specify exactly the activities by commodity for one agriculture household.
- If necessary, additional samples for kabupaten potential commodities can be selected and alternative sample selection techniques can be used. For example, two or more independent sub-samples can be selected (depending on the number of potential commodities in the kabupaten, e.g. 2 or 3 for systematic arrangement with different potential crops in each kabupaten).

Chart 4 : Multipurpose Double Sampling



1) Sample selection of census block by sub-sector at the kabupaten level

2) - Agriculture household by potential commodity as the ultimate sampling unit will be selected by integrated census block (unlike selection of census blocks by commodity as in Chart 2).
- The sample of agriculture households, by national and kabupaten potential commodities, are determined first at the kabupaten level and the selection of sample is then done independently for each potential commodity (target by commodity).

3) - Produce data on agriculture households by commodity at the national level for national potential commodities and at the kabupaten level for kabupaten potential commodities.
- No evaluation by sub-sector (e.g. activities within sub-sector), especially for mixed crops and livestock/poultry. But we need to arrange variables in the questionnaire into core and module variables to know the structure of agriculture households compared with the commodity agriculture households (general information and detailed commodity information).

A. Sampling Unit

Comparing the three methods above, the uni-purpose design will be the most inefficient because it will require big samples and the level of estimation will only cover lower level administrative units. Using the multi-purpose two-stage sampling design is recommended because it has the simplest sample selection, estimation procedure and field operations. In the uni-purpose design, controlling sampling selection requires attention commodity by commodity. Table 1 below compares the total number of sampling units by kabupaten and type of sampling unit selected under the two-stage sampling and the double sampling designs.

Table 1
Sampling Unit Under the Two-Stage Sampling and Double Sampling Designs

Type of sampling unit	Two stage sampling	Double Sampling
1. First stage sampling unit	Census Block	Census Block
2. Second stage sampling unit	Agriculture household by sub-sectors (9 sub-sectors)	Agriculture household by commodity 1. Paddy : 1 2. Other main food crops : 6+2 3. Horticulture :20+2 4. Animal husbandry :11+2 5. Estate crops :12+2 6. Fish culture : 2+2 7. Wood forestry culture : 1 8. Fishery : 2 9. Forestry(hunting, wood cutting etc) : 1 Total : 56+10
3. Sample selection by kabupaten	9 sub-sectors	Maximum 66 types, total types per kabupaten depend on the national potential and kabupaten potential, varies from one kabupaten to another.

In the two-stage sampling design, one can arrange sample selection of agriculture households by type of main potential commodity in kabupatens with several potential commodities. The probability selection of potential commodities is high (for example by area/production). Tables 2 and 3 show the distribution of vegetables and fruits production in Jambi published by BPS in “Jambi in Figures” every year as an example.

Table 2
Production of Vegetables by Commodity in Jambi 1999 (Tons)

Commodity	Kabupaten/Kota						
	Kerinci	Bungo Tebo	Sarola-ngun bango	Batang hari	Tanjung jabung	Kodya Jambi	Total
1. Onion* **	781	-	35	-	-	-	816
2. Garlic	28	-	-	-	-	-	28
3. Shallot	347	-	130	-	-	-	477
4. Potatoes* **	23038	-	3136	-	-	-	26174
5. Cabbage* **	4816	-	349	-	-	-	5165
6.Mustrad green**	181	-	142	-	-	383	705
7.Carrots **	60	-	13	-	-		73
8.Radish	96	-	-	-	-		96
9. Red Kidney	530	-	-	-	10		540
10.Longnourish-ing bean*	111	896	602	717	564	367	3258
11.Chilli*	2869	1992	1939	1471	849	345	9465
12Tomatoes*	430	326	393	199	290	205	1843
13.Eggplants	174	887	255	253	374	92	2036
14.Greenbeans	1124	12	426	716	205	679	3162
15.Cucumber*	136	550	278	146	263	233	1605
16.Labu siam	10	1	-	-	19	16	46
17.Kangkung	74	408	324	61	200	121	1087
18.Spinach*	80	172	138	10	117	112	629

Source: Jambi in Figures, 1999
*) Pre-coded commodities in ST2003-L2
**) Commodities in Statistics Indonesia Publication

Table 3
Production of Fruits by Commodity in Jambi 1999

Commodity	Kabupaten/Kota						
	Kerinci	B. Tebo	S Bango	B.Hari	Jabung	K Jambi	Total
1. Avocado **	978	44	310	40	9	12	1393
2. Manggo * **	114	1674	1547	2974	552	92	6952
3. Rambutan* **	379	5682	1314	3465	838	290	11969
4. Duku **	33	1801	637	17374	2794	4	22642
5. Orange * **	404	311	163	350	331	92	1651
6. Durian * **	977	6177	2136	3595	470	54	12809
7. Jambu air	222	340	324	273	386	36	1580
8. Sapodilla **	5	282	139	514	171	49	1159
9. Papaya * **	2175	6135	2381	1279	835	1883	14688
10.Banana * **	677	7494	1298	5659	3928	454	19511
11. Pineapple* **	103	8251	109	11408	318	1643	21833

Source: Jambi in Figures, 1999
*) Pre-coded commodities in ST2003-L2
**) Commodities in Statistics Indonesia Publication

A small province like Jambi, for example, needs data on 18 vegetable commodities and 11 fruit commodities. Not all these commodities are included in the BPS national publication. A comparison of commodities included in “Jambi in Figures”, the BPS “Statistics Indonesia” publication and ST2003-L2 for vegetables is presented in Table 4, and for fruits in Table 5.

Table 4
Comparison of Vegetables in ST2003-L2, Statistics Indonesia and Jambi in Figure

Commodity	ST2003-L2	Statistics Indonesia	Jambi in Figure
1. Potato 2. Long Nourishing Bean 3. Chilli 4. Cabbage 5. Tomato			
6. Cucumber 7. Spinach 8. Onion 9. Rose 10. Orchard			
11. Jasmine 12. Ginger 13. Turmeric 14. Garlic 15. Shallot			
16. Mustard Greens 17. Carrots 18. Radish 19. Red Kidney 20. Eggplant			
21. Green Bean 22. Labu Siam 23. Kangkung			

Table 5
Comparison of Fruits in ST2003-L2, Statistics Indonesia and Jambi in Figure

Commodity	ST2003-L2	Statistics Indonesia	Jambi in Figure
1. Durian 2. Orange 3. Mangoes 4. Pineapple 5. Papaya			
6. Banana 7. Rambutan 8. Avocado 9. Duku 10. Jambu air			
11.Sapodilla 12. Salacia 13. Jambu biji (Guava)			

For vegetables, note that:

a. Only 8 out of 18 commodities of Jambi mentioned in “Jambi in Figures” publication are included in ST2003-L2, and only 5 are in “Statistics Indonesia”.
b. Only 3 commodities are covered in all three sources.

For fruits,

a. Only 7 out of 11 commodities of Jambi mentioned in "Jambi in Figures" publication are included in ST2003-L2, but 10 are in "Statistics Indonesia".
b. Seven commodities are covered in all three sources.

Variation in commodity type is high among kabupatens and provinces. Thus each kabupaten will have to determine its potential commodity for comparison with the national potential commodities.

B. Level of Estimation

As mentioned above, the level of estimation needs to be done at the national, provincial and kabupaten levels for potential crops in each area, but for determining potential commodities/crops and limit the number of commodities covered, only national and kabupaten levels need to be done. During the tabulation stage, one can measure provincial standard errors to determined whether commodities meet requirements for provincial estimation.

Preparation of commodity types by sub-sector to be tabulated is necessary before sample selection, since two types of commodity groups will be selected:

- potential commodities at the kabupaten level which are rare (i.e. in only in small parts of the kabupaten).
- potential commodities at the kabupaten level which are spread throughout the kabupaten.

The second group will present no problem in sample selection if we use census blocks with agriculture household sub-sector as a sampling unit. But the first group may need additional samples in the kabupaten.

1. Two Stage Sampling Design

In the case of a two-stage sampling design, the sampling unit by sub-sector where estimation is done at the kabupaten, provincial and national levels, no problem exists because the total sample is big enough at the provincial and national levels, and at the kabupaten level additional samples may be selected. Another advantage of this design is that, besides data for potential commodities, it will be possible to tabulate other commodities.

2. Double Sampling

In the case of double sampling, the potential commodities will be determined before sample selection and the sample list will vary from one kabupaten to another. The data to be produced will only cover potential commodities.

IV CONCLUSIONS

1. Using the uni-purpose sampling design is not efficient when one looks at the potential complications in sample selection, estimation procedures and cost.

2. Double sampling is still complicated and does provide flexibility in allowing estimation of all commodities and in comparing the structure of commodities within a sub-sector. The two-stage sampling design is simpler and allows the production of data on potential commodities and other commodities.

3. A core and a module questionnaire needs to be produced, and different questionnaires needs to be designed for different sub-sectors.

4. Produce a manual documenting sample selection, sample list, method of estimation and monitoring system for kabupaten, province and national level in order to ensure that at consistent methodology is applied.

5. Produce a list of potential commodities by province and kabupaten and highlight national potential commodities. It is important to ensure completeness of estimations and tabulations.

SAMPLE DESIGN FOR THE SECOND PILOT PLANNED UNDER THE AGRICULTURE CENSUS 2003

Report # 43

by
Sri Budianti Sukmadi

November, 2001

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 283

II. SAMPLING DESIGN FOR URBAN AREAS 284

III. SAMPLING DESIGN FOR SUB-SECTORS 285
- **A. Paddy** 285
  - **1. One-Stage Design** 285
  - **2. Two-Stage Design** 286
- **B. Other Main Food Crops** 286
- **C. Horticulture** 286

IV. PRELIMINARY RESULTS 287
- **A. General Frame** 287
- **B. Sub-sector Frame** 292
- **C. Second Pilot** 294

V. CONCLUSIONS 297

I. INTRODUCTION

Two earlier reports[1] discussed the choice of sampling unit and selection method for the Agriculture Census 2003 sample census. The main conclusion was not to use uni-purpose design by type of commodity for each sub-sector with different treatment of potential crops in different kabupatens. The reason being that such a design would produce only data per crop which could not be summed up at the provincial or national level. Detailed characteristics for the sample census will be divided into 9 sub-sectors.

1. Paddy
2. Other main fixed crops
3. Horticulture
4. Animal husbandry
5. Estate crops
6. Fish culture
7. Wood forestry culture
8. Fishing
9. Forestry (hunting, wood cutting, etc)

Three pilot studies are planned:

1. The first was conducted in Kabupaten Subang, Jawa Barat, in May 2003. It focused on updating Census Block (CB) maps, the listing questionnaire and sub-sector questionnaires.
2. The second pilot was conducted in Kabupaten Sukabumi in August 2001. It involved a complete count in Kecamatan Cisaat rural areas, a sample census in urban areas by using probability sample and a purposive sample for areas outside Kecamatan Cisaat (Kecamatan Pelabuhan Ratu) which focused on updating CB maps, listing questionnaire and sub-sector questionnaires. The pilot study conducted in Kecamatan Pelabuhan Ratu covered only a test of the questionnaire.
3. The third pilot will be conducted in Kabupaten Takalar, Gowa, and Maros, in the province of Sulawesi Selatan. It will cover one kecamatan in each kabupaten, involving a complete count with ST2003-L1 and ST2003-L2 (listing questionnaire) and a sample for the detailed questionnaire ST2003-S. The ST2003-Podes (Village Potential) will be included in this pilot.

This report discusses the sample design for the second pilot.

---

[1] *Choice of Sampling Unit in the Agriculture Sample Census 2003*, STAT Project Report #32, May 2001 and *Sample Selection Method for Agricultural Sample Census 2003*, STAT Project Report #37, July 2001.

II. SAMPLING DESIGN FOR URBAN AREAS

Because of budget limitations, the pilot study in Kecamatan Cisaat was conducted on a sample basis, with 50 of the 254 CB's selected following a stratification procedure. Table 1 provides the frequency distribution of CB's by group of agricultural households in the pilot study.

Table 1
Total Number of CB's by Group of Agricultural Households in Kecamatan Cisaat Urban Areas

Number of Agricultural Households	Number of CB's	Percentage
0	13	5.1
1 - 9	134	52.7
10 - 19	49	19.3
20 - 29	27	10.6
30 - 39	20	7.9
40 - 49	3	1.2
50 - 59	3	1.2
60 - 69	2	0.8
70 - 79	1	0.4
80 -	2	0.8
Total	254	100

Four strata were constructed from the frame:

1. Strata with no agricultural household
2. Strata with 1 - 19 agricultural households
3. Strata with 20 - 29 agricultural households
4. Strata with 30 or more agricultural households.

Stratification used the Dellinius method. Strata number 3 and 4 were considered "potential" strata with the total number of agricultural households in one CB of 20 or more (around 22.9 percent of all CB's).

Because a complete count is too costly, this stratification provides a useful and reliable alternative. It may be worth noting that most countries do the same: although they nominally call it an agricultural "census", the operation is often conducted on a sample basis, with a big sample. Complete count in an agricultural census, especially in urban areas, is not effective because the total number of agricultural households is not big.

After processing the pilot study ST2003-L1 and ST2003-L2, alternatives designs will be evaluated to compare the efficiency of design, particularly for basic information:

- using the complete count method in both urban and rural areas.
- using the complete count method in rural areas and a sample in urban areas.
- using the sample census method in both urban and rural areas.

National Sakernas figures for the year 2000 show that 24.2 percent of the population aged 15 year and older were self-employed or as own-account workers, self-employed assisted by family member(s) or temporary employee(s) or employee(s) in an agricultural activity. This figure is consistent with the percentage of agricultural households in Kecamatan Cisaat (22.9 percent of CB's are agricultural household CB's). Such consistency supports the idea of sampling for the agricultural census if budgets are limited.

III. SAMPLING DESIGN FOR SUB-SECTORS

Sampling design in the pilot study covers only the sub-sectors of paddy, other main foods crops and horticulture.

A. Paddy

Two alternatives for sampling design were applied in the pilot study: the one-stage and the two-stage designs. Evaluating these alternatives for this sub-sector is very important, especially in order to determine which one would provide less difficulties during implementation.

1. One-Stage Design

The procedure for this design includes the following steps:

1. All CB's in Kecamatan Cisaat rural areas were listed by questionnaire ST2003-L1 and ST2003-L2. Three categories of paddy fields were considered: wet land with irrigation, wet land with no irrigation, and dry land. A fraction (interval) for sampling selection of paddy agricultural households needs to be determined. A rough approximation can be made from BPS's routine surveys by using the area of paddy fields divided by average area of paddy per agricultural household (4.9). That is, about one-fifth of paddy agricultural households have to be selected.
2. The estimated number of paddy agricultural households using the above interval was around 320 (which became the target sample for the pilot). However, during implementation the total number of paddy agricultural households was found to be almost twice as high (579 households). This problem arose because data on paddy areas were not accurate, mainly because of weaknesses in the register. Applying this procedure was important, especially because of its budget implication: the sampling fraction calculated *after* field operation may be very different from the one estimated *before* field operation. That is why this procedure cannot be used when conducting large-scale activities such as censuses or large sample surveys.

2. Two-Stage Design

Two methods were applied for evaluating this design: proportion allocation to type of irrigation and fixed samples for each type of irrigation to compare which one is the best. The fixed sample allocation by type of irrigation turned out to be too complicated, especially for targeting the sample size for each type of irrigation, because we have to wait until the processing of ST2003-L1 is finished while the time lag between the listing activities and the implementation of sample census for detailed paddy agriculture households is only one month.

The third pilot in Sulawesi Selatan rural areas will use the two-stage sampling design with paddy agricultural households as the ultimate sampling unit, but will still use the type of irrigation for statistical arrangement only and CB with paddy agricultural households as the primary sampling unit. Hopefully, coefficients of variation and relative standard errors from this pilot can be used to estimate the minimal sample sizes.

B. Other Main Food Crops

Application of the two-stage sampling design will be evaluated for this sub-sector. In the first stage, CB's are selected with probability proportional to size measured by the total number of other main food crop agricultural households derived from listing results (ST2003-L1 and ST2003-L2). The procedure in the pilot study basically follows suggestions mentioned in the earlier report, that is by using two-stage sampling with multipurpose design in each sub-sector and the selection of ultimate sampling units by using double sampling (targeted samples for each commodity) and proportional allocation.

There are 3 alternatives that will be evaluated for other main food crop agricultural households.

- Agricultural household by crop/commodity is treated as the ultimate sampling unit and a target sample is determined for each commodity.
- Agricultural household of main food crop is treated as the ultimate sampling unit and agricultural households are selected by arranging first with main commodity of food crops and targeting the samples by commodity.
- Agricultural household of main food crop is treated as the ultimate sampling unit and agricultural households are selected by arranging first with main commodity of food crops and selecting ultimate sampling unit systematically (automatically it will be a proportional allocation and no targeting for each commodity will be necessary).

C. Horticulture

The two-stage sampling design will be evaluated by using the three alternatives used for the main food crops. However, there are special considerations to be given for horticulture. Because of budget constraints, not all commodities are pre-coded in ST2003-L2 (only 20

commodities are), and only six commodities are included in the pilot and are divided into two categories:

- National potential commodities: papaya, rambutan, banana, and manggis.
- Local potential commodities: mustard green, spring onion.

These six commodities will be evaluated for the detailed questions. Targeting samples by commodity and agricultural household may present more problems in horticulture compared to other main food crops because the types of commodity has greater variation in horticulture.

IV. PRELIMINARY RESULTS

A. General Frame

Having a proper master frame is necessary for choosing a proper sampling method and an adequate sample. That is why one needs to take particular care in evaluating the frame to be used. By using data from SP2000 and the pilot study, this section provides a number of tabulations which were obtained with the help of the Methodology Directorate.

In the pilot study in Kecamatan Cisaat, the total number of agricultural household CB's in urban areas was around 22.9 percent. That can be compared with CB's from Population Census 2000 in bigger areas with less urban concentrations, such as in the province of Sulawesi Tenggara. Kecamatan Cisaat is predominantly agricultural, although it has some urban parts. Table 2 shows the number of CB's by group of agricultural households in Sulawesi Tenggara divided into urban and rural areas.

Given available budget, data in the table can help determine strata for groups of agricultural households and the sampling fraction for first-stage selection.

Table 2
Number of CB's by Group of Agricultural Households
Province: Sulawesi Tenggara

Number of Agricultural Households	Urban		Rural		Total	
	CB's	Percent	CB's	Percent	CB's	Percent
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0	445	73.8	473	14.9	918	24.3
1 - 4	27	4.5	15	0.5	42	1.1
5 - 9	31	5.1	70	2.2	101	2.7
10 - 19	58	9.6	1040	32.8	1098	29.1
20 - 39	41	6.8	1530	48.2	1571	41.6
40 - 59	1	0.2	40	1.3	41	1.1
60 - 99	0	0.0	3	0.1	3	0.1
100 and over	0	0.0	0	0.0	0	0.0
Total	603	100	3171	100	3774	100

Source: SP2000-L2, Preliminary Figure

Table 3
Number of CB's by Group of Agricultural Households
Province Sulawesi Tenggara

Number of Agricultural Households	Kabupaten				Kota	Total
	Buton	Muna	Kendari	Kolaka	Kendari	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0	236 (22.9)	106 (17.3)	158 (14.5)	111 (16.8)	307 (80.2)	918 (24.3)
1 - 4	9 (0.9)	4 (0.7)	5 (0.5)	7 (1.1)	17 (4.4)	42 (1.1)
5 - 9	20 (1.9)	10 (1.6)	42 (3.9)	15 (2.3)	14 (3.7)	101 (2.7)
10 - 19	248 (24.0)	217 (35.4)	433 (39.9)	172 (26.0)	28 (7.3)	1098 (29.1)
20 - 39	486 (47.0)	274 (44.8)	446 (41.1)	348 (52.6)	17 (4.4)	1571 (41.6)
40 - 59	31 (3.0)	1 (0.2)	1 (0.1)	8 (1.2)	0 (0.0)	41 (1.1)
60 - 99	3 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.1)
100 and over	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total	1033 (100.0)	612 (100.0)	1085 (100.0)	661 (100.0)	383 (100.0)	3774 (100.0)

Source: SP2000-L2 Preliminary Figure

The distribution of the total number of CB's for that province by kabupaten is provided in Table 3.

As was mentioned earlier, three alternative methods can be adopted:

1. Complete count in both rural and urban areas
2. Complete count in rural areas and a sample in urban areas with CB as a sampling unit. The number of CB's in the sample depends on the coefficient of variation of characteristics and the degree of precision desired. The coefficient of variation can be estimated from pilot study data and also from SP2000-L2 limited to the total number of agricultural households.
3. Sample in either rural or urban areas by using stratification, for example by group of agricultural households from SP2000-L2 data. Agriculture Census in concentration/potential CB's having agricultural activity can be conducted on a complete count and the rest of CB's are on a sample basis.

Some simulations were done with data from Sulawesi Tenggara:

- First: using complete count method the all CB's (3774 CB's). Results will replicate the distribution in Table 2.
- Second: complete count for rural areas and a sample for urban areas. Table 4 shows the result of a 50% sample in some strata and 100 in others.

Table 4
Number of CB's for Second Alternative

Number of Agricultural Households	Rural	Urban			
	Complete Count	CB's	Strata	Sample (%)	Sampled CB's
(1)	(2)	(3)	(4)	(5)	(6)
0 1 - 4 5 - 9	473 15 70	445 27 31	I	50	252
10 - 19 20 - 39	1040 1530	58 41	II	100	99
40 - 59 60 - 99 100 and over	40 3 0	1 0 0	III	100	1
Total	3171	603		58	352

- Third: sample for both rural and urban areas but with different fractions (i.e. two simulations). Tables 5.1 and 5.2 show the results of using a third and fourth simulation respectively.

Table 5.1.
Number of CB's Third Simulation Applied to the Third Alternative

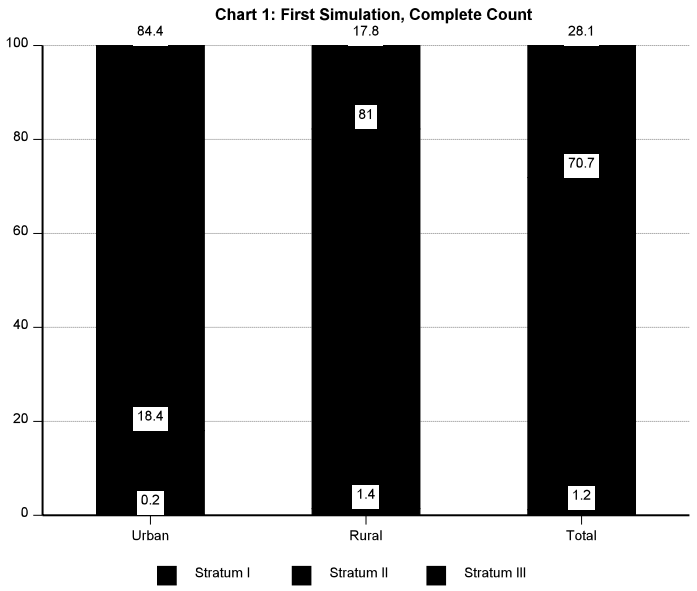
Number of Agriculture Households	CB's		Strata	%	Sampled CB's		Total
	Urban	Rural			Urban	Rural	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0 1 - 4 5 - 9	445 27 31	473 15 70	I	50	252	279	531
10 - 19 20 - 39	58 41	1040 1530	II	100	99	2570	2669
40 - 59 60 - 99 100 and over	1 0 0	40 3 0	III	100	1	43	44
Total	603	3171			352	2892	3244 (85%)

Table 5.2.
Number of CB's Fourth Simulation Applied to the Third Alternative

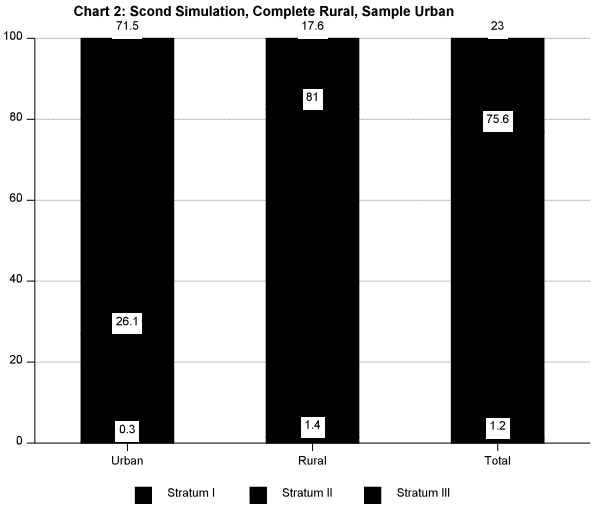
Number of Agriculture Households	CB's		Strata	%	Sampled CB's		Total
	Urban	Rural			Urban	Rural	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0 1 - 4 5 - 9	445 27 31	473 15 70	I	25	126	140	266
10 - 19 20 - 39	58 41	1040 1530	II	50	50	1285	1335
40 - 59 60 - 99 100 and over	1 0 0	40 3 0	III	100	1	43	44
Total	603	3171			177	1468	1645 (44%)

Charts 1-4 provide graphical illustration of results using these alternatives.

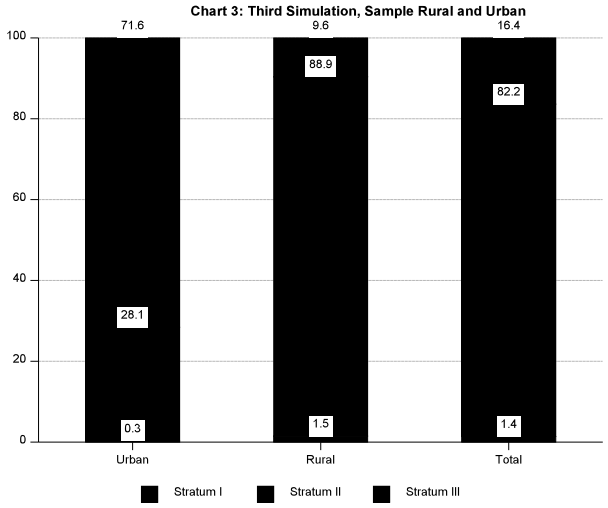
Number of CB's
Urban : 603
Rural : 3171
Total : 3774



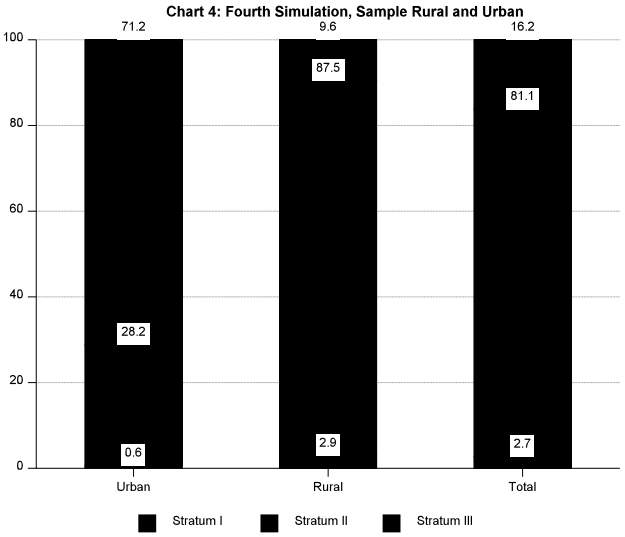
Number of CB's
Urban : 352
Rural : 3171
Total : 3523



Number of CB's
Urban : 352
Rural : 2892
Total : 3244



Number of CB's
Urban : 177
Rural : 1468
Total : 1645



Of course, one can have many more types of simulations depending on conditions prevailing in different provinces. Since the sample size is large, using stratified systematic sampling design can be applied by arranging CB's according to the total number of agricultural households and the procedure of estimation become simple:

$$\hat{Y}_h = \frac{N_h}{n_h}\sum_{i=1}^{n_h} y_{hi} \qquad \text{so} \qquad \hat{Y} = \sum_{h=1}^{H} \hat{Y}_h$$

The variance can be obtained with stratified simple random sampling by regional, province, and national figure. Other sampling method also can be applied by probability to size with agriculture households as size.

B. Sub-sector Frame

From 3774 CB's in Table 2 we can produce more detailed tables to know the total number of CB's by sub-sector and also the coefficient of variation (CV) of the total number of agricultural households as a whole and by sub-sector at the CB level. Tables 6 and 7 provide such information. Table 6 shows the total number of agricultural households with the same concept of main activity as in SP2000 in the selection of CB's for the sample census. If we use the SP2000 CB frame we have to be careful when we deal with particular sectors, for example animal husbandry. Although some agricultural households have livestock/poultry but not as a main activity, they are still included as animal husbandry agricultural households because they may have enough livestock/poultry to qualify for the minimum requirement according to the definition.

Table 6
Total Number of Agricultural Household CB's by Sub-sector
Province: Sulawesi Tenggara

Agricultural Household	Number of CB's		Total
	With No Agricultural Households	With Agricultural Households	
(1)	(2)	(3)	(4)
1. General	918	2856 (75.7%)	3774
2. Sub-sector			
a. Food crop	506	2350 (82.3%)	2856
b. Horticulture	1196	1660 (58.1%)	2856
c. Plantation	1918	938 (32.8%)	2856
d. Fishery	2706	150 (5.3%)	2856
e. Animal husbandry	2026	830 (29.1%)	2856

Table 6 shows total number of CB's with animal husbandry agricultural households is only 29.1 percent. Their CV is 6.93 (see Table 7). The CV of CB's with horticulture agricultural households is 1.55 and from the Kecamatan Cisaat pilot study is 1.68 (Table 7). These figures are in the same order of magnitude. But when taken by type of commodity, CV's vary, sometimes significantly. For example the CV for papaya is 2.97, spring onions 4.05 and mustard green 2.58. Using the CB as a primary sampling unit is a good method, but we have to take great care in using the master frame: it depends on which variable we choose.

Table 7
Coefficient of Variation (CV) of Total Number of Agricultural Households
Province: Sulawesi Tenggara

Kabupaten/Kota	General	Sub-sector				
		Food crops	Horti-culture	Plantation	Fishery	Animal husbandry
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Kabupaten						
1. Buton	0.38	0.79	2.53	4.46	1.82	6.40
2. Muna	0.30	0.50	3.51	5.09	2.92	11.20
3. Kendari	0.32	0.86	1.43	2.92	3.36	7.70
4. Kolaka	0.33	1.85	0.62	4.59	3.63	6.91
Kota						
1. Kendari	0.65	1.48	2.11	1.84	1.81	2.60
Total	0.36	0.91	1.55	3.7	2.64	6.93

Two alternative frames can be used for the Agriculture Census: SP2000-L2 with agriculture as the main activity of household members of household and ST2003 itself (using the results of ST2003-L1 and ST2003-L2).

1. **SP2000-L2**: a frame constructed from SP2000-L2 will be useful for selection CB's mainly if the budget is severely constrained. According to our preliminary evaluation, if there are agricultural household CB's in SP2000-L2, there will also be agricultural household CB's in the pilot study. But the selection of agricultural sub-sector CB's becomes crucial, because one agricultural household may have more than one sub-sectoral activity (the SP2000-L2 only covers the main activity). The CB's with food crops and horticulture agricultural households are not as crucial as animal husbandry used for agriculture census sample. For example the CB's with food crop agriculture households are 82.3 percent of the total number of CB's with agriculture households or 62.3 percent of the total number of CB's.

2. **ST2003-L1 and ST2003-L2**: a frame derived from ST2003-L1 and ST2003-L2 will very efficient for use for sub-sector agriculture household sample selection because the data automatically contain variables which are very closely related to the data being collected. ST2003-L1 and ST2003-L2 was conducted in October 2001 and the individual questionnaire for every agricultural household ST2003-S (sample census) will be conducted in October 2004 (except for paddy, which will be conducted one month after ST2003-L1 and L2). Since paddy is only a single crop, it will be easier. But for other sub-sectors we have to use the frame from ST2003-L1 and L2 and there will be some problems in constructing the frame if we have to wait until the processing is finished. To avoid such problems, advance processing of variables used for sample selection will be needed or one could force all processing to be completed by July 2004. The construction of a sampling frame and sample selection (including evaluation) need two months. The CB's sample list has to be finished in August 2004 to recruit field workers and to prepare the training. The frame constructed from ST2003-L1 and L2 will be used not only for the agriculture sample census but also for the following agriculture surveys.

C. Second Pilot

As was mentioned above, the pilot focused only on three areas (paddy, other food crops and horticulture) since these are the only "potential" crops in Kecamatan Cisaat. One-stage sampling was used for paddy agricultural households, systematically arranging by type of irrigation. But during implementation it was found that the real samples were twice the size of the target sample, which makes it difficult to predict the total number of paddy households (see Section III.A.2).

The total number of other main food crops and horticulture agricultural households from 67 CB's in rural areas (complete count) is as follows:

Maize	133	Manggis	4
Soyabean	1	Banana	58
Peanut	7	Papaya	2
Green bean	1	Rambutan	0
Cassava	488	Spring onion	19
Sweet potatoes	80	Mustard green	79
		Other horticulture	75
Other main food crops	596	Horticulture	225

In addition to paddy, Kecamatan Cisaat has a potential only for maize and cassava. Most of the other main food crop agricultural households are cassava agricultural households (488 out of 596 households, or 82 percent). The second largest proportion is made up of households with maize crops (133 out of 596 households, or 22 percent). What this shows is that using an integrated sampling unit is more efficient, because if an area has a potential in some crop, the estimate for that crop can be automatically produced (i.e. do not worry that these crops will not be included in the sample). Another advantage is that with some modification of the questionnaire we can produce the estimate of other crops and automatically produce data at the kabupaten, province, and national levels.

In case of soybeans, peanuts and green beans, it is not possible to estimation small domains because these crops are not "potential" ones for that area. Estimation for these cases is good only for higher levels of aggregation (certainly at the national level).

The same applies to horticulture crops. Moreover, if one looks at more detailed information on crops, some useful observations can be made:

- For main food crops: besides the pre-coded crops, one food crop turns out to be important, talas.

- For horticulture: 13 commodities (in addition to the 6 pre-coded ones) turn out to be important in Cisaat (shallot, kidney bean, red chili, long bean, spinach, cucumber, eggplant, tomato, durian, orange, mango, jack fruit, pineapple).

Table 8 shows the distribution of agricultural households by sub-sector and type of crop for every CB in Kecamatan Cisaat. Of the 6 potential crops, only two are covered by horticulture agricultural households having a high percentage (mustard green 35.1 percent and banana 25.8 percent). On the other hand, some outside potential crops with a high percentage are durian (18.7 percent), cucumber (12.4 percent), and long bean (10.7 percent). This strongly suggests that two-stage multi-purpose sampling design and the choice of horticulture agricultural household as the sampling unit with a multi-purpose questionnaire for sub-sectors is very efficient. Double sampling has a weakness in terms of the probability of all commodities appearing in the sample (especially if we use pre-coded potential crops) and in terms of the complications it introduces (this point was discussed in detail in the previous report).

Table 8
Percentage of Agricultural Households to Total Horticulture Agricultural Households by Type of Commodity

Vegetables		Fruits	
Commodity	%	Commodity	%
(1)	(2)	(3)	(4)
1. Spring onion	8.4	11. Durian	18.7
2. Shallot	1.8	12. Orange	1.9
3. Kidney bean	0.9	13. Mango	4.9
4. Red chili	3.6	14. Manggis	1.7
5. Long bean	10.7	15. Jack fruit	2.2
6. Spinach	4.9	16. Pineapple	3.6
7. Cucumber	12.4	17. Papaya	0.9
8. Mustard green	35.1	18. Banana	25.8
9. Eggplant	3.6	19. Rambutan	0.0
10. Tomato	2.2	Total	100.0

Results of interviews using both the listing (ST2003-L1 and L2) and the sample census (ST2003-S) questionnaires suggest that most sub-sector agricultural households have only one or two commodities. Consequently the questionnaire is not so complicated to fill and therefore will not burden the interviewer. Table 9 shows the correlation of agricultural households by commodity.

Table 9
Correlation of Agricultural Households by Commodity.

Commodity	Manggis	Banana	Papaya	Rambut -an	Spring onion	Mustard green	Others
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manggis	-	0.131	-0.031	0.000	-0.052	-0.071	0.069
Banana		-	0.321	0.000	-0.096	-0.098	-0.055
Papaya			-	0.000	-0.045	-0.062	-0.040
Rambutan				-	0.000	0.000	0.000
Spring onion					-	0.475	0.179
Mustard green						-	0.513
Others							-

One can see clearly that most commodities have very low correlations (the only exceptions are spring onions with mustard green and mustard green with “others”). In other words, most agricultural households only cultivate one commodity.

The main conclusion to derive from this analysis is that commodity agricultural households having a small percentage in one area is not possible to estimate for a small domain. Estimation in this case can only be done adequately for higher levels of aggregation (e.g. provincial and national levels). Commodities with a high percentage in the area will automatically have data estimated for the domain and then the data can be summed up at provincial and national levels.

V. CONCLUSIONS

To summarize, two main conclusions can be derived from the above analysis:

- **The multi-purpose method by sub-sector is the best method**. This involves using the sub-sector agricultural household as the ultimate sampling unit and modifying the questionnaire so that it is not used for only one commodity, but rather for other commodities managed by agricultural households (at least for basic information).
- **The systematic arrangement according to main commodity will make the design more efficient**. Potential crops in the kabupaten will be included automatically in the sample and kabupaten estimation can then be automatically produced. For other crops, estimation can be done at least for the national and provincial levels, and by modifying the questionnaire estimation at the kabupaten may also be possible.

A further steps in evaluating the pilot would include a closer look at ways to minimize the sample size for sub-sectors other than paddy, main food crops and horticulture (e.g. estate crops, animal husbandry, fishery, etc).

GEOGRAPHICAL INTEGRATION OF HOUSEHOLD SURVEYS

Report # 50

by
Sri Budianti Sukmadi

March, 2002

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 300

II. PROBLEM 300

III. SOLUTION: INTEGRATED SAMPLE DESIGN 302
- **A. Units in Master Frame** 302
- **B. Evaluation of Parameters** 303
- **C. Standardization** 305
- **D. Sample Design** 305
  - **1. Parameter Selection** 306
  - **2. Minimum Sample Size** 307

IV. CONCLUSIONS & RECOMMENDATIONS 311

APPENDIX A ANNEX TABLES 313

I. INTRODUCTION

Traditionally, household sample surveys of BPS were designed to produce national level data and at times corresponding provincial level data. From time to time, however, surveys would also attempt coverage at the kabupaten/kota level. Representation of national samples was obviously designed with a higher level of precision (i.e. lower relative standard error) while representation of provincial samples tolerated a lower level of precision. In support of the Government's decentralization policy, BPS has had to think of ways to design samples representative at the kabupaten/kota, provincial as well as national levels with acceptable levels of precision at all levels. To do that in the most efficient way, BPS needs to have survey methodologies which design geographically integrated samples. This report provides general guidelines for selecting geographically integrated demographic and socio-economic household surveys based on the most complete frame available, namely the 2000 population census (SP 2000).

In the development of an integrated design to meet the requirements at the various geographical levels, the proposed approach is to begin from the requirements at the national level, then adjust and expand these to meet provincial requirements, and then requirements at the level of individual kabupaten/kota, even for individual kecamatan in some cases.

Integration of the design is required in both dimensions: spatial and in time. For instance, data may be aggregated over time as we move to lower levels geographically. The present paper addresses only the first, namely the spatial dimension. Sampling over time will be discussed in a later paper.

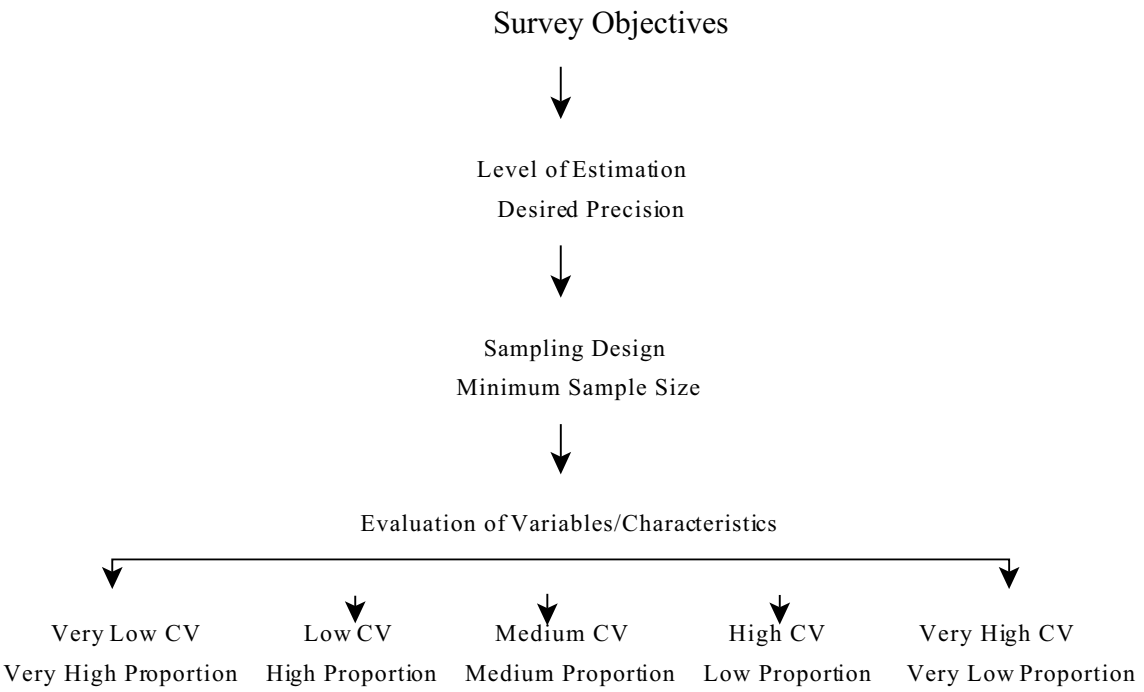
II. PROBLEM

In designing surveys, the main problem is usually choosing a sample of respondents which allows us to make inferences from the sample value to the population value given a limited number of respondents and a limited budget. Survey objectives should determine what sample design to be adopted, but it is often the case that with serious budget limitations, sample design influences and sometimes forces changes in survey objectives. It is not realistic to design a sample where all variables have the same level of precision in the estimation, because not all variables have the same standard deviation. There are rare case variables, variables with very high coefficients of variation in the population (e.g. births and deaths, expenditure on durable goods) and variables with more detailed categories (e.g. industry or occupation at four-digit classifications).

Specifying precision for detailed classifications has to be clearly stated based on the objectives of the survey. Some variables in a survey may be represented well enough for national and regional estimations, others only for national estimation, and still other may continue to have a high relative standard error even for national estimation. Figure 1 explains the decisions one has to make: objectives determine the level of estimation and the desired level of precision; which in turn determine the minimum sample size required. At one extreme, a "very high" level of precision requires a "very low cv"; at the other extreme, a "very low" level of precision requires a "very high cv". In practice, compromises are always sought and these are the levels which have been achieved in the past:

- national, provincial, kabupaten/kota and kecamatan estimation: for variables with "very low" or "low" cv's
- national, provincial and kabupaten/kota estimation: for variables with "medium" cv's
- national and provincial estimation: for variables with "high" and "very high" cv's

Figure 1



It is worth noting that these problems of adequate representation come up in every survey, be it a household or an establishment survey. Users always want the survey objectives to be as detailed as possible, and BPS always attempts to determine, given available budgets, the most appropriate level of representation.

III. SOLUTION: INTEGRATED SAMPLE DESIGN

To ensure consistency as well as efficiency of samples designed to meet national objectives with those designed to meet kabupaten/kota objectives, sample design needs to be geographically integrated. The SP2000-L1 and SP2000-L2 can provide the basic frame for selection of various level samples. Basic indicators (e.g. mean, median, coefficient of variation, etc.) of relevant variables can be calculated for the purpose of identifying those with a high correlation with the items in the surveys. These can then be classified according to the categories in Figure 1.

There are four steps for activity identification:

1. Province and kabupaten/kota which have to have a particular level of estimation

2. Census Block as the primary sampling unit. It is also possible to have kecamatan estimation by grouping samples and taking one or two sub-samples. By grouping kecamatan before sampling selection we may obtain more efficient results (i.e. a smaller relative standard errors).

3. Households as the ultimate sampling unit.

A. Units in Master Frame

The most complete master frame that can be used is the complete count data of SP2000. Census blocks (CB) are probably the most efficient sampling units to use. They can be defined as clusters constructed during SP2000. Before using the SP2000 results for this purpose, however, a number of steps need to be undertaken:

1. Check the completeness of CB's by desa/kelurahan and kecamatan according to SP2000-L1 and SP2000-L2.

2. Check the completeness of available CB maps and boundaries, their consistency with SP2000-L1 and L2 and if necessary update them during the preparation of Agriculture Census 2003.

3. Identify the total number of households used for sample selection, especially by using probability proportional to size.

4. Monitor changes of administrative units and CB's as a routine activity since this condition influences the master frame and the estimation.

5. Identify the variables used for information on how big the minimum sample size needed is (as in Figure 1).

6. Calculate the proportions/ratios by variable in SP2000 and group them according to the level of heterogeneity/proportion/ratio.

7. Evaluate standard errors of sampling proportions by using simple random sampling in various sample sizes as shown in Table A.5.

8. By using sample sizes in the table and combining them with ratios of cases of variables in the frame, we can estimate the rough sample sizes needed with a simple method.

9. By using the design effect from surveys which have been conducted by BPS (e.g. Susenas, Sakernas, Demographic and Health Survey etc.), more accurate minimum sample sizes can be estimated to figure out the desired precision.

Once these activities are completed, then two more steps remain

- Evaluation of parameters (e.g. distribution, coefficient of variation, correlation, ratio, etc.). Some variables will then need to be selected which are sufficient for a rough estimation of the sample size needed.

- A standardized sampling design methodology needs to be developed and a prototype sample can be documented in a manual to be used for national, regional as well as kabupaten/kota selection and estimation.

These will be discussed in more detail in the following sections.

B. Evaluation of Parameters

There are two groups of parameters in the Population Census 2000: those in SP2000-L1 (building and household listing) and those in SP2000-L2 (characteristics of individual household members). The first set of variables are useful for stratification and the second for targeting of sample sizes.

The list of variables in SP2000-L1 (building and household listing) includes:

1. Name/number of local administrative unit
2. Segment number
3. Serial number of physical building
4. Serial number of census building
5. Usage of census building : Non dwelling
   Dwelling
   Mixed
6. Dwelling building component: Permanent/non permanent
   With facility : Bedroom
   Kitchen
   Bathroom
   Toilet
   Water
   Electricity
   Occupied/not occupied
7. Total number of households in the census building

8. Serial number of household
9. Type of household: ordinary or special
10. Name of head of household
11. Total number of household members (male, female, total)
12. Ownership status of census building (self owned or else). If else : rent/contract or others.

Variables in SP2000-L1 are very useful if used for stratification before sampling selection. That would produce a more efficient sampling design. CB's can then also be stratified by:

- Urban and rural areas
- Type of CB's: common, special and preparation CB
- Total number of dwelling buildings: permanent/non permanent with facilities. This will be useful for stratification of household expenditure surveys, for example.
- Total number of households and household members

Variables in SP2000-L2 (household member characteristics) include:

1. Total number of household members
2. Names of household members
3. Relationship to head of household: Head
   Spouse (wife/husband)
   Children
   Others
4. Sex: Male
   Female
5. Place of birth: Province, Kabupaten/Kota
6. Month and year of birth and age
7. Marital status: Single
   Married
   Divorced
   Widowed
8. Religion: Moslem
   Catholic
   Protestant
   Hindu
   Buddhist
   Others
9. Citizenship / ethnic: Indonesia (specify the ethnic)
   Foreigner (specify the citizenship)
10. Place of residence 5 years ago : Province, Kabupaten/Kota
11. Education or highest certificate of education attained : Not going to school
    Primary school
    Junior high school
    Senior high school
    Academy/DIII
    University/DIV
12. Activities during previous week: Employed
    Looking for work
    School
    Others

13. Type of main industry: Food crops
    Horticulture
    Fishery
    Animal husbandry
    Plantation
    Manufacturing
    Trade
    Services
    Transportation
    Others
14. Employment status: Self employed
    Self employed assisted by family/temporary employee
    Employer
    Employee
    Unpaid worker
15. Total number of children ever born
16. Total number of children still alive

These variables would be useful for targeting sample sizes by calculating ratios or percentages such as: sex ratio, child women ratio, percentage of illiterate, labor force participation rate, unemployed rate, etc. by kecamatan, kabupaten/kota, province and at the national level.

C. Standardization

The purpose of integrating survey design is to ensure consistency and efficiency of samples. The multi-stage sampling design discussed in this report is not only for national samples, but can also be applied for provincial and kabupaten/kota samples as well. The main point made in this report is that by standardizing all procedures (design, selection, operation) according to the guidelines provided here, BPS will be able to produce samples which will allow kabupaten/kota, provincial and national representation at the same time with the minimum required sizes. Kabupaten/kota offices can increase their samples, if they want to improve their level of precision, by following the same methodology discussed in this report, without compromising coverage, representativeness or comparability with the national sample. At the same time, by allowing district offices to design samples according to their own needs, provided they use the BPS standardized procedures, they will relieve the BPS Directorate of Methodology from the impossible burden of having to design 341 separate samples with different characteristics.

D. Sample Design

The simplest way to get a rough idea of the needed size of a particular household sample is to look at proportions: e.g. proportion of population with particular levels of education, health, social welfare etc. It is easy to link such proportions to the size of samples needed with specified standard errors (see Table A.5). The exceptions are income and expenditure variables.

In the current Susenas, the total number of estimations are:

a. National level (national, urban, and rural)	:	3
b. Province level (32 provinces by total, urban, and rural)	:	96
c. Kabupaten/Kota level (total)	:	341
Total	:	440

Because budgets are very limited, most BPS household surveys ate limited to national and provincial level estimations. Some provinces now have larger budgets that allow them to require kabupaten/kota-level estimations; others have budgets allowing coverage of some kabupaten/kota but not all; and others cannot fund any. Some kabupaten/kota, in turn, have budgets that allow them to require kecamatan-level estimation, but others don't.

To ensure consistency of selection procedures at all levels, we will start with the primary sampling units and work our way up. For estimation at the kabupaten/kota level within a particular province, calculation of the minimum sample size needs to take the following into consideration:

- number of kecamatan and number of CB's in each kabupaten/kota. The reason is that they have high correlations with the total number of households or the population, the sample design, and the sampling fractions.
- main characteristics and the level of heterogeneity within the kabupaten/kota
- intra-class correlation coefficients of some basic variables for calculating the design effect using cluster sampling and determining cluster sizes.

In what follows, we will present an example of the steps to be taken for sample design using the SP 2000 parameters.

1. Parameter Selection

The Methodology Directorate, which is in charge of compiling the master sampling frame, needs to produce the list of parameters from SP2000-L2 by calculating the ratio of events by desa/kelurahan, kecamatan, kabupaten/kota and province. These ratios can then be evaluated in depth and grouped by the five categories in Table 1 ("very low", "low", "medium", "high" and "very high"). Results can then be compared with Table A.5 (standard error of estimate of proportion in simple random sampling) to have some idea of how big the rough sample sizes by kecamatan are.

This kecamatan-level procedure can be easily linked with the existing sampling design for Susenas which uses the kecamatan as the primary sampling unit. With this design, it is possible to integrate the national design, which goes down until the kabupaten/kota level, and the regional design down to the kecamatan level.

Selection of kecamatan by independent sub-samples will help increase sample sizes in kecamatan to produce kecamatan-level estimation by selecting additional sub-samples of CB's with minimum requirement sample sizes. The integrated design will produce:

- Kabupaten/kota, province and national level estimation for core questions with the kecamatan as the primary sampling unit

- Sufficient simultaneous sub-samples for more than one year. One or two independent sub-samples are selected for one year or we may arrange the rotation method such as in the Sakernas Survey.

- Extra sub-samples if needed for lower level estimation of domains.

Households will be used as the ultimate sampling unit.

2. Minimum Sample Size

The minimum sample size needed will depend on the level of estimation and the condition of rare/non-rare case variables and heterogeneity/homogeneity of characteristics. The simple formula for a rough individual sample size of each domain of estimation is:

$$Z_{\alpha/2}\sqrt{\frac{pq}{n}} = k, \text{ so then } \quad n = \frac{Z^2_{\alpha/2}pq}{k^2}$$

where k : margin of error

p : proportion of observation unit of certain variable/item in the sample

q : 1 - p

$Z_{/2}$: value of Z at (1 -) level of confidence (at 95% level of confidence, the value of $Z_{0.025} = 1.96$)

In most household surveys the design effect (ratio of actual to SRS variance deft ^{2}) is around 2. That means that, due to multistage clustering of the sample, the variance increases by a factor of about 2 over that in an equivalent simple random sampling. Then the total sample size becomes:

$$n = \frac{Z^2_{\alpha/2}pq}{k^2} \text{ x design effect}$$

By using households as ultimate sampling units there are two steps for sample size calculation:

- First: calculate the sample size (n) as above to know how big the samples of observation/individual unit for a certain variable/characteristic.

- Second: estimate the size of the sample households by using the proportion of that variable in households.

Here is a numerical example.

1. Proportion of variable in sub-classes: there are two percentages, percentage in the elements and percentage in sampling units. Variables have ratios around 0.8 and 0.9 (high proportion/percentage in domain), but small percentage in sampling unit, for example variable "one year old children" with variable in health such as "immunization." The proportion of "one year old children", for example, is only around 12 percent in the households.
   The ultimate sample size needed is:

   a. With relative standard error around 5 percent:
      k = (0.05)(0.8) = 0.04
      $$n = \frac{(0.8)(0.2)}{0.04^2} x2 = 200$$ ("one year old children")
      or around 1650 households
      If we want to estimate at the 95% confidence level, the sample size of households is around 6400.

   b. With relative standard error around 10 percent
      k = (0.10)(0.8) = 0.08
      $$n = \frac{(0.8)(0.2)}{0.08^2} x2 = 50$$ ("one year old children")
      or around 420 households
      If we want to estimate at the 95% confidence level, the sample size is around 1613.

   c. With relative standard error around 15 percent
      k = (0.15)(0.8) = 0.12
      $$n = \frac{(0.8)(0.2)}{0.12^2} x2 = 22$$ ("one year old children")
      or around 184 households
      If we want to estimate at the 95% confidence level, the sample size of households is around 707

2. Proportion directly from sampling unit: for example, "use of drinking water by households". In this case the household is an observation unit and an ultimate unit. If the proportion is 73 percent, the sample size needed is:

   a. With relative standard error around 5 percent
      k = (0.05)(0.73) = 0.0365
      $$n = \frac{(0.73)(0.27)}{0.0365^2} x2 = 295$$ households
      If we want to estimate at the 95% confidence level, the sample size of household is around 1137.

   b. With relative standard error around 10 percent
      k = (0.10)(0.73) = 0.073
      $$n = \frac{(0.73)(0.27)}{0.073^2} x2 = 74$$ households

If we want to estimate at the 95% confidence level,
the sample size of households is around 284.

From these results we have a rough idea that a kecamatan estimation with a relative standard error of around 10 percent would require a total number of sampled households around 150. For big kecamatan we can increase the sample size to 200, for example.

The example also shows that there has to be some agreement on the size of the relative standard error for non-rare case variables and rare case variables by level of estimation. The relative standard error may be bigger for lower domains of estimation: for example 15 - 20 percent for kecamatan level, 10 - 14 percent for kabupaten/kota level, 5 - 9 percent for province level and less than 5 percent for national level for non-rare cases. By observing the SP2000 parameters using the above procedure, it is easy to determine the sample size for each domain of study (i.e. by calculating the ratio of those variables by level of domain such as in Table 2).

Table 2
Ratio of Variables by Domain/Unit

Parameter variables	CB	Desa/ kelurahan	Kecamatan	Kabupaten/ Kota	Province
(1)	(2)	(3)	(4)	(5)	(6)
1. Number of CB's	X	X	X	X	X
2. Number of buildings	X	X	X	X	X
3. Number of households	X	X	X	X	X
4. Number of household members	X	X	X	X	X
5. Percentage of permanent building by facility	X	X	X	X	X
6. Percentage of own house	X	X	X	X	X
7. Sex ratio	X	X	X	X	X
8. Education by level	X	X	X	X	X
9. Percentage of urban population	-	-	X	X	X
10. Child women ratio	-	-	X	X	X
11. Old dependency ratio	-	-	X	X	X
12. Labor force participation rate	-	-	X	X	X
13. Unemployment rate etc	-	-	X	X	X

Calculating the ratios in Table 2 will help determine the approximate sample size for each domain of study from the lowest geographical breakdown up to the national level:

a. By evaluating the ratio of variables by kecamatan (if the lowest domain is kecamatan) and grouping the ratio into the five categories of Figure 1.

b. Kecamatan within a particular group will require the same minimum sample size for such variables. Kabupaten/kota and national level estimation will be the sum of kecamatan levels.

c. If the estimation domain is the kabupaten/kota, the evaluation becomes easier because we can start with the ratio at the kabupaten/kota level. If the estimation is until the province level, then the evaluation would start from the province.

d. If only some kecamatan or kabupaten/kota have enough budget, the calculation of sample size can start from the higher domain. For the particular kecamatan or kabupaten/kota, the additional sample can be taken by adding sub-samples using the same procedure.

For illustration, Tables A.1 and A.2 show the potential workload involved in evaluating SP 2000 parameters at the national and provincial levels when compared, for example, with that of doing it by kabupaten/kota for a province like Lampung (Tables A.3.A through A.3.C). As the average number of households per CB and cv's for national and provincial figures are more or less the same (about 100 for the former and 40 percent for the latter), the main issue is how big the needed sample size is. Tables 3 and 4 provide the main parameters from Susenas 2002 respectively for Lampung province and for kabupaten Lampung Selatan. Samples in this survey were designed for national, provincial and kabupaten/kota representation.

If the Lampung Selatan Kabupaten office decided, for example, to have estimation at the kecamatan level, this can be done using the following steps:

- First, select a sub-sample using the above procedure (adding about 150 households)
- Give the sample list the random number of the existing samples and the frame used for the sample selection.
- Put new random numbers for the additional sub-sample selected by the kabupaten/kota office and continue with the sample selection.
- Do estimation for each sub-sample and their average (independent sub-samples)

As Table 4 shows, while an additional sample is needed for kecamatan level estimation in Lampung Selatan, no additional sample will be needed for Kota Metro because samples in the kota are already big enough (208 households for Kecamatan Bantul and 752 for Kecamatan Metro Raya).

Table 3
Susenas 2002 Sample Sizes per Kabupaten/Kota, Lampung Province

Kabupaten/Kota	Sample			Sampled households
	Kecamatan	CB's	Average CB's per kecamatan	
(1)	(2)	(3)	(4)	(5)
1. Kab. Lampung Barat	6	40	6.7	640
2. Kab. Tanggamus	11	38	3.5	608
3. Kab. Lampung Selatan	13	34	2.6	544
4. Kab. Lampung Timur	12	60	5.0	960
5. Kab. Lampung Tengah	13	60	4.6	960
6. Kab. Lampung Utara	8	34	4.3	544
7. Kab. Way Kanan	6	34	5.7	544
8. Kab. Tulang Bawang	8	38	4.8	608
9. Kodya Bandar Lampung	9	40	4.4	640
10. Kodya Metro	2	60	30.0	960
Total	88	438	5.0	7008

Tabel 4
Susenas 2002 Sample Sizes for Kabupaten Lampung Selatan & Kota Metro

Kabupaten/Kota	Kecamatan	CB's	Households
(1)	(2)	(3)	(4)
1. Kab. Lampung Selatan	1. Padang Cermin	4	64
	2. Kedundung	2	32
	3. Gedung Tataan	2	32
	4. Negeri Katon	2	32
	5. Tegineneng	2	32
	6. Natar	4	64
	7. Jati Agung	2	32
	8. Tanjung Bunting	2	32
	9. Katibung	4	64
	10. Sidumulyo	3	48
	11. Kalianda	3	48
	12. Padas	2	32
	13. Panengahan	2	32
	Total	34	544
2. Kota Metro	1. Bantul	13	208
	2. Metro Raya	47	752
	Total	60	960

IV. CONCLUSIONS & RECOMMENDATIONS

Two clear conclusions can be drawn from the above discussion:

- To ensure consistency and efficiency of sample design, it is essential to have a geographically integrated standardized methodology applying to sample selection at the national,

provincial, kabupaten/kota and kecamatan levels. Such a methodology is certainly feasible and needs to be produced as quickly as possible.

- The best source to use as a master frame to apply this methodology is the SP 2000 (both L1 and L2 results). For selecting kabupaten/kota (kecamatan) samples, tables (derived from this master frame) need to be produced including parameter ratios grouped by kabupaten/kota (kecamatan).

To achieve this, our main recommendation is for the BPS Methodology Directorate to produce a detailed manual of this methodology, based on the guidelines enunciated in this report, with a prototype sampling selection procedure and send to all provincial and kabupaten/kota offices. The second step would be to evaluate feedback from these offices before adopting it as a firm policy over the next few years.

APPENDIX A
ANNEX TABLES

Table A.1
Number of Kabupaten/Kota, Kecamatan, and Desa/Kelurahan by Province

Province	Kabupaten/ kota	Kecamatan	Desa/kelurahan		
			Urban	Rural	Total
(1)	(2)	(3)	(4)	(5)	(6)
11. Dista Aceh	13	146	440	5156	5596
12. Sumatera Utara	19	265	758	4585	5343
13. Sumatera Barat	15	120	429	1747	2176
14. Riau	15	96	209	1259	1468
15. Jambi	10	60	123	1038	1161
16. Sumatera Selatan	7	88	258	2386	2644
17. Bengkulu	4	32	118	1044	1162
18. Lampung	10	88	178	1894	2072
19. Bangka Belitung	3	23	104	224	328
31. DKI Jakarta	5	43	265	-	265
32. Jawa Barat	22	447	1833	3915	5748
33. Jawa Tengah	35	534	2398	6145	8543
34. DI Yogyakarta	5	75	169	269	438
35. Jawa Timur	37	624	2399	6052	8451
36. Banten	6	96	400	1076	1476
51. Bali	9	53	231	447	678
52. NTB	7	62	168	535	703
53. NTT	14	124	147	2368	2515
61. Kalimantan Barat	9	128	98	1346	1444
62. Kalimantan Tengah	6	85	59	1269	1328
63. Kalimantan Selatan	11	117	202	1951	2153
64. Kalimantan Timur	12	88	140	1138	1278
71. Sulawesi Selatan	5	72	190	964	1154
72. Sulawesi Tengah	8	69	93	1343	1436
73. Sulawesi Selatan	24	197	406	2724	3130
74. Sulawesi Tenggara	5	67	106	1445	1551
75. Gorontalo	3	21	73	299	372
81. Maluku	5	29	67	778	845
82. Maluku Utara	3	27	62	662	724
91. Irian Jaya Barat	4	44	22	730	752
92. Irian Jaya Tengah	5	44	27	615	642
93. Irian Jaya Timur	5	85	39	1435	1474
Total	341	4049	12211	56839	69050

Source: BPS, Kerangka Contoh Induk Blok Sensus, Tahun 2000

Table A.2
Number of Common CB's by Group of Households, Indonesia

Total number of households	Urban	Rural	Total
(1)	(2)	(3)	(4)
< 21	4719	8215	12934
21 - 40	6473	6316	12789
41 - 60	16057	16702	32759
61 - 80	33196	40585	73781
81 - 100	46708	68951	115659
101- 120	40292	66726	107018
121 - 140	25822	42990	68812
>140	32977	36295	69272
Total	206244	286780	493024
Average households per CB's	106	105	105
Coefficient of variation (%)	47.22	35.93	41.41

Source: BPS, Kerangka Contoh Induk, Tahun 2000.

Table A.3.A
Number of Common CB's by Group of Households, Lampung Province

Total number of households	Urban	Rural	Total
(1)	(2)	(3)	(4)
< 21	69	219	288
21 - 40	145	162	307
41 - 60	325	544	869
61 - 80	514	1423	1937
81 - 100	693	2935	3628
101- 120	607	2886	3493
121 - 140	415	1994	2409
>140	464	1657	2121
Total	3232	11820	15052
Average households per CB's	101	109	107
Coefficient of variation (%)	45.01	34.95	37.30

Source: BPS, Kerangka Contoh Induk, Tahun 2000.

Table A.3.B
Number of Kecamatan and Desa/Kelurahan by Kabupaten/Kota, Lampung

Kabupaten/Kota	Kecamatan	Desa/Kelurahan		
		Urban	Rural	Total
(1)	(2)	(3)	(4)	(5)
1. Kab. Lampung Barat	6	4	167	171
2. Kab. Tanggamus	11	22	291	313
3. Kab. Lampung Selatan	13	22	330	352
4. Kab. Lampung Timur	12	7	225	232
5. Kab. Lampung Tengah	13	10	276	286
6. Kab. Lampung Utara	8	16	192	208
7. Kab. Way Kanan	6	-	192	192
8. Kab. Tulang Bawang	8	7	214	221
9. Kodya. Bandar Lampung	9	82	2	84
10. Kodya Metro	2	7	5	12
Total	88	177	1894	2071

Table A.3.C
Number of CB's by Type, Lampung

Kabupaten/Kota	Type of CB			Total
	Common	Special	Preparation	
(1)	(2)	(3)	(4)	(5)
1. Kab. Lampung Barat	763	1	83	847
2. Kab. Tanggamus	1754	-	445	2199
3. Kab. Lampung Selatan	2484	9	453	2946
4. Kab. Lampung Timur	2011	4	582	2597
5. Kab. Lampung Tengah	2529	6	487	3022
6. Kab. Lampung Utara	1106	1	295	1402
7. Kab. Way Kanan	773	1	320	1094
8. Kab. Tulang Bawang	1613	4	768	2385
9. Kodya. Bandar Lampung	1787	10	30	1827
10. Kodya Metro	232	1	8	241
Total	15052	37	3471	18560

Table A.4
Estimates of Sampling Error for Education Variable

Province	No education				Secondary education and higher			
	Value (R)	SE	Deft	RSE	Value (R)	SE	Deft	RSE
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Jawa-Bali	0.138	0.008	2.049	0.055	0.246	0.009	2.002	0.038
DKI Jakarta	0.038	0.006	1.241	0.149	0.543	0.020	1.680	0.036
Jawa Barat	0.122	0.013	1.548	0.104	0.203	0.015	1.497	0.074
Jawa Tengah	0.148	0.015	1.582	0.099	0.228	0.023	2.107	0.101
DI Yogyakarta	0.102	0.017	1.784	0.163	0.453	0.026	1.710	0.058
Jawa Timur	0.163	0.016	1.738	0.101	0.224	0.015	1.440	0.069
Bali	0.209	0.020	1.749	0.095	0.285	0.020	1.615	0.071
2. Outer Jawa-Bali I	0.118	0.006	1.896	0.050	0.342	0.010	2.261	0.030
Dista Aceh	0.137	0.014	1.402	0.104	0.347	0.025	1.779	0.072
Sumatera Utara	0.061	0.015	2.367	0.247	0.409	0.031	2.345	0.075
Sumatera Barat	0.032	0.010	1.673	0.308	0.458	0.031	1.893	0.068
Sematera Selatan	0.096	0.015	1.720	0.159	0.327	0.033	2.309	0.100
Lampung	0.079	0.012	1.400	0.151	0.252	0.022	1.574	0.086
NTB	0.334	0.023	1.649	0.070	0.216	0.022	1.789	0.103
Kalimantan Barat	0.248	0.025	1.817	0.100	0.275	0.024	1.690	0.087
Kalimantan Selatan	0.085	0.011	1.241	0.125	0.334	0.025	1.702	0.074
Sulawesi Utara	0.014	0.004	0.888	0.257	0.513	0.032	1.903	0.063
Sulawesi Selatan	0.162	0.021	2.007	0.128	0.311	0.029	2.205	0.092
3. Outer Jawa-Bali II	0.135	0.011	3.012	0.079	0.333	0.012	2.365	0.035
Riau	0.139	0.015	1.367	0.108	0.320	0.022	1.505	0.069
Jambi	0.179	0.064	4.986	0.359	0.266	0.046	3.077	0.174
Bengkulu	0.102	0.013	1.220	0.132	0.338	0.027	1.571	0.080
NTT	0.156	0.021	1.694	0.137	0.233	0.028	1.926	0.121
Timor Timur	0.540	0.030	1.832	0.056	0.206	0.027	2.027	0.131
Kalimantan Tengah	0.078	0.024	2.447	0.305	0.309	0.031	1.882	0.102
Kalimantan Timur	0.064	0.011	1.365	0.178	0.449	0.037	2.199	0.083
Sulawesi Tengah	0.039	0.012	1.903	0.319	0.372	0.035	2.123	0.093
Sulawesi Tenggara	0.094	0.015	1.421	0.163	0.416	0.037	2.032	0.089
Maluku	0.044	0.008	1.112	0.183	0.451	0.022	1.243	0.048
Irian Jaya	0.221	0.027	1.815	0.121	0.364	0.031	1.837	0.086
Indonesia	0.132	0.005	2.566	0.039	0.281	0.007	2.459	0.023

Source: Demographic and Health Survey 1997

Total number of sample households : 1650 - 2050 in Jawa - Bali
1000 - 1500 in Outer Jawa - Bali

Table A.5
Standard Error of Estimate of Proportion in Simple Random Sampling

($\sigma_p = \sqrt{\frac{PQ}{n}}$ for specified value of P and n)

n = number of sample cases	P = Proportion of units having a characteristic (Q = 1 - P has the same standard error, but not relative standard error)								
	.001 or .999	.01 or .99	.03 or .97	.05 or .95	.10 or .90	.15 or .85	.20 or .80	.30 or .70	.50
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
50	.0045	.0141	.0240	.0420	.0510	.0570	.0570	.0650	.0710
100	.0032	.0099	.0170	.0300	.0360	.0400	.0400	.0460	.0500
200	.0022	.0071	.0120	.0210	.0250	.0280	.0280	.0330	.0350
300	.0018	.0058	.0099	.0170	.0210	.0230	.0230	.0270	.0290
400	.0016	.0050	.0086	.0150	.0180	.0200	.0200	.0230	.0250
500	.0014	.0045	.0076	.0130	.0160	.0180	.0180	.0210	.0220
600	.0013	.0041	.0070	.0120	.0150	.0160	.0160	.0190	.0200
700	.0012	.0038	.0065	.0110	.0140	.0150	.0150	.0170	.0190
800	.0011	.0035	.0061	.0110	.0130	.0140	.0140	.0160	.0180
1000	.0010	.0032	.0054	.0095	.0110	.0130	.0130	.0150	.0160
1200	.0009	.0029	.0049	.0087	.0100	.0120	.0120	.0130	.0140
1500	.0008	.0026	.0044	.0077	.0093	.0100	.0100	.0120	.0130
1700	.0008	.0024	.0042	.0073	.0087	.0097	.0097	.0110	.0120
2000	.0007	.0022	.0038	.0067	.0081	.0090	.0090	.0100	.0110
2500	.0006	.0022	.0034	.0060	.0072	.0080	.0080	.0092	.0100
3000	.0006	.0018	.0031	.0055	.0066	.0073	.0073	.0084	.0092
3500	.0005	.0017	.0029	.0051	.0061	.0068	.0068	.0078	.0084
4000	.0005	.0016	.0027	.0047	.0057	.0063	.0063	.0079	.0079
4500	.0005	.0015	.0025	.0045	.0054	.0060	.0060	.0069	.0074
5000	.0004	.0014	.0024	.0042	.0051	.0057	.0057	.0065	.0071

1) Values of n greater than 5000 : When n is multiplied by 100, the standard error is divided by 10
2) Practically the sample value of p would be used, since the value of P from population is not known.

USING ROLLING SAMPLES IN SELECTED BPS SURVEYS

Report # 52

by
Sri Budianti Sukmadi

May, 2002

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 320

II. WHICH SURVEYS CAN BE AFFECTED? 320
- **A. Susenas** 320
- **B. Sakernas** 324
- **C. Susi** 324
- **D. Crop-Cutting Survey** 324

III. ILLUSTRATIVE CONSEQUENCES OF DESIGN 325

IV. CONCLUSIONS 332

I. INTRODUCTION

BPS conducts regularly two major types of household surveys: the National Socio-Economic Survey (Susenas) and the Labor Force Survey (Sakernas). These are conducted annually (or quarterly) between census years and are generally designed to produce estimates at the national and provincial levels. The only exception to this rule is the core Susenas survey, which has been designed in the past few years to produce kabupaten/kota level estimates. Even with the core Susenas survey, difficulties remain in estimating kabupaten/kota level estimates for rare case variables.

Given the increasing importance of providing data at the district level and the increasing likelihood of continued budgetary constraints, BPS began evaluating ways to integrate design of its household surveys. Integration can cover two dimensions: space and time. An earlier report[1] addressed the first dimension. This report addresses the second, namely integration over time.

Integration of samples over time is done by using what is referred to as "rolling samples," that is using different samples at different times (e.g. monthly, quarterly or annually) which are specifically designed to be combined over a particular period (e.g. one year, three years) in order to produce a large enough sample size with data of acceptable precision even at a disaggregated geographical level (e.g. district or sub-district). In other words, by carefully selecting periodic (e.g. quarterly) samples, one can continue to use the same current periodic national/provincial sample sizes but would maximize the possibility of producing reliable annual data for lower-level geographic units (e.g. kabupaten/kota). This paper discusses selected issues related to design of rolling samples.

II. WHICH SURVEYS CAN BE AFFECTED?

The rolling sample design can be applied to several household or household-type surveys currently conducted by BPS: Susenas, Sakernas, Susi and crop-cutting surveys. The following sections will discuss in detail how that design can be applied to the particular survey.

A. Susenas

The National Socio-Economic Survey (Susenas) is the largest regular household survey conducted by BPS. It consists of two major parts:

---

[1] Sri Budianti Sukmadi, *Geographical Integration of Household Surveys*, STAT Project Report #50, March, 2002.

a. "Core" surveys: these are conducted annually covering over 200,000 households in about 340 kabupaten/kota. The questionnaire is more or less the same every year and contains more than 75 variables including:

- main characteristics of household members (relation to the head of household, sex, age, marital status, etc)
- individual health (type of health complaint, type of treatment or checkup etc.)
- education of household members aged 5 years and over (education level/degree etc.)
- activities of household members aged 10 years and over (type of activity, industrial status, employment status, etc)
- fertility and family planning (first marriage, children ever born, contraceptive type etc.)
- housing type, activity and environment (construction material, lighting, drinking water etc.)
- consumption and expenditure by group and main source of income.

b. "Module" surveys: these are surveys designed to cover a smaller number of households (62,000) but in greater depth. These are covered on a rotational basis once every three years:

- one covering welfare, criminality and tourism
- one covering health, nutrition, education and home improvement
- one covering income and expenditure.

With such a design, the Susenas "core" survey is probably ideal for applying the rolling sample design for several reasons:

- the same variables are covered from year to year, making results for a particular variable comparable between years.
- different districts are covered more fully in one particular year while at the same time coverage by province and for the nation is adequate every year. By combining results over many years, coverage of districts increases substantially without an increase in the sample size for any particular year.

- by spreading the workload over a longer time period, the burden of data collection and processing in any one period would be reduced. Needless to say that the ensuing budget savings would be substantial.

Thus application of the rolling sample design to the Susenas core surveys would facilitate the joint satisfaction of three important objectives: spatial representation, temporal representation and a substantial increase in the richness of the data. Figures 1 and 2 illustrate graphically two scenarios: one combining three-year rolling samples and the other combining five-year rolling samples. If data are combined for three years, we would have a sample of about 630,000 households, which would not only allow representation at the kabupaten/kota level, but may even allow representation for some kecamatan. If data are combined for five years, such representation would be even more feasible.

As for combining Susenas "module" surveys, one can still combine, for example, the 1996, 1999 and 2002 "income and expenditure" surveys, which would provide a sample of 200,000 households. Such combination, despite the 6-year span of these surveys, would still provide a useful data set which allows kabupaten-level representation.

Figure 1
First Alternative (3 year rolling)

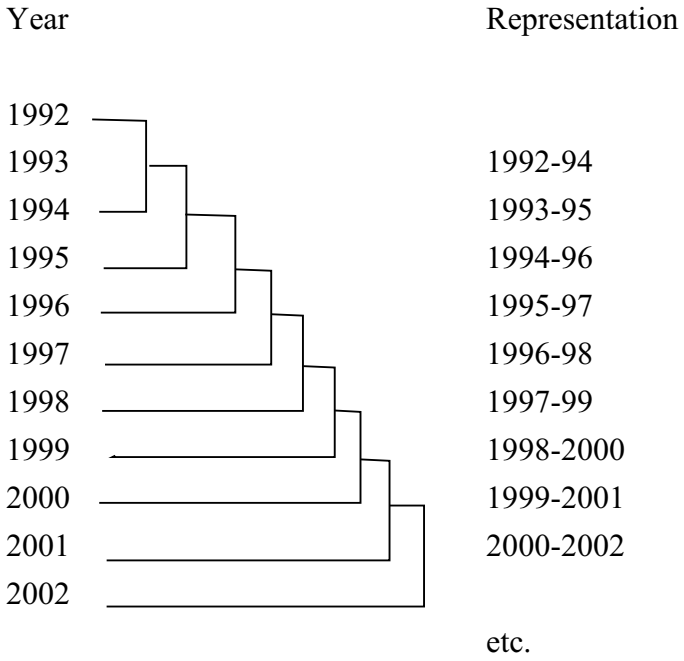
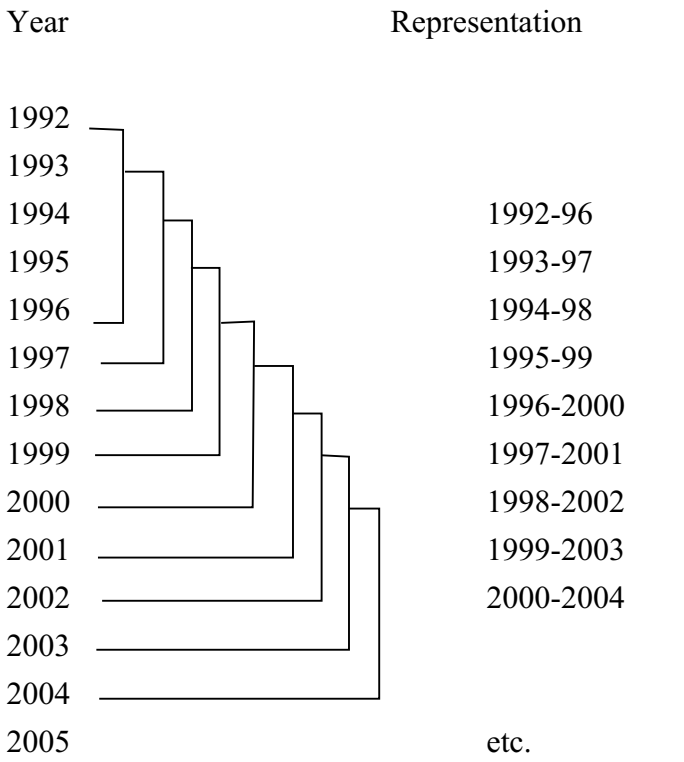


Figure 2
Second Alternative (5 year rolling)



B. Sakernas

The National Labor Force Surveys (Sakernas) are periodic surveys which, unlike Susenas, have undergone substantial changes over time. Since 1994 they have been conducted annually, but in 2002 a new design is being tested. The format being tested involves independent quarterly sub-samples designed to be representative of the nation as a whole, with a 50% overlap between two consecutive quarters to allow reliable tracking of quarterly growth, and a full sample over four quarters which provides adequate representative by province.[2]

The new design provides a perfect example of how rolling samples can be applied to provide highly efficient samples at a time of significant budget constraints. A sample of 20,000 households is covered every quarter, with 10,000 overlapping between two consecutive quarters. For the whole year, a cumulative total of 50,000 households would be obtained, and if combined over 3 years, a cumulative total of 150,000 households. Results of this pilot will be carefully evaluated and compared with the point estimates obtained from the annual survey.

C. Susi

The Survei Usaha Terintegrasi (Susi) has been conducted every year since 1998 to cover small-scale/household-type establishments outside agriculture. Coverage has depended on available budgets: in 1998 it covered 120,000 establishments and data collection was conducted in November; since then it was divided into quarterly sub-samples covering an annual cumulative sample of 90,000 in 1999 and 60,000 in 2000. The use of quarterly sub-samples continues as of today for the purpose of designing the next sample and an evaluation of the methodology is underway. Once decision is made as to the frequency of future surveys, and regardless of whether they will use quarterly or annual samples, it will still be possible to continue to use rolling samples (whether quarterly or annual).

The reason for having rolling samples in Susi is the same as that in Susenas: to allow cumulation of detailed data to enable estimation for smaller administrative units. Local (kabupaten) governments may need data now only available at the national or provincial level: e.g. by 1-digit ISIC, or covering more variables than simply the number of establishments and employment (e.g. owner's education, worker education, age, sex, wages etc.).

D. Crop-Cutting Survey

The crop-cutting survey for main food crops (paddy, maize, cassava, sweet potatoes, soybeans and green beans) is a monthly survey which has a different objective than the above three BPS surveys: although conducted monthly, the survey's main objective is to estimate annual

---

[2] Vijay Verma, *Selected Issues in labor Force Statistics*, STAT Project Report #25, March 2001.

production of these crops. BPS has been estimating annual production by kabupaten only for paddy, and by province for all other food crops. With the government's autonomy policy, some provinces have expressed their need for kabupaten level data for all crops. The procedure in the past year or so has been the following: if a province/kabupaten has the budget to finance coverage of additional households, then the procedure applied nationally would then be applied to those kabupaten; if no budget is available then nothing more is done and consequently no representation would be available for that kabupaten for anything other than paddy.

The crop-cutting survey has been conducted annually for more than 10 years using the same procedure. The sample every year covers around 45,000 plots with an area of 2,5 x 2,5 m^2. Since the same procedures are followed, rolling samples can easily be applied over, for example, three years to produce kabupaten level data. If this is done, then we would have a cumulative sample of 135,000 areas, which would probably be more than sufficient for producing yield data per kabupaten.

III. ILLUSTRATIVE CONSEQUENCES OF DESIGN

If BPS decides to apply the rolling sample design discussed in the previous section to the above four surveys, then the annual timetable that it will face would be similar to that illustrated in Table 1.

Table 1
Annual Timetable for Fieldwork if Rolling Sample Design was Applied

Period of Survey	Survey			
	Susenas	Sakernas	Susi	Crop-cutting
January				X
February	X	X	X	X
March				X
April				X
May		X	X	X
June				X
July				X
August		X	X	X
September				X
October				X
November		X	X	X
December				X

Table 2 shows the current status, in terms of sample size, for the four surveys. Table 3 shows the effect, on the sample size, of combining data over three years.

Table 2
Current Sample Size for the Four Main BPS Surveys

Province	Susenas (households)		Sakernas (households per quarter)	Susi (establishments)	Crop-cutting (plots per year)
	Core	Module			
(1)	(2)	(3)	(4)	(5)	(6)
11. DI Aceh [1])	480	-	-	1309	376
12. Sumatera Utara	13120	2944	880	5889	3542
13. Sumatera Barat	8640	1792	536	1318	1843
14. Riau	9472	1664	496	1084	609
15. Jambi	6080	1152	344	1017	579
16. Sumatera Selatan	4992	1824	656	1478	776
17. Bengkulu	2432	1024	304	779	384
18. Lampung	7008	2176	656	1403	1893
19. Bangka Belitung	1856	800	128	- [2])	490
31. DKI Jakarta	6080	3072	920	4443	369
32. Jawa Barat	18112	7200	2256	5302	5456
33. Jawa Tengah	25248	7552	2264	9354	7113
34. DI Yogyakarta	3456	2304	688	2181	978
35. Jawa Timur	29024	8832	2648	8734	7854
36. Banten	4864	1920	480	- [3])	1176
51. Bali	5728	1920	576	1901	1101
52. Nusa Tenggara Barat	4512	2176	656	1511	1217
53. Nusa Tenggara Timur	8512	1792	536	1247	711
61. Kalimantan Barat	5472	1920	576	842	495
62. Kalimantan Tengah	3680	1152	344	1244	450
63. Kalimantan Selatan	6688	1792	536	1477	1767
64. Kalimantan Timur	6976	1152	344	961	279
71. Sulawesi Utara	3168	1152	400	1102	304
72. Sulawesi Tengah	4992	1152	344	975	816
73. Sulawesi Selatan	15328	2304	688	1589	3606
74. Sulawesi Tenggara	3712	1152	344	839	390
75. Gorontalo	2016	800	184	- [4])	224
81. Maluku [1])	416	-	-	781	58
82. Maluku Utara [1])	416	-	-		58
91. Irian Jaya [1])	1664	-	-	792	86
Total	214144	62720	18784	59482	45000

[1]) Critical areas, reduce samples
[2]) Integrated with Sumatera Selatan
[3]) Integrated with Jawa Barat
[4]) Integrated with Sulawesi Utara

Table 3
Sample Size for the Four Main BPS Surveys if Data were Combined for 3 Years

Province	Susenas (households)		Sakernas (households per quarter)[5]	Susi (establishments)[5]	Crop-cutting (plots per year)[5]
	Core	Module			
(1)	(2)	(3)	(4)	(5)	(6)
11. DI Aceh [1])	1440	-	-	3927	1128
12. Sumatera Utara	39360	8832	2640	17667	10629
13. Sumatera Barat	25920	5376	1608	3954	5529
14. Riau	28416	4992	1488	3252	1827
15. Jambi	18240	3456	1032	3051	1737
16. Sumatera Selatan	14976	5472	1968	4434	2328
17. Bengkulu	7296	3072	912	2337	1152
18. Lampung	21024	6528	1968	4209	5679
19. Bangka Belitung	5568	2400	384	-	1470
31. DKI Jakarta	182240	9216	2760	13329	1107
32. Jawa Barat	39336	21600	6768	15906	16368
33. Jawa Tengah	75744	22656	6792	28062	21339
34. DI Yogyakarta	10368	6912	2064	6543	2934
35. Jawa Timur	87072	26496	7944	26202	23562
36. Banten	14592	5760	1440	-	3528
51. Bali	17184	5760	1728	5703	3303
52. Nusa Tenggara Barat	13536	6528	1968	4533	3651
53. Nusa Tenggara Timur	25536	5376	1608	3741	2313
61. Kalimantan Barat	16416	5760	1728	2526	1485
62. Kalimantan Tengah	11040	3456	1032	3732	1350
63. Kalimantan Selatan	20064	5376	1608	4431	5301
64. Kalimantan Timur	20928	3456	1032	2883	837
71. Sulawesi Utara	9504	3456	1200	3306	912
72. Sulawesi Tengah	14976	3456	1032	2925	2448
73. Sulawesi Selatan	45984	6912	2064	4767	10818
74. Sulawesi Tenggara	11136	3456	1032	2517	1170
75. Gorontalo	6036	2400	552	-	672
81. Maluku [1])	1248	-	-	2343	174
82. Maluku Utara [1])	1248	-	-	-	174
91. Irian Jaya [1])	4992	-	-	2376	258
Total	642432	188160	56362	178446	135000

[1]) Critical areas, reduce samples
[2]) Integrated with Sumatera Selatan
[3]) Integrated with Jawa Barat
[4]) Integrated with Sulawesi Utara
[5]) These figures cover the overlapping portions of the samples.

Instead of the current annual sample size of 214,144 households, the Susenas "core" survey would cover a sample of 642,432 households. Similarly, the Susenas "module" survey would allow coverage of 188,160 instead of the current 62,720 households. For the crop cutting survey, the rolling sample design would substantially increase coverage at the commodity level, as a comparison of Tables 4 and 5 shows.

Table 4
Current Sample Size of Crop-cutting Survey by Type of Crops

Province	Paddy	Maize	Cassava	Sweet Potatoes	Peanut	Soybean	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
11. DI Aceh	280	20	0	0	16	60	376
12. Sumatera Utara	2550	551	150	60	107	124	3542
13. Sumatera Barat	1506	127	83	16	65	46	1843
14. Riau	477	50	42	0	8	32	609
15. Jambi	406	50	45	9	17	52	579
16. Sumatera Selatan	583	62	52	11	29	39	776
17. Bengkulu	221	66	22	28	25	22	384
18. Lampung	722	567	379	42	52	131	1893
19. Bangka Belitung	258	25	142	39	19	7	490
31. DKI Jakarta	360	0	9	0	0	0	369
32. Jawa Barat	4165	355	363	144	221	208	5456
33. Jawa Tengah	3882	1490	706	107	412	516	7113
34. DI Yogyakarta	322	177	162	6	140	171	978
35. Jawa Timur	3313	2450	650	100	414	927	7854
36. Banten	922	55	57	38	76	28	1176
51. Bali	595	183	101	40	76	106	1101
52. Nusa Tenggara Barat	713	88	48	18	61	289	1217
53. Nusa Tenggara Timur	249	257	129	21	35	20	711
61. Kalimantan Barat	418	41	24	0	0	12	495
62. Kalimantan Tengah	357	36	37	0	6	14	450
63. Kalimantan Selatan	1421	109	93	16	84	44	1767
64. Kalimantan Timur	219	28	17	6	6	3	279
71. Sulawesi Utara	160	70	19	12	18	25	304
72. Sulawesi Tengah	605	69	47	21	37	37	816
73. Sulawesi Selatan	2110	895	243	72	169	117	3606
74. Sulawesi Tenggara	184	111	40	7	24	24	390
75. Gorontalo	96	59	15	3	21	30	224
81. Maluku	10	15	33	0	0	0	58
82. Maluku Utara	22	12	8	6	3	3	58
91. Irian Jaya	35	8	4	39	0	0	86
Total	27161	8026	3724	861	2141	3087	45000

Table 5
Sample Size of Crop-cutting Survey by Type of Crops if Data were Combined for 3 Years

Province	Paddy	Maize	Cassava	Sweet Potatoes	Peanut	Soya bean	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
11. DI Aceh	840	60	0	0	48	180	1128
12. Sumatera Utara	7650	1653	450	180	321	372	10629
13. Sumatera Barat	4518	381	249	48	195	138	5529
14. Riau	1431	150	126	0	24	96	1829
15. Jambi	1218	150	135	27	51	156	1737
16. Sumatera Selatan	1749	186	156	33	87	117	2328
17. Bengkulu	663	198	66	84	75	66	1152
18. Lampung	2166	1701	1137	126	156	393	5679
19. Bangka Belitung	774	75	426	117	57	21	1470
31. DKI Jakarta	1080	0	27	0	0	0	1107
32. Jawa Barat	12495	1065	1089	432	663	624	16368
33. Jawa Tengah	11646	4470	2118	321	1236	1548	21339
34. DI Yogyakarta	966	531	486	24	420	513	2934
35. Jawa Timur	9939	7350	1950	300	1242	2781	23562
36. Banten	2766	165	171	114	228	84	3528
51. Bali	2088	549	303	120	228	318	3303
52. Nusa Tenggara Barat	2139	264	144	54	183	867	3651
53. Nusa Tenggara Timur	747	771	387	63	105	60	2133
61. Kalimantan Barat	1254	123	72	0	0	36	1485
62. Kalimantan Tengah	714	108	111	0	24	42	1350
63. Kalimantan Selatan	4263	327	279	48	256	132	5301
64. Kalimantan Timur	657	84	51	24	24	9	837
71. Sulawesi Utara	480	210	57	36	54	75	912
72. Sulawesi Tengah	1815	207	141	63	111	111	2448
73. Sulawesi Selatan	6330	2685	729	216	507	351	10818
74. Sulawesi Tenggara	552	333	120	21	72	72	1170
75. Gorontalo	288	177	45	9	63	90	672
81. Maluku	30	45	99	0	0	0	174
82. Maluku Utara	66	36	24	18	9	9	174
91. Irian Jaya	105	24	12	117	0	0	258
Total	81483	24078	11172	2583	6423	9261	135000

Increased sample size would have a significant impact on coverage at the kabupaten level, increasing the precision of kabupaten estimates. Tables 6 and 7 illustrate that impact in the case of the kabupatens in the province of Lampung (on the Susenas "core" and "module" surveys repectively): samples would become three times or five times the size of the current yearly samples, making sampling fractions 3 (or 5) times bigger and consequently the inflation factors 1/3 (or 1/5) the current size.

Table 6
Cumulative Households in Lampung Province in Core Susenas 2002 if Rolling Samples Were Used

Kabupaten/Kota	Yearly Sample Households	First Alternative (three year rolling)	Second Alternative (five year rolling)
(1)	(2)	(3)	(4)
1. Kab. Lampung Barat	640	1920	3200
2. Kab. Tanggamus	608	1824	3040
3. Kab. Lampung Selatan	544	1632	2720
4. Kab. Lampung Timur	960	2880	4800
5. Kab. Lampung Tengah	960	2880	4800
6. Kab. Lampung Utara	544	1632	2720
7. Kab. Way Kanan	544	1632	2720
8. Kab. Tulang Bawang	608	1824	3040
9. Kota Bandar Lampung	640	1920	3200
10. Kota Metro	960	2880	4800
Total	7008	21024	35040

Table 7
Cumulative Households in Lampung Province in Module Susenas 2002 if Rolling Samples Were Used

Kabupaten/Kota	Yearly Sample Households	Alternative 1 (3 year rolling)	Alternative 2 (5 year rolling)
(1)	(2)	(3)	(4)
1. Kab. Lampung Barat	128	384	640
2. Kab. Tanggamus	256	768	1280
3. Kab. Lampung Selatan	352	1056	1760
4. Kab. Lampung Timur	288	864	1440
5. Kab. Lampung Tengah	352	1056	1760
6. Kab. Lampung Utara	160	480	800
7. Kab. Way Kanan	96	288	480
8. Kab. Tulang Bawang	224	672	1120
9. Kota Bandar Lampung	256	768	1280
10. Kota Metro	64	192	320
Total	2176	6528	10880

Table 8 illustrates the effect of applying rolling samples to the crop-cutting survey in two provinces: Jawa Timur and Nusa Tenggara Barat.

Table 8
Crop-cutting Samples by Type of Crop in Two Provinces

Type of crops	Jawa Timur	Nusa Tenggara Barat
(1)	(2)	(3)
1.Paddy	3313	713
2. Maize	2450	88
3. Cassava	650	48
4. Sweet Potatoes	100	18
5. Peanut	414	61
6. Soyabean	927	289
Total	7854	1217

At the kecamatan level, the impact would be significant as well, as illustrated in Table 9 in the case of the Susenas "core" for two selected kabupaten/kotas.

Table 7
Cumulative Households in Kabupaten Lampung Selatan and Kota Metro in Core Susenas 2002 if Rolling Samples Were Used

Kabupaten/Kota	Kecamatan	Yearly Sample Household	Alternative 1 (3 year rolling)	Alternative 2 (5 year rolling)
(1)	(2)	(3)	(4)	(5)
1. Kab. Lampung Selatan	1. Padang Cumin	64	192	320
	2. Kedundung	32	96	160
	3. Gedung Tataan	32	96	160
	4. Negeri Katon	32	96	160
	5. Tegineneng	32	96	160
	6. Natar	64	192	320
	7. Jati Agung	32	96	160
	8. Tanjung Bunting	32	96	160
	9. Katibung	64	192	320
	10. Sidumulyo	48	144	240
	11. Kalianda	48	144	240
	12. Padas	32	96	160
	13. Panengahan	32	96	160
	Total	544	624	2720
2. Kota Metro	1. Bantul	208	624	1040
	2. Metro Raya	752	2256	3760
	Total	960	2880	4800

IV. CONCLUSIONS

The main conclusion of this report is this: given the likelihood of continuing budget constraints for BPS and the increasing demand for lower-level geographic units (e.g. kabupaten/kota or kecamatan), the use of rolling sample offers a feasible cost-effective solution. The report showed that this methodology can be safely applied to four important BPS surveys: Susenas, Sakernas, Susi and crop cutting. If multi-year data were combined, then the resulting cumulative sample can provide adequate representation at the kabupaten level, and in some cases at even the kecamatan level.

LINKING DESIGN AND ORGANIZATION IN THE AGRICULTURE CENSUS SAMPLE

Report # 58

by
Sri Budianti Sukmadi

August, 2002

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

- 3 with the possibility of integration of activities in only one group (areas of A, B and C with no overlap)
- 3 with the possibility of integrating activities of 2 groups (areas AB, AC and BC) and
- 1 with no possibility of integration of activities (area ABC).

Then ideally one would plan workload, training and enumeration to cover to the extent possible areas with no overlap, or at most areas with 2 overlaps, and avoid to the extent possible areas with 3 overlaps to minimize the possibility of having one CB enumerated by more than one enumerator.

C. Consistency with SP2000 Figures

Results of the pilot study in Sulwesi Selatan show a high consistency in sub-sectoral distribution with results of the 2000 population census. As Table 5 shows, the number of CB's in Food Crops are very similar, but for the other sub-sectors the SP2000 includes less CB's than the pilot study. One reason for the discrepancy between the pilot and population census results may be that the SP2000 classifies households according to their main activity only whereas the pilot study included other activities as well.

Table 5
Number of CB's by Sub-sector: SP2000 vs ST2003 Pilot Study

Sub-sector ST2003	Sub-sector SP2000				All sub-sectors
	Food crops *)	Estate crops	Animal husbandry	Fishery	
1. Paddy	237	66	76	77	238
2. Other main crops	238	66	76	77	239
3. Horticulture	223	64	74	76	225
4. Estate crops	112	47	47	40	114
5. Animal husbandry	215	63	72	73	216
All sub-sectors	238	66	76	77	240

*) Includes paddy, other main food crops, and horticulture
Source: Sulawesi Selatan Pilot Study. The number of CB's covered was 240

If we examine numbers more closely, we realize that integration will allow substantial cross-subsectoral coverage. Table 6, for example, shows that the 238 Paddy CB's not only cover paddy, but also other main crops; 223 of them also cover horticulture; 113 cover estate crops and 216 cover animal husbandry. The same applies to other sub-sectoral CB's.

Table 6
Number of CB's by Sub-Sectoral Allocation and Coverage

CB Allocation by Sub-sector	CB Coverage by Sub-sector				
	Paddy	Other main crops	Horticulture	Estate crops	Animal husbandry
1. Paddy CB	238	238	223	113	216
2. Other main crops CB	238	239	224	113	216
3. Horticulture CB	223	224	225	110	206
4. Estate crops CB	113	113	110	114	109
5. Animal husbandry CB	216	216	206	109	216

Total number of CB's: 240
Source: Sulawesi Selatan Pilot Study 2003

III. SURVEY ORGANIZATION

A. Time & Budget Constraints

Given the timetable mentioned above, it is important to ensure that a sample list (with a clear identification of locations and sub-sectors) is fixed before regional offices start recruiting field workers. Such a list can be prepared with the help of a matrix of location by sub-sector. This provides two major advantages:

- It can be used for grouping field worker training and for allocating the work load of enumerators.
- It will allow integration of training since census materials for sub-sectors have similar concepts and definitions, and most sub-sector questionnaires have the same core questions.

These can provide substantial savings in time and budget.

B. Sample Selection Stages

Sample selection involves four stages:

- Evaluation of results of the 2001 and 2002 pilot studies to ensure the feasibility of constructing a frame, designing the sample and to determine the needed sample sizes.
- Checking the completeness of CB's as a basic frame and incorporating SP2000-L2 results, particularly if the paddy sub-sector is conducted in 2003 and integrated with ST2003-L1 and L2.
- Selection of paddy CB's with probability proportional to size by food crop agricultural households derived from SP2000-L2. Then in the selected paddy CB's the paddy agricultural households will

be selected from the frame of ST2003-L2 and enumerated directly with detailed questionnaires (long forms). Sample selection of CB's is done by the BPS Central Office, while selection of paddy agricultural households is done by regional offices. Based on the sample list, regional offices will have to give a priority for the selected CB's to enumerate first with ST2003-L1. The frame of paddy agricultural households is then derived from ST2003-L2 results.

- Construction of a master frame for the non-paddy sub-sector CB's is also derived from ST2003-L2. This frame will be used in 2004 for the non-paddy sub-sector census sample. Listing households from selected non-paddy CB's will be done in 2004 to produce a frame of non-paddy agricultural households. Selection of non-paddy agricultural households is done directly after completing the frame. The number of non-paddy agricultural households to be selected will need to be determined earlier by the BPS Central Office.

The stages of sample selection can change if the census sample for the paddy sub-sector is to be implemented in 2004 and integrated with the other sub-sectoral census sample. Figures 1 and 2 provides an organizational scheme for various stages involved in sample selection under two alternative scenarios: the first assumes that paddy is covered in 2003 and the other that all sub-sectors are covered in 2004.

Figure 1: Stages of Sample Selection ST03 (Paddy in 2003 and Non Paddy in 2004)

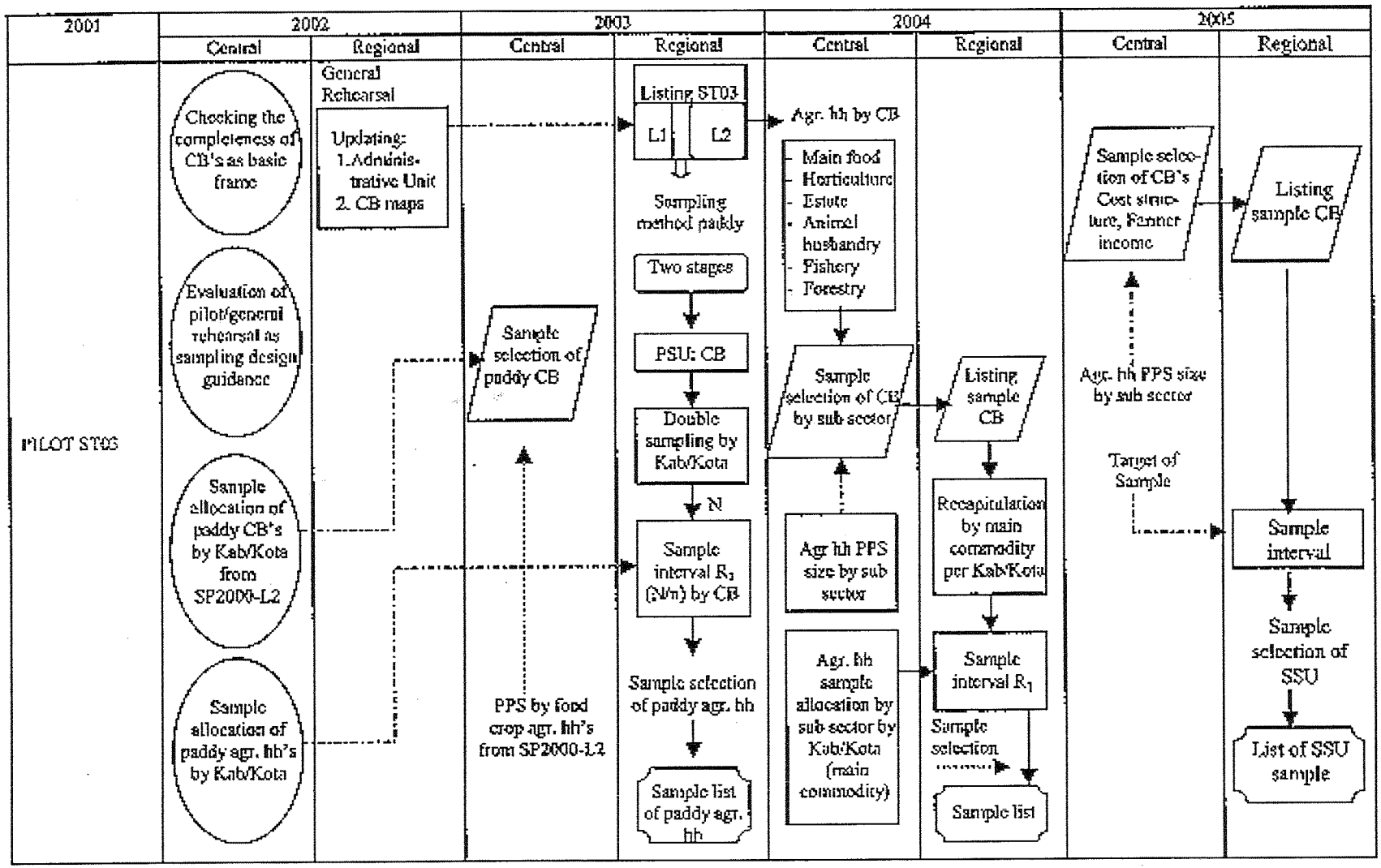


Figure 2: Stages of Sample Selection ST03 (All Sub-Sectors in 2004)

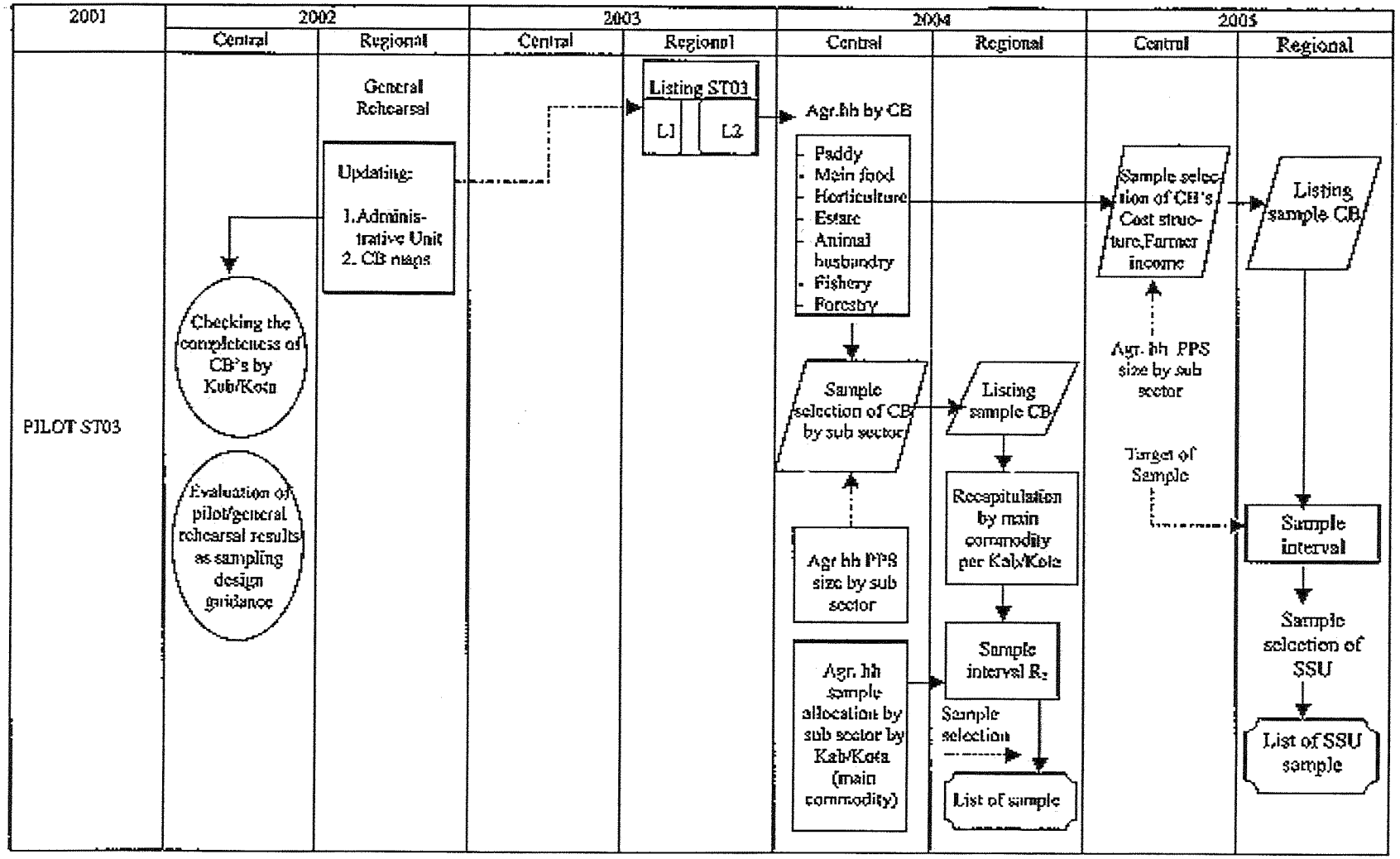


Figure 3 shows the flow of activities under the first alternative and Figure 4 shows it under the second alternative.

Figure 3
Activity Flow Under First Alternative ST03-L1 and L2

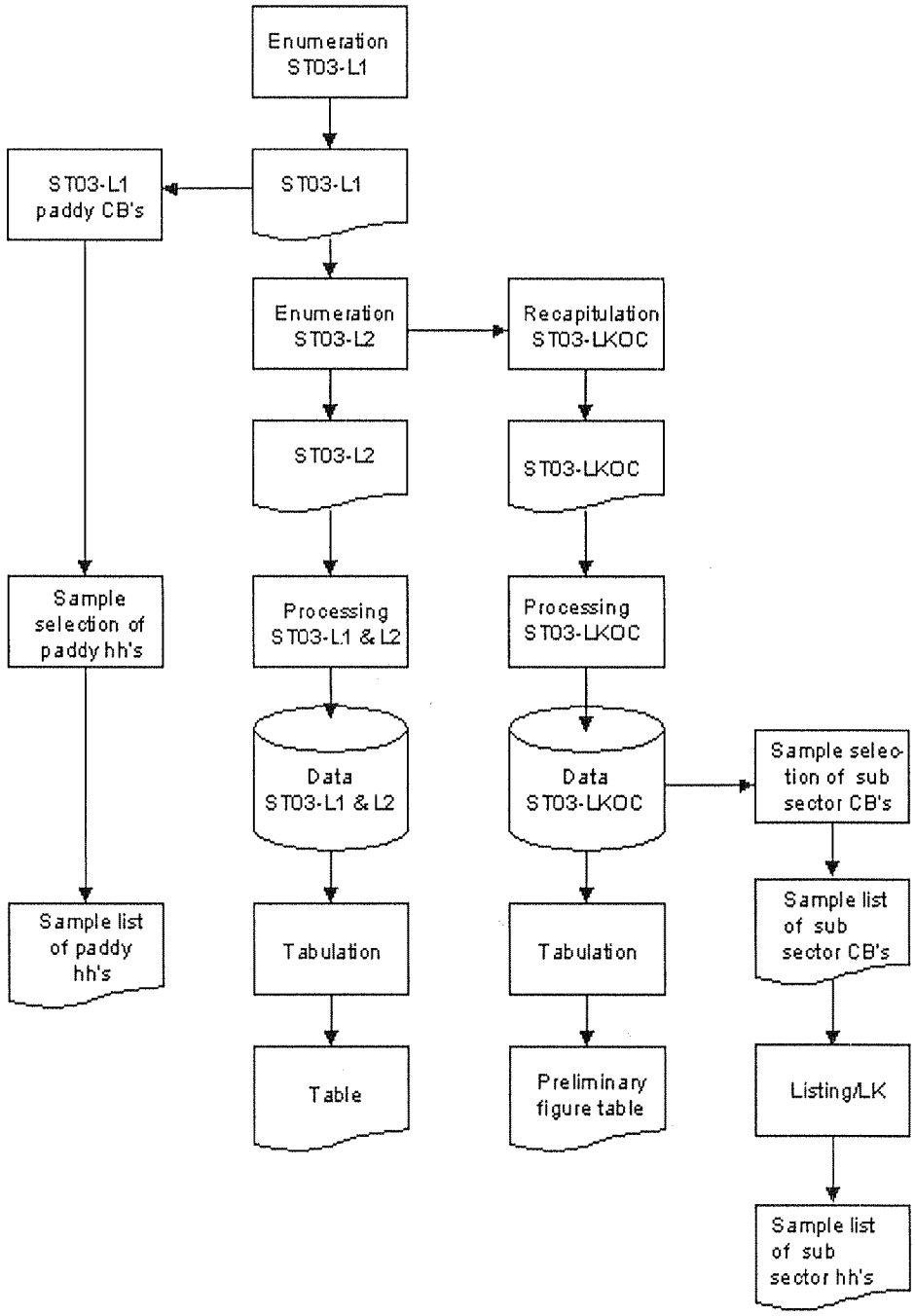
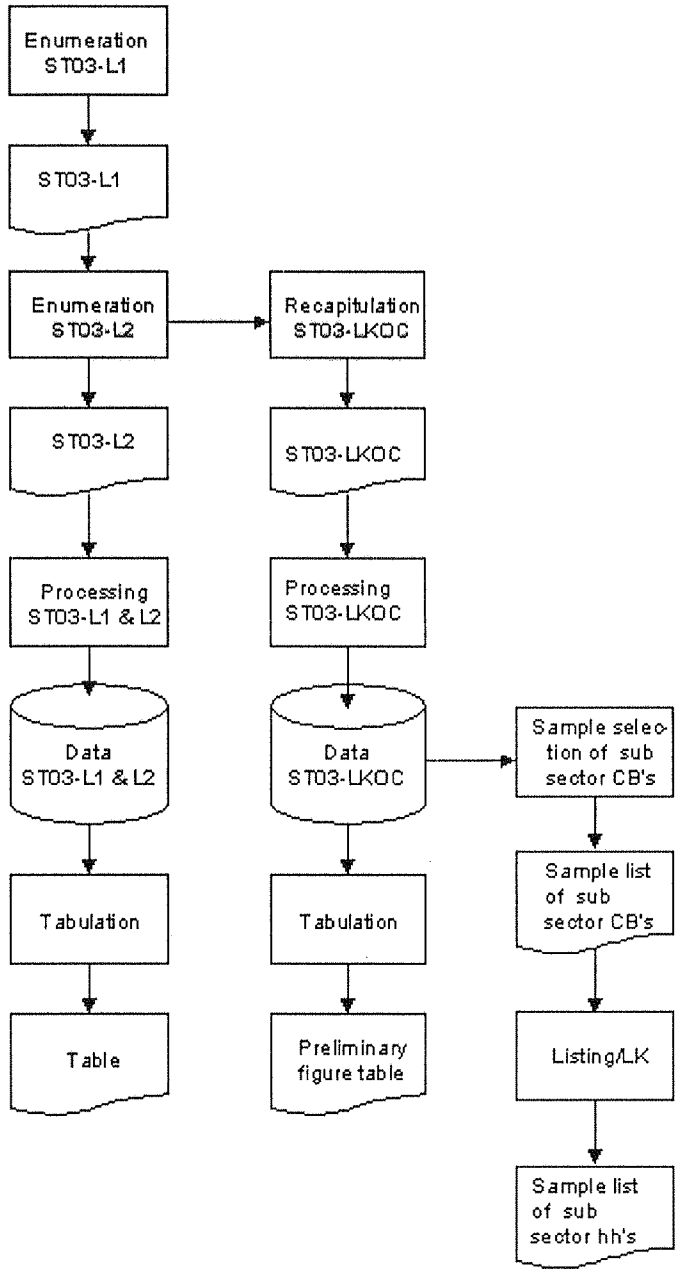


Figure 4
Activity Flow Under Second Alternative ST03-L1 and L2

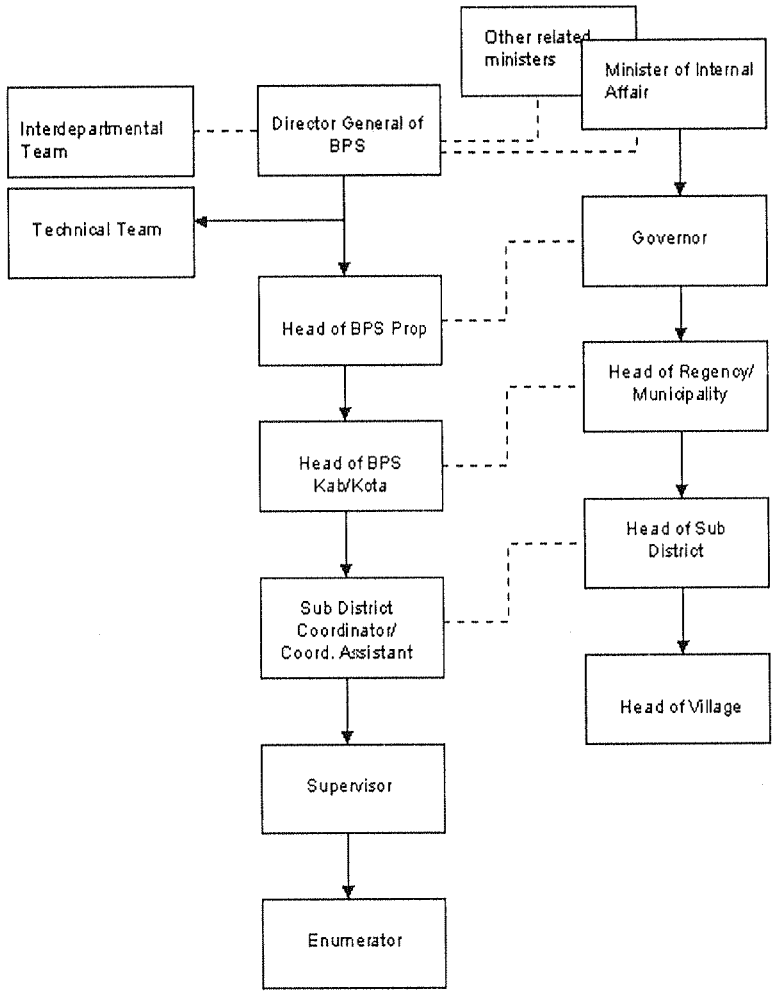


C. Field Operations

Figure 5 presents an overall organizational chart of field operations. To conduct these activities, BPS needs support from institutions such as the Ministry of Internal Affairs, the Ministry of Agriculture, the Ministry of Forestry and other related agencies at the national and provincial levels. This includes not only administrative support, but also technical support from professionals with long experience in their fields. That is why BPS appointed inter-departmental

and technical teams including members from BPS and other institutions from the national down to the desa/kelurahan level. While small kecamatan can function adequately with one Mantri Statistik, large ones may need additional people.

Figure 5
ST03 Field Organization

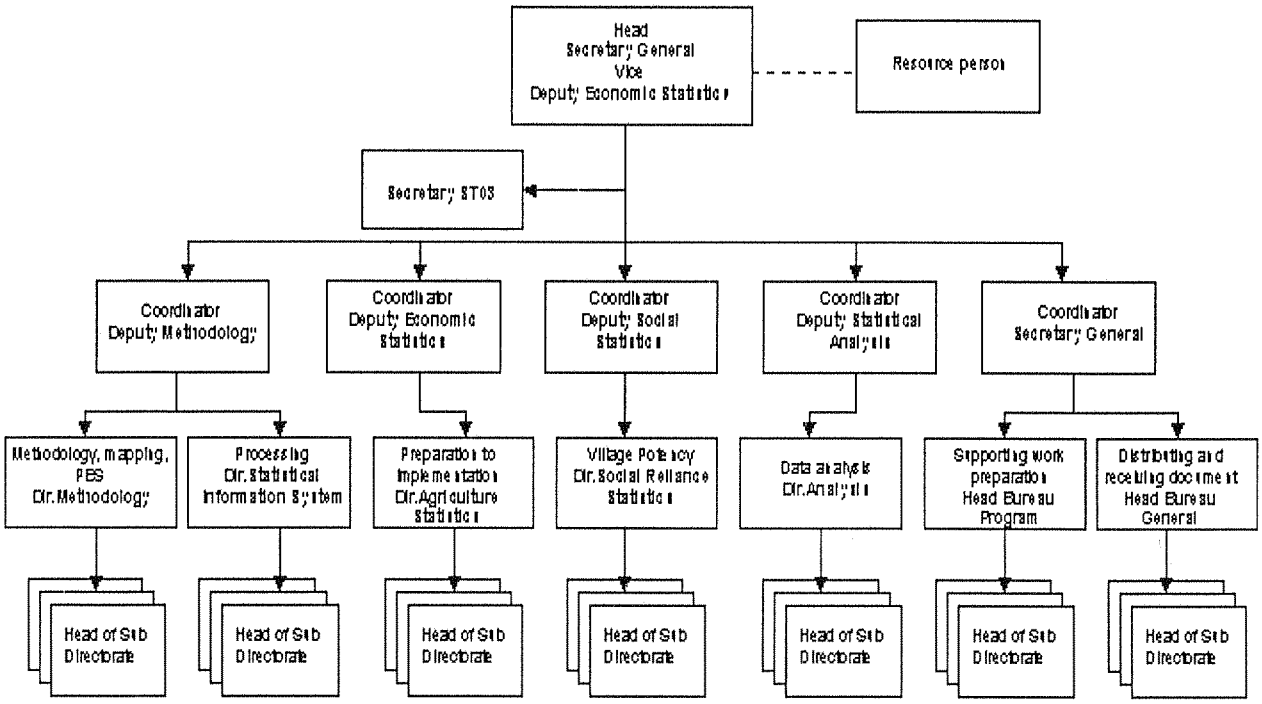


D. Technical Team

The organizational chart of the technical team responsible for census activities is provided in Figure 6. Note that the chart mirrors closely the organizational structure of BPS. The head of the technical team is the BPS Secretary General. He is responsible for coordinating all activities. He is assisted by four deputies with respective teams under them covering particular technical areas. A Secretariate is responsible for monitoring implementation of daily activities.

The person responsible for developing methodologies is the Director for Methodology. His job includes constructing the master frame, sample design, concept definition and classification, and mapping of administrative units and CB's.

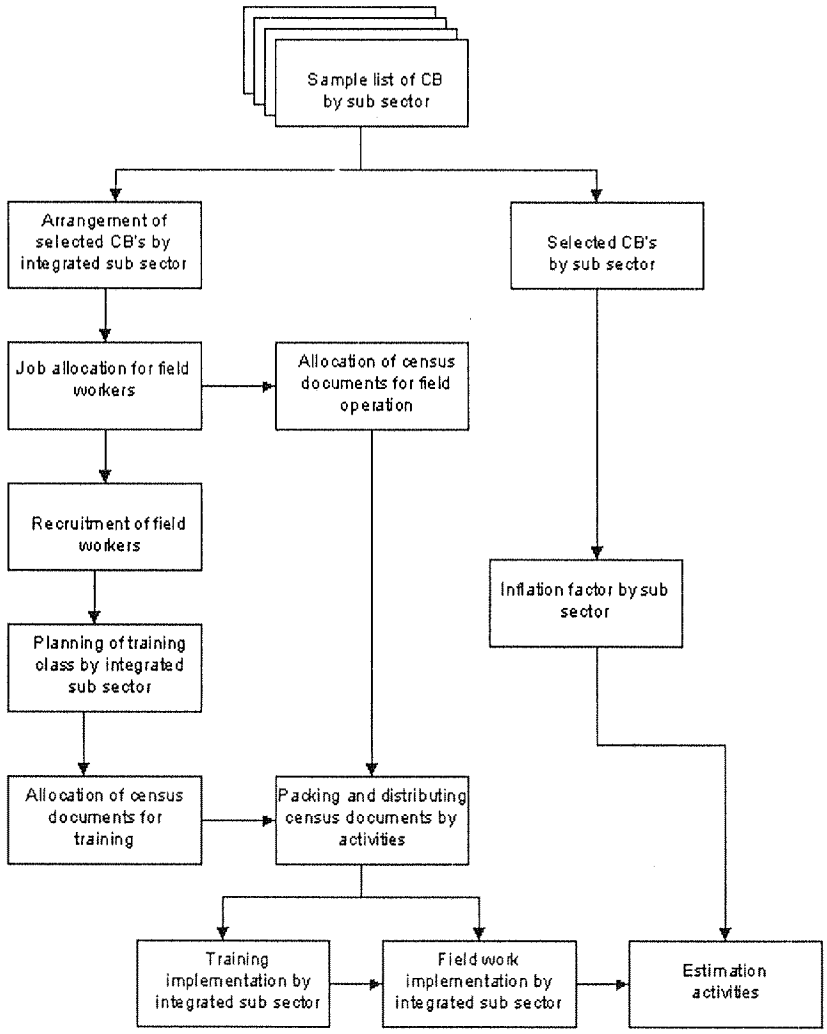
Figure 6
Technical Team ST03



E. Link with Sample Design

The sample list by sub-sector can be converted into a sample list by integrated sub-sector, following the methodology described in Section II.B above. Then activities such as enumerator training, enumerator distribution and document allocation can be be planned in a way that minimizes response/enumeration burdens. A schematic representation of the link between sample design and survey organization is presented in Figure 7.

Figure 7
Linkage between Sampling Design and Survey Organization



I. FURTHER TO SUSI ESTIMATION PROCEDURE

This note clarifies the application of the estimation procedure in SUSI, and gives the relevant variance estimation formulae.

Step 1: The first step is to modify the weights of SUSI units such that the sample areas selected into SUSI give the same estimates *of employment measured at the time of the census* as the total census values of employment by ISIC and province. Rather than applying ratio adjustments after the SUSI results have been tabulated, this procedure achieves exactly the same results by making appropriate adjustment to the sample weights. The advantage of this alternative approach is its great operational convenience.

Step 2: Using the weights (w) redefined in step 1 (rather than the original SUSI design weights), repeat the procedure for the *number of establishments measured at the time of the census* by ISIC and province. This gives an alternative set of weights (e).
(Note. The major shortcomings of the SUSI design for detailed domains is expected to be taken care of by step 1; hence the further adjustment from weights w from step 1 to weights e in step 2 is expected to be generally small. The further adjustment arises only from the fact that the average establishment size in SUSI areas, as measured at the time of the census, may not be identical to that average size when all the census areas are considered.)

Step 3: SUSI estimates of employment by province, ISIC, their combination, or any other domains are made using weights (w) defined from step 1. There is no further ratio adjustment required. Similarly, SUSI estimates of number of establishments by province, ISIC, their combination, or any other domains are made using weights (e) defined from step 2.

Step4: All other characteristics, measured in SUSI but not available (or available reliably) from the census may be divided into two groups: (i) characteristics which are more closely related to employment, and (ii) characteristics more closely related to the number of establishments. The first group is likely to be much larger and more important, and include measures such as output, value-added, turnover.... The second group may include characteristics which are meaningfully defined only at the establishment level, such as ownership, age, distribution by size ...

Step 5: For group (i) compute the ratio to employment, using weights (w) from step 1. For group (ii) compute the ratio to number of establishments, using weights (e) from step 2.

A. Step 1

Incorporating Census Data on Employment by Province/ISIC into Sampling Weights

Let

k province

h one or more geographical strata within province (if applicable)

j sample area

i ISIC

$w^{(1)}_{khji}$ sample design weight (from SUSI design, by ISIC and area)

f_{khji} sampling rate for enumerating establishments of ISIC i WITHIN SUSI area j
= number listed / number successfully enumerated in the area

$y^{(1)}_{khji}$ total employment of enumerated establishments within the area (SUSI)

y_{khji} estimate of total employment by ISIC in the area (SUSI) = $y^{(1)}_{khji} / f_{khji}$

$w^{(2)}_{khji}$ sample design weight for area j ISIC i = $f_{khji} * w^{(1)}_{khji}$

x_{khji} total employment by ISIC in the area (CENSUS)

x_{ki} estimate of total employment by ISIC in province k ISIC i (CENSUS)
= $_{h,j}(w^{(2)}_{khji} \cdot x_{khji})$

X_{ki} actual total employment by ISIC in province k ISIC i (CENSUS)

x, X the above summed over all ISIC and provinces

C overall adjustment ratio X /x

C_{ki} ratio adjustment for province k ISIC i (on the basis of census employment)
[initially computed as X_{ki} / x_{ki} then adjusted iteratively to make the SUSI estimates of y_{ki} by province and ISIC additive so as to give the overall estimate X, as explained in a separate report; also, these factors may be subject to maximum and minimum limits as noted in that report]

w_{khji} 'employment adjusted' sample design weight for area j ISIC i = $C_{ki} * w^{(2)}_{khji}$
[these adjusted weights ensure that overall $(w_{khji} \cdot x_{khji}) = X$ and also that approximately $_{h,j}(w_{khji} \cdot x_{khji}) = X_{ki}$ in the province-by-ISIC matrix; the approximation comes only from the additivity constraint mentioned above]

Y_d estimate of employment for ANY domain 'd'
= $[\ _d(w_{khji} \cdot y_{khji}) /\ _d(w_{khji} \cdot x_{khji})] * X_d$
[a 'domain' corresponds to any combination of provinces/strata/areas and ISIC's]
[for all Indonesia all ISIC, the denominator in the first term is identical to the second term; it is nearly so for most of the individual province-by-ISIC cells; and generally for any other reporting domain with some approximation]

Variance estimation

Variance of Y_d is obtained as for a normal ratio estimate, with usual assumptions:

$var(Y_d) = X^2_d * var(r_d)$

with

$x_d = {}_d(w_{khji} \cdot x_{khji})$,
$r_d = [\ {}_d(w_{khji} \cdot y_{khji}) / \ {}_d(w_{khji} \cdot x_{khji})] = y_d / x_d$

$$z_{hij \in d} = \frac{1}{x_d} \cdot \left(y_{hji} - r_d \cdot x_{hji}\right), \quad z_h = \Sigma_{ji \in h} z_{hji}$$

[sum "ji h" refers to the sum over all ISIC's i in the domain of interest, and over all sample areas in geographical stratum h. In general, a domain may be divided into a number of such strata, depending on SUSI sample design]

and

$$var(r_d) = \sum_{h \in d} \left[(1 - f_h) \cdot \frac{a_h}{a_h - 1} \cdot \Sigma_{ji} \left(z_{hji} - \frac{z_h}{a_h} \right)^2 \right].$$

B. Step 2

Further Adjustment of Weights to Match the Census Number of Establishments

In addition to the above symbols, let

$v^{(1)}_{khji}$ total number of establishments enumerated within the area (SUSI)

v_{khji} estimate of number of establishments by ISIC in the area (SUSI)
$= y^{(1)}_{khji} / f_{khji}$

u_{khji} total number of establishments by ISIC in the area (CENSUS)

u_{ki} estimate of total number of establishments by ISIC in province k ISIC i (CENSUS)
$= {}_{hj}(w_{khji} \cdot u_{khji})$
[note that the weights used are 'employment adjusted' weights as defined in the previous step]

U_{ki} actual total number of establishments by ISIC in province k ISIC i (CENSUS)

u, U the above summed over all ISIC and provinces

D overall adjustment ratio U /u

D_{ki} ratio adjustment for province k ISIC i (on the basis of census number of establishments) [initially computed as U_{ki} / u_{ki} then adjusted iteratively so as to make the SUSI estimates of v_{ki} by province and ISIC additive so as to give the overall estimate U, as explained in a separate report; also, these factors may be subject to maximum and minimum limits as

noted in that report]

e_{khji} further or 'establishment adjusted' sample design weight for area j ISIC i
$= D_{ki} * w_{khji}$
[these adjusted weights ensure that overall $(e_{khji} \cdot u_{khji}) = U$ and also that approximately $_{hj}(e_{khji} \cdot u_{khji}) = U_{ki}$ in the province-by-ISIC matrix; the approximation comes only from the additivity constraint mentioned above]

V_d estimate of employment for ANY domain 'd'
$= [\ _d(e_{khji} \cdot v_{khji}) / \ _d(e_{khji} \cdot u_{khji})] * U_d$
[for all Indonesia all ISIC, the denominator in the first term is identical to the second term; it is nearly so for most of the individual province-by-ISIC cells; and generally for any other reporting domain with some approximation]

Variance estimation

Variance of V_d is obtained in exactly the same way as that for Y_d above.

$var(V_d) = U^2_d * var(r_d)$

with

$r_d = [\ _d(e_{khji} \cdot v_{khji}) / \ _d(e_{khji} \cdot u_{khji})]$,

$$z_{hij\in d} = \frac{1}{u_d}.\left(v_{hji} - r_d.u_{hji}\right); \quad z_h = \Sigma_{ji} z_{hji}$$

and

$$var(r_d) = \sum_{h\in d}\left[(1-f_h).\frac{a_h}{a_h - 1}.\Sigma_{ji}\left(z_{hji} - \frac{z_h}{a_h}\right)^2\right]$$

C. Step 3

Estimating Employment and Number of Establishments for Any Domain from SUSI

For estimating employment, weights from step 1 are used. For estimating the number of establishments in any domain, the modified weights from step 2 are used.

D. Step 4

Classification of All SUSI Characteristics According to Whether they are Expected to be More Closely Related to Employment or to the Number of Establishments

Statistical as well as substantive arguments may be invoked for this classification. There

should be no 'residual' category: if in doubt include the characteristics concerned in the 'employment dependent' group.

E. Step 5

SUSI Estimation for Any Characteristics Other Than Employment or Number of Establishments, for Any Domain d

A 'domain' may refer to province, ISIC, their cross-classification, or any other cell in the tabulation of SUSI results.

1. Characteristics Related to Employment

For any characteristic m related more closely to employment than to the number of establishments, estimate of employment for ANY domain 'd' is provided by:

$$M_d = [\ _d(w_{khji} \cdot m_{khji}) / \ _d(w_{khji} \cdot y_{khji})] * Y_d$$
$$= [r_d] * Y_d \text{, say.}$$

Here Y_d is the estimate of employment for domain d from SUSI, as in step 3, and w are the weights as defined in step 1.

Variance estimation

Variance for this estimate is obtained by the (very reasonable) assumption that the estimates of a ratio r_d and a total Y_d are un-correlated. This gives

$$\text{var}(M_d) = Y_d^2 . \text{var}(r_d) + r_d^2 . \text{var}(Y_d)$$

The second term is as defined in Step1. The first term is computed in the usual way for the ratio as follows.

Defining

$$z_{hij \in d} = \frac{1}{y_d} . (m_{hji} - r_d . y_{hji}); \quad z_h = \Sigma_{ji} z_{hji}$$

gives

$$var(r_d) = \sum_{h \in d} \left[(1 - f_h) . \frac{a_h}{a_h - 1} . \Sigma_{ji} \left(z_{hji} - \frac{z_h}{a_h} \right)^2 \right]$$

2. Characteristics Related More Closely to Number of Establishments

For any characteristic s related more closely to the number of establishments than to employment, estimate of employment for ANY domain 'd' is provided by:

$$S_d = [\ _d(e_{khji} \cdot s_{khji}) / \ _d(e_{khji} \cdot v_{khji})] * V_d$$
$$= [r_d] * V_d \text{, say.}$$

Here V_d is the estimate of number of establishments for domain d from SUSI, as in step 3, and e are the weights as defined in step 2.

Variance estimation

Variance for this estimate is obtained by the (very reasonable) assumption that the estimates of a ratio r_d and a total V_d are un-correlated. This gives

$$\text{var}\left(S_d\right) = V_d^2 . \text{var}\left(r_d\right) + r_d^2 . \text{var}\left(V_d\right)$$

The second term is as defined in step 2. The first term is computed in the usual way for the ratio as follows.

Defining

$$z_{hijed} = \frac{1}{y_d}.\left(s_{hji} - r_d.v_{hji}\right); \quad z_h = \Sigma_{ji} z_{hji}$$

gives

$$var(r_d) = \sum_{h \in d}\left[(1-f_h).\frac{a_h}{a_h - 1}.\Sigma_{ji}\left(z_{hji} - \frac{z_h}{a_h}\right)^2\right]$$

II. CUMULATING DATA OVER YEARS IN A HOUSEHOLD CONSUMPTION EXPENDITURE SURVEY

A. Introduction

A survey like Susenas may be able to produce annual estimates with sufficient reliability at the national level and for major provinces, but the sample is likely to be too small to produce sufficiently reliable annual estimates for small provinces and certainly in the case of individual kabupaten. *For this latter purpose it will be necessary to cumulate data over a number of years.* This note discusses some of the issues involved concerning sampling and estimation etc. in such cumulation.

Essentially, the idea is of a survey of modest size (modest for individual kabupaten or

other such domains) conducted on a continuous or regular basis, data from which can be *cumulated over years* to achieve more adequate sample sizes. To what extent can the data from individual years (rounds) be cumulated in such cases? Essential elements of the situation may be described in terms of the following illustration.

Suppose that the survey is conducted on a continuous or regular (say annual) basis with a representative sample of 1,000 households every year.[1] The sample is designed such that the information can be efficiently cumulated over time to achieve sufficient sample sizes, and the results are reported on a regular (annual) basis in the form of 'moving averages' over a number of most recent years.

The major advantage of cumulating data over time is that geographically more detailed information can be produced with a smaller sample size. However, possible disadvantages or problems should also be recognised. The lack of flexibility can be a possible disadvantage when the results have to be cumulated over a number of years. Major redesigns are more easily accommodated when the survey results are used one year at a time. By contrast, in using the same survey for a number of years, it is necessary to carefully regulate and control changes in content, design and procedures. This is for operational as well as substantive reasons. In any case, sudden or major changes have to be avoided so as not to disrupt the system. This requires that the survey is started only after very careful preparation and testing of all procedures. It is very difficult to rectify the situation once the survey is 'locked into' a particular design. In this sense, the situation is similar to that in a panel survey.

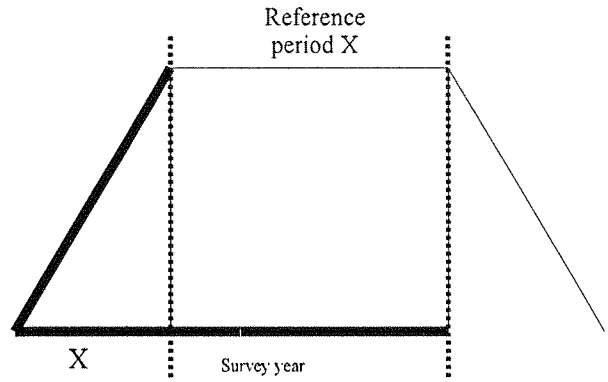
B. Some Basic Sampling Considerations

To cumulate the survey results over time, it is necessary that the sample be representative simultaneously over space and time. This means that for annual surveys for instance, the sample

---

1 The survey may be conducted regularly once a year, with field work concentrated over a relatively short period of time (as at present). An alternative and attractive model would be to conduct the survey on a continuous basis, with field work spread out through the year. During the year the workload may also distributed more or less uniformly, enumerating one-twelfth of annual sample size per month, in which case the work can be conducted by a much smaller team of interviewers deployed permanently for the task. With a continuous flow of data from the field, data preparation and processing also becomes an ongoing operation. Potential advantages of the design may be summarised as follows. The main operational advantage spring from the relatively moderate but regular workload. The number of survey staff required is reduced, who can then be better (and more cheaply) trained and supervised. Engagement on a permanent basis also tends to generate a sense of loyalty, responsibility and satisfaction, and is conducive to building more enduring survey capability. There can be major substantive advantages as well. The results can be regularly updated and reported in a more timely manner.
The points made in this note apply to both these situations: concentrated or distributed field work.

Time Period to Which the Data Pertain in 'normal' survey (Y=1)



It can be seen that with different values of the reference period for different types of items, the situation with cumulation over a number of years (Y>1) is similar in form to that with a conventional estimate over a single year (Y=1). Actually, with increasing Y, the period covered becomes less sensitive to differences in reference periods for different types of items in the same survey. This greater uniformity of the time-periods covered for different types of items is in fact an *advantage* of increasing Y (the period of cumulation).

For a *fixed reference period* such as the preceding calendar year, the situation is similar but simpler. For a conventional one-year survey, the period covered is centred at the middle of the calendar year. We have:

for Y =1: $P_M = 1$; $P_O = 1/2$.

More generally, with cumulation over Y>1 years we have,

$P_M = (P+1)/2$; $P_O = P/2$,

giving, for instance,

P_M = mid-point of the second year when cumulated over Y = 3, and
P_M = beginning of the second year when cumulated over Y = 4 years.

D. Price Adjustment

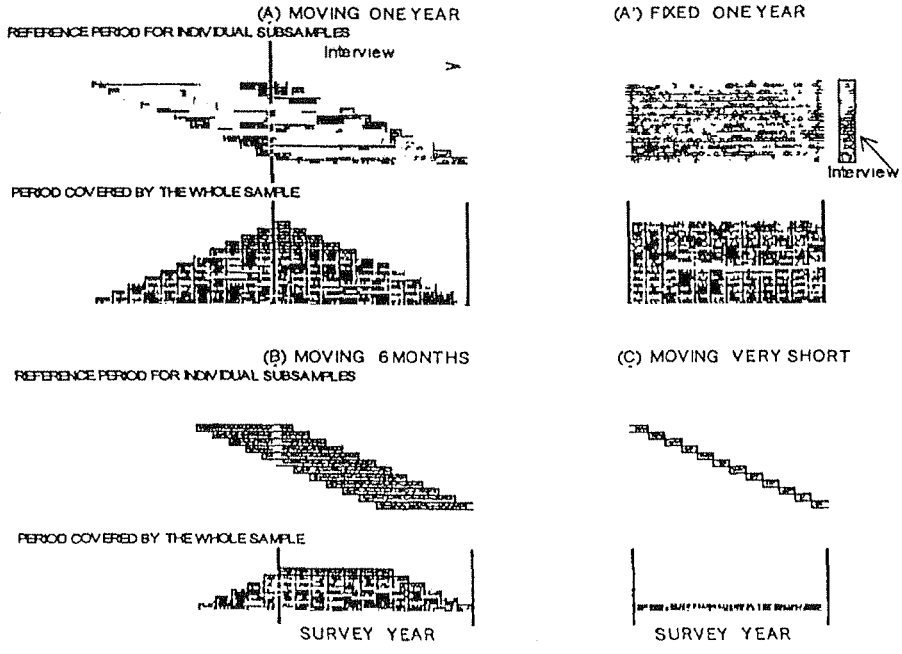
In any household consumption/expenditure survey, prices (whether actual or imputed) for all items of consumption or expenditure need to be adjusted in accordance with the periods they refer to. At the individual level, the relevant price is the one prevailing at the mid-point of the

reference period for the item concerned. Exactly the same procedure as that for a single-year survey applies to any annual or continuous survey involving cumulation over a number of years.

In adjusting prices, it is important to note that *a single, common adjustment factor - reflecting the overall consumer price index for private households - applies to all types of items and all categories of households*. (Using different adjustment factors for different categories will fail to reflect changes in the structure of consumption in terms of values.) This fact considerably simplifies the adjustment process.

On the other hand, if certain items of consumption such as imputed rent are obtained from an external source and refer to a different period than the reference period of the survey, price adjustments to bring them in line with the survey period have to be made using appropriate quantity, price and quality indicators *specific to the item concerned*.

EFFECT OF TYPE AND LENGTH OF REFERENCE PERIOD



The table shows the effect of the type and length of the reference period on the period covered in the whole survey The type refers to whether the reference period is 'moving' or 'fixed' A fixed reference period means a period of the same duration and timing for all units in the sample, irrespective of the timing of their interviewing

With a moving reference period the duration is the same, but the actual time covered differs from one unit or subsample to another depending on the timing of its interview With interviewing distributed over the whole survey year, the period covered for the sample as a whole includes the survey year, and extends beyond it into the preceding year depending on the length of the moving reference period. With the reference period also of one year, the period covered in the survey extends over two years with a triangular density distribution (diagram A) As the length of the moving reference period is reduced, the coverage becomes more concentrated in the actual survey year, and more rectangular in shape (diagrams B and C) With a fixed reference period, the coverage is always confined to the fixed survey year, and is rectangular in shape (diagram A') irrespective of the length of the fixed reference period.

Another effect of the length is on the 'quantity' of the time covered, as indicated by the area of the diagram representing the period covered by the whole sample In a sense, this is the effective sample size This quantity does not depend on the type of the reference period but only on its length: for example, it is the same for diagrams (A) and (A') It is reduced in proportion to the length of the reference period (as in diagrams B and C)

III. RAKING OR ITERATIVE PROPORTIONAL FITTING

A convenient method of adjusting the sample distribution to a number of external controls simultaneously is the classical *Iterative Proportional Fitting* or *Raking* method originally proposed by Deming[4]. The basic idea is to reweight the sample to make the sample distribution agree with the external distribution for each control variable in turn, and then to repeat the whole process till sufficiently close agreement is obtained for all the variables concerned.

The procedure may be applied with any type of units; for convenience, the following description is in terms of households. The control distributions may be in terms of numbers of units, or in terms of aggregated values of variables measured on each unit.

A. Notation

h	Subscript indicating a particular sample household
y_h	Value of a certain variable (y) for household h, e.g. the count of number of households (in which case, $y_h \equiv 1$), household income, household size, number of members with specified characteristics (such as persons of a specified sex and age group).
d_h	Sample weight for household h
u, v, w ...	Classification variables, into categories of each of which the total sample households are divided

i, j, k ...	Subscripts indicating a particular category, respectively of u, v, w ...; thus
	i=1 to I are I categories of u;
	j=1 to J are J categories of v;
	k=1 to K are K categories of w, etc.

h(i)	where necessary, to indicate that household h belongs to category i of variable u; =1 if h belongs, =0 otherwise; similarly h(j) for h belonging to category j of variable v h(k) for h belonging to category k of variable w, etc.
P_i	Proportion of households (or of the sum of household values on a certain variable y) in category i of classification variable u in the *control distribution.* Similarly, P_j, P_k .. in categories j, k ... of classification variables v, w ....

---

4 The original description is in Deming and Stephan (1940), "On a least square adjustment of a sample frequency table when the expected marginal totals are known", Annals of Mathematical Statistics, 11, pp. 427-444.
A more accessible source is Deming, W. E. (1943) Statistical Adjustment of Data, Chapter VII. New York: John Wiley. Paperback edition of this book has been published by Dover Publications New York, first in 1964.

p_i The same proportion estimated (with appropriate weights d_h) from the sample:

$$p_i = \frac{\sum_{h \in i} d_h . y_h}{\sum_{h \in s} d_h . y_h} = \frac{\sum_{h \in i} d_h . y_h}{Y_s}$$

The numerator involves the sum of weighted values for all households *h* which belong to category *i* of classification variable *u*. The denominator involves the sum over all *i*, i.e. over all households in the sample, giving the total Y_s.

B. Outline of the Procedure

The objective is to adjust the sample weights such that

$p_i = P_i$ for all *i*, and

simultaneously to ensure that for categories of all other classification variables

$p_j = P_j$ for all *j* ,

$p_k = P_k$ for all *k*, ...

Starting with the initial weights d_{ho}, the weights are to be adjusted by factors f_h to give revised weights

$$d_h = f_h . d_{ho}$$

such that the re-computed marginal distributions simultaneously satisfy the (I+J+K+...) constraints

$$\begin{aligned} p_i &= \sum_{h \in i} d_h . y_h \Big/ Y_s = P_i \quad \text{with } i = 1 \text{ to I} \\ p_j &= \sum_{h \in j} d_h . y_h \Big/ Y_s = P_j \quad \text{with } j = 1 \text{ to J} \\ p_k &= \sum_{h \in k} d_h . y_h \Big/ Y_s = P_k \quad \text{with } k = 1 \text{ to K, etc.} \end{aligned} \qquad [1]$$

Any adjustment to the weights can always (and are assumed to) be scaled such that this total Y_s for the whole sample remains unchanged, and does not depend on the particular classification (*u, v, w...*) being considered. When y_h represents simply a count variable (identically = 1), then Y_s is the sum of weights of households in the sample, i.e. the weighted sample size. Generally, it is the weighted sum of the variable y of interest overall the whole sample. In all summations, the reference is to the sample households for which the household interview has been completed. The factors by which the weights are adjusted, f_h, in fact do not vary by individual household (h), but depend on the 'cell' to which a household belongs in the cross-classification (IxJxK ...). Hence, more correctly, f_h has been written as an abbreviation for $f_h = f(i,j,k ...)$.

Various models have been proposed for the form of this functional relationship. For instance, it may be assumed that the effects are additive. The more common (and preferred)

model is to assume that the effects are multiplicative. Within each class of models, there is also a choice in the functional form to be used, as well as the numerical procedure for the solution. In addition, restrictions may be placed on the values of weights permitted, such as that the weights must be positive and lie within specified minimum and maximum limits.

C. The "Classical" Iterative Procedure

For the multiplicative type of model, the following describes the 'classical' iterative procedure.

1. Notation

d_{ho} The weight given to the household at all preceding steps, i.e., the starting weights for the iterative procedure.

$d_{(h)}$ In the case of an iterative procedure, the weight for household *h* as determined at the end of the previous iteration.

$p_{(i)}$ The sample proportion similar to p_i, but estimated using weights $d_{(h)}$ as determined at the end of the previous iteration.

$R_{(i)}$ Ratio of the control to the sample marginal distributions for classification variable u (with categories i), $=P_i/p_{(i)}$. as determined using the weights $d_{(h)}$ determined at the end of the previous iteration.

Similarly $R_{(j)}$, $R_{(k)}$ ... for other classification variables, such as

d_h Modified sample weight for household h *after* a particular iteration

At the beginning of the iterative procedure, d_{ho} are the weights for households h as determined at the previous steps of weighting. Let (p_{io}, p_{jo}, p_{ko} ...) be the proportions computed using these weights for classification by variables (u, v, w...) respectively. An iterative cycle consists of determining the weights for each category of each of the classification variables in turn. To begin with, define $d_{(h)}=d_{ho}$ and $p_{(i)}=p_{io}$. With these staring assumptions, the iteration to determine the factors (f) by which the weights are adjusted at each step proceeds as follows.

2. First Cycle

This is applied to each classification variable in turn.

Step 1: Start with categories of the first control variable, say *u*. The weighting adjustment is a function only of the category *i* to which the household belongs. Compute the weighting adjustment f_i by substituting $d_h=f_i.d_{(h)}$ into (1):

$$P_i = \sum_{h\in i} d_h . y_h \Big/ Y_s = f_i . \sum_{h\in i} d_{(h)} . y_h \Big/ Y_s = f_i . p_{(i)}$$

i.e.

$$f_i = \frac{P_i}{p_{(i)}} = R_{(i)}\ ,\ \ h \in i \quad \text{for each i = 1 to I.} \qquad [2]$$

Step 2: Then *using the new weights*, i.e. $d_{(h)}$ redefined as $d_{(h)}.f_i$, determine the marginal distribution according to the next variable, say *v*, as

$$p_{(j)} = \sum_{h \in j} d_{(h)}.y_h \Big/ Y_s$$

As above, the further adjustment to be applied to meet constraint (1) is

$$f_j = R_{(j)}\ ,\ \ h \in j \quad \text{for each j = 1 to J.}$$

Step 3: Next, *using all weights adjusted* by the above steps, determine the marginal distribution according to the next variable (w), and as above, this gives the further adjustment to be applied as

$$f_k = R_{(k)} = \frac{P_k}{p_{(k)}}\ ,\ \ h \in k \quad \text{for each k = 1 to K.}$$

Step 4: Repeat this procedure for each classification variable in turn. This completes the first cycle of the iteration. The total adjustment to be applied to the weight for household h (in cell i,j,k ...) is the product of the adjustments given above for individual classification variables:

$$(f_h)_{\eta=1} = (f_i.f_j.f_k.f_l....)_{\eta=1} \qquad [3]$$

where index η=1 indicates the results of the first iteration.

3. Second and Subsequent Cycles

Step 5: Using all weights $d_{(h)}$ adjusted by factor [3], i.e. $d_{(h)}=d_{ho}.(f_h)_{\eta=1}$, re-compute the marginal distribution $p_{(i)}$ etc., and repeat steps (1)-(4) for the second iteration.

Step 6: Using the results of the first two iterations, steps (1)-(4) can be repeated for further iterations. The process is stopped when all factors in equation (2) become arbitrarily close to 1.0 within a specified numerical limit, say ±0.1%. Normally only a small number of iterations are required.

Step 7: The final weights are the product of the original weights, and the adjustments made at each of the iterations:

$$d_h = f_h.d_{ho} \quad \text{with} f_h = (f_h)_{\eta=1}.(f_h)_{\eta=2}.(f_h)_{\eta=3}..... \qquad [4]$$

4. Order of the Classification Variables (u, v, w, ...)

Sometimes it is more important to ensure more accurate agreement of the final distribution with the control distribution of some classification variables than of other. Since any iteration "disturbs" the results of the previous iterations, it may be considered desirable to order the variables from the least important to the most important, so that at the end, the required controls for the more important variables are satisfied more precisely than for those less important variables in the above sense.

D. Constraining Minimum and Maximum Values of Weights

The 'classical' method described above always gives positive weights, but it places no limits on the minimum and maximum values any weight may take. Uncontrolled variation in weights is not desirable as it can inflate the variance of the survey estimates. In order to incorporate lower and upper limits to be imposed on the resulting weights, we note that the weighting adjustments involved in the multiplicative unconstrained procedure described above may be written in the form:

$$f_h = f(i,j,k,...) = e^{u_i}.e^{v_j}.e^{w_k}....= e^{(u_i+v_j+w_k+...)}.$$

where f_h is the final adjustment to weight of case h as defined in equation [4], and u_i, v_j, w_k ... are the implied adjustment factors determining the final f_h (i,j,k ..) as a function of the categories (i,j,k ..) of the classification variable to which unit h belongs .

The above can be modified to gives the weight adjustments within *specified upper and lower limits* [L, U] with L<1 and U>1 as follows:

$$f_h = f(u_i, v_j, w_k, ...) = \frac{L(U-1)+U(1-L).e_h}{(U-1)+(1-L).e_h}$$

where $$e_h = e^{A(u_i+v_j+w_k...)} \text{ with } A = \frac{(U-L)}{(1-L).(U-1)}.$$

It can be seen that irrespective of the value of e_h, the lower limit for f_h in the above is L (coresponding to e_h=0) and the uper limit is U (coresponding to infinitely large e_h).

In the implementation of the above form, Program CALMAR from INSEE uses an iterative approach, making use of the Newton's method followed by a matrix inversion[5]. The CALMAR procedure requires repeated inversion of rather large matrices, and hence is time consuming. I have found it adequate to use a more straightforward iterative method, requiring a

---

5 *La macro CALMAR*. INSEE: Document no F9310, November 1993.

greatly reduced computational time, as follows. Rather than directly using the computed marginal distribution $R_{(i)}$ to compute f_h as in equation [2] above, an intermediate variable e_h is first computed as below. Otherwise the iterative procedure remains as above. The procedure starts with initial value $e_{(h)}=1$.

An improved value of e_h from value in the previous interation, $e_{(h)}$, is provided by

$$e_h = e_{(h)} \cdot \left(\frac{R_{io} - L}{R_{io}/R_{(i)} - L} \right) \Bigg/ \left(\frac{U - R_{io}}{U - R_{io}/R_{(i)}} \right) \qquad [5]$$

giving an improved value for the weight adjustment factor

$$f_h = \frac{L(U-1) + U(1-L).e_h}{(U-1) + (1-L).e_h}. \qquad [6]$$

The iteration is repeated for categories of each control variable in turn as before.

Notes

1. Here R_{io} is the <u>initial value</u> of the ratio of control to sample proportions P_{io}/p_{i0}, while $R_{(i)}$ is defined as before, i.e. the ratio $P_i/p_{(i)}$ at the end of the previous step in the iteration.
2. As the iterative procedure converges, $R_{(i)}$ tends to 1, so that from [5], e_h tends to some stable value, unchanged from the previous value $e_{(h)}$. The convergence condition is the same a before, namely $R_{(i)}$ arbitrarily close to 1.
3. At the convergence of the iteration, f_h in [6] is the <u>final adjustment factor</u> by which the weight of unit h is adjusted. Unlike equation [4], it is not multiplication of the factors coming out from each iteration.

Limits on the constraints (L and U)

Convergence of the procedure requires that L and U satisfy the constraints for all values of i, j, k ... :

$$L_{\max} < \min(R_{io}),\ \min(R_{jo}),\ \min(R_{ko}),\ \ldots$$

i.e. L must be <u>less than</u> the smallest value in any classification, and

$$U_{\min} > \max(R_{io}),\ \max(R_{jo}),\ \max(R_{ko}),\ \ldots$$

i.e. U must be <u>greater than</u> the largest value in any classification.

The permissible limits within which the weight adjustments can be confined therefore depends on the nature of the data at hand, in particular how far the marginal distributions in the sample according to the classification variables selected differ from those in the external control source. In practice, one may choose the limits of the range of weights to be applied on the basis of what is considered reasonable and acceptable, and then examine the marginal distributions ratios R_{io} etc. to verify that *all* are larger than the desired L and smaller than the desired U. Offending categories, which tend to be small in any case, can be collapsed with other categories for the purpose of weighting controls.

Choosing wide limits, i.e. permitting a large range in the applied weights, can be undesirable as it tends to inflate the increased variance associated with weighting; it may also be objected to on other grounds. On the other hand, confining the weights to a narrow range simply forces a lot of cases to be concentrated near the specified limits, possibly making the weighting adjustments less effective in correcting the distortion present in the sample distributions. Also, choosing tight limits tends to slow down the convergence of the iterative procedure. In practice I have found limits something like (L=0.4, U=2.5) useful, still permitting over 1:6 variation in the range of weights.

IMPLEMENTATION OF SUSI ESTIMATION PROCEDURES

Report # 63

by
Vijay Verma

October, 2002

Statistical Assistance to the Government of Indonesia (STAT) Project
USAID Contract No. PCE-I-00-99-00009-00

TABLE OF CONTENTS

I. INTRODUCTION 374

II. BASIC WEIGHTS 374
A. Sample Design 374
B. Weighted Estimates of Aggregates 375

III. RATIO ESTIMATION 376

IV. MODIFICATION OF WEIGHTS TO DIRECTLY PRODUCE RATIO ESTIMATES 380

V. CONSISTENCY OF RATIO ESTIMATES OF TOTALS 381

VI. MODIFICATION OF WEIGHTS TO DIRECTLY ACHIEVE ADDITIVE RATIO ESTIMATES 385

VII. SUSI ESTIMATION FOR OTHER CHARACTERISTICS FOR ANY DOMAIN D 385

I. INTRODUCTION

The recommended procedure for estimation in SUSI has been described in previous notes and reports. In particular, the basic ratio-estimation procedure is detailed (with illustrations) in *SUSI Estimation Procedure*,[1] and in note "Further to the estimation procedure for SUSI".[2] This report puts together the various ideas elaborated earlier for the estimation of totals, ratios and similar statistics from SUSI, with a view to give systematic guidelines for implementation. Some corrections and clarifications are also offered. Variance estimation, discussed earlier, is not considered in this report.

II. BASIC WEIGHTS

The first step in estimation of course is to compute sample weights taking into account the selection probabilities and response rates in SUSI. These "basic weights" are the inverse of the product of unit selection probabilities and response rates for appropriately defined domains or weighting classes.

A. Sample Design

The SUSI sample has been selected as follows. The sampling frame is provided by the 1996 Economic Census, which itself is based on a sample of areas (wilcah). The census covered all wilcah in urban areas and other areas of concentration of non-agricultural activity, while a 20% sample was taken in other areas. Overall, more than 50% of the wilcah were included. Following the 1996 Economic Census, the first three applications of SUSI (1996, 1998, 1999) were sub-sampled from the census areas as follows.

- For SUSI 1996, a sample of areas (wilcah) was selected with probability proportional to a measure of size (PPS), systematically from lists ordered by geographical location, and all establishments within each selected wilcah were taken (compact cluster sampling).

- In SUSI 1998 a second sampling stage was introduced: within each sample area, establishments were listed and divided into three groups of ISIC (i=1,2,3), and a certain number selected with equal probability from each group. In both the above surveys, fieldwork was confined to a single period during the year.

---

1 Vijay Verma, STAT Report #21, February 2001.

2 Verma, Vijay, *Notes on Estimation Procedures*, STAT Project Report #62, October, 2002.

- The design was similar for SUSI 1999, except that the sample wilcah were divided into 4 replications (representative sub-samples), and one replication was enumerated during each quarter of the year. Thus SUSI became a quarterly survey. However, considering the year as a whole, the sample structure remains unchanged.

B. Weighted Estimates of Aggregates

It is convenient to begin with the simplest form of 'unbiased linear' estimator (which in fact has been hitherto used for the 1996 Economic Census as well as for each round of SUSI). The contribution of a census/sample area k to the estimate of an aggregate y is given by

[1] $\mathbf{Y_{k,i} = y_{k,j} * W_{k,j}}$

where

y an aggregate quantity such as the number of establishments, employment, or output, etc.

$y_{k,j}$ the quantity measured in area (wilcah) k for ISIC(s) j

$W_{k,j}$ weighting or inflation factor for area k for ISIC(s) j, as defined below.

The weighting or inflation factor $W_{k,j}$ is defined as follows.

<u>For the 1996 Economic Census</u>

[2] $W_{k,.} = 1 / (P_{1k} * R_{1k})$

where

P_{1k} the probability of selection of area k into the census (=1.0 for urban and other areas of concentration; =0.2 for other areas)

$R_{1,k}$ response rate in area k (proportion of establishments successfully enumerated) in the census

Note that W is the same for all ISIC's, its value depending only on the wilcah.

Usually, it is not desirable to compute the response rates separately for each individual area k, but to take the average value of groups of areas so as to avoid extreme values. Such groups, or "weighting classes", may for instance be formed by grouping 5-15 adjacent wilcah in the census/sample.

<u>For SUSI 1998</u>

Up to three different sub-sampling rates ($P_{3k,i}$, i=1-3) were used within each sample area for the selection of establishments. (Subscript i represent a group of ISIC's)

[3] $\mathbf{W_{k,i} = 1 / (P_{1k} * P_{2k} * P_{3k,i} * R_{3k,i})}$

n_j number of establishments of *individual* ISIC j in the area according to the 1996 Economic Census .

e_j' total employment in enumerated establishments of individual ISIC j in the area

e_j employment in ISIC j in the area according to the 1996 Economic Census

Beginning with the 'unbiased linear estimator', the contribution of the area to estimates of the number of establishments and employment as measured in SUSI are, respectively, as follows.

[7] $N_j' = W_{j(i)} . n_j'$ from SUSI data

$$E_j' = W_{j(i)} . e_j'$$

where subscript j(i) indicates that the weights are determined by the group i to which ISIC j belongs. It is the same for all ISIC's in the group, and equals W_i defined above.

Similarly, the contribution of the same area to estimates of the number of establishments and employment from the 1996 Economic Census are, respectively:

[8] $N_j = n_j / P$ from the 1996 Economic Census data, *using SUSI areas only*

$$E_j = e_j / P$$

The above quantities may be summed up over all sample areas (k) in a province to provide estimates (of the number of establishments and employment) by ISICxProvince. They may be further aggregated to other domains (d), such as by province to obtain national-level estimates by ISIC, or by ISIC to obtain total (all ISIC) provincial estimates, or both by ISIC and province to obtain total national estimates. For example, from SUSI [7], we have estimates of such domain totals:

[9] $$N_{(d)}' = \sum_{k,j \in d} N_{k,j}' ,$$

$$E_{(d)}' = \sum_{k,j \in d} E_{k,j}' ,$$

where the sum is over all sample areas (k) and/or ISIC's (j) in the domain (d).

Similarly, the 1996 Economic Census gives the following estimates <u>from SUSI areas</u> in the domain:

[10] $$N_{(d)} = \sum_{k,j \in d} N_{k,j} ,$$

$$E_{(d)} = \sum_{k,j \in d} E_{k,j} \,,$$

Note: In either of the above sets, any areas selected but not enumerated in SUSI are excluded from the sum. However, any areas enumerated but found to be empty (no establishments) are to be included with zero value.

We also have the full 1996 Economic Census estimates for the domain (i.e. from *all census areas* in the domain – not just those in the SUSI sample, i.e. the census data weighted by the inverse of the census selection probabilities P_1 defined above). Let these be:

[11] $N^{EC}_{(d)}, \; E^{EC}_{(d)}$

for the number of establishments and employment, respectively.

The corresponding *ratio estimates* from SUSI are:

[12] $$N^{S}_{(d)} = \left(\frac{N'_{(d)}}{N_{(d)}} \right) . N^{EC}_{(d)} \,,$$

$$E^{S}_{(d)} = \left(\frac{E'_{(d)}}{E_{(d)}} \right) . E^{EC}_{(d)} \,,$$

Estimates of proportionate change from the 1996 Economic Census are, obviously:

[13] $$\frac{N^{S}_{(d)}}{N^{EC}_{(d)}} = \left(\frac{N'_{(d)}}{N_{(d)}} \right) ,$$

$$\frac{E^{S}_{(d)}}{E^{EC}_{(d)}} = \left(\frac{E'_{(d)}}{E_{(d)}} \right) ,$$

Note. If a denominator in [12] is zero, i.e. all SUSI areas in the domain were empty *at the time of the 1996 Economic Census*, the linear form [7] is the only one which can be used to provide the required estimate. No estimate of the proportionate change [13] is possible. By contrast, if a domain is empty in SUSI, the estimates are all zero.

Extreme values

Because the present SUSI sample is not fully controlled for distribution by ISIC and province, the direct (linear) estimates for individual ISICxProvince cells can vary unacceptably in some cases. An indication of this is the ratio of SUSI and Census estimates, both for SUSI sample areas (see equation [12]):

$$ratio = \left(\frac{E'_{(d)}}{E_{(d)}} \right).$$

Cells for which this ratio is outside the range, say, 0.5-2.0 should be marked and not published separately.

IV. MODIFICATION OF WEIGHTS TO DIRECTLY PRODUCE RATIO ESTIMATES

In practical applications, ratio estimates of the type described above can be produced directly by using slightly modified versions of the sample weights W defined above. Once the weights are modified at the micro level as below, the production of ratio estimates for the ISICxProvince domains is no more complicated than the production of linear (but less efficient) estimates. This follows from equation [12].

For clarity, we will use the following subscripts.

j	an individual ISIC
i	any grouping of ISIC's (this may refer to individual ISIC's, of course)
j(i)	an individual ISIC j, which belongs to group i
k	an individual SUSI sample area
p	grouping of all areas in province p
k(p)	an individual SUSI sample area, which happens to be in province p
(d)	a domain, such as cell in a ISICxProvince classification (i,p)
$W_{k,j}$	the "basic" sample weight for establishments in area k, ISIC j, taking into account sampling probabilities and response rates, as defined above

First considering employment as the variable, we define the modified weight as

[14] $$W^{S}_{k(p),j(i)} = \left(\frac{E^{EC}_{p,i}}{E_{p,i}} \right) . W_{k(p),j(i)} ,$$

Estimate {7] with the original weights (using the subscripts defined here) for ISIC j province k may be written as:

[7'] $$E'_{k,j} = W_{k,j}.e'_{k,j}.$$

The estimate with the modified weights is

[15] $$E^S_{k,j} = W^S_{k,j}.e'_{k,j} = \left(\frac{E^{EC}_{p,i}}{E_{p,i}}\right).E'_{k,j}.$$

Summing over (k,j) to domain (p,i) gives

[16] $$E^S_{p,i} = \Sigma_{k,j} E^S_{k,j} = \left(\frac{E^{EC}_{p,i}}{E_{p,i}}\right).E'_{p,i}.$$

which is exactly the same as the ratio estimate [12].
The first equation in [16] shows how with modified weights the ratio estimate takes the form of a simple linear estimate.
Ratio estimation of the number of establishments may be simplified in the same way by using a slightly different modification of the weights:

[14'] $$W^S_{k(p),j(i)} = \left(\frac{N^{EC}_{p,i}}{N_{p,i}}\right).W_{k(p),j(i)}.$$

V. CONSISTENCY OF RATIO ESTIMATES OF TOTALS

Consistency of estimates produced from the survey is obviously a desirable property. A special aspect of consistency is the *additivity* of estimates produced at different levels of aggregation. Consistency requires that statistics for mutually exclusive domains should add up to the estimates given for their sum, for instance provincial estimates adding up exactly to the national level estimate provided, or separate estimates by ISIC adding up to the estimate for all sectors combined.

In this sense, ratio estimates, though more precise than simple unbiased estimates (i.e. direct or 'linear' estimates of totals), may not be additive. A slight modification of the weights at the estimation stage may be introduced to ensure such consistency, as explained below. Of course, often such refinement is of no substantive significance, but it is a desirable feature in publishing official results.

The following description of the procedure is taken from my earlier report. In the next

section, it is explained how this may be applied in a simple way by appropriate modification of the sample weights.

Suppose ratio estimates of totals of some variable are reported for a matrix of domains (i,j). For example, i may refer to ISIC and j to province. Thus we are required to produce estimates of totals of some variable Y by ISIC and province (Y_{ij}); also by ISIC for the whole country ($Y_{i.}$); by province for all ISIC's combined ($Y_{.j}$); and the overall total for all ISIC's combined for the whole country ($Y_{..}$).

With linear estimates, it is automatically ensured that the Y_{ij} summed over provinces equal $Y_{i.}$, and summed over ISIC's equal $Y_{.j}$. Similarly, the $Y_{i.}$ summed over ISIC's and the $Y_{.j}$ summed over provinces both equal the overall total $Y_{..}$.

Non-linear estimates such as ratio estimates may not be exactly consistent in the above sense. Consider a ratio estimate for a particular ISIC (i) and province (j)

$$[17] \qquad Y''_{ij} = \frac{y_{ij}}{x_{ij}} . X_{ij} \cdot$$

Its sum over provinces for a given ISIC is

$$[18] \qquad Y''_{i.} = \sum_j \left(\frac{y_{ij}}{x_{ij}} . X_{ij} \right);$$

the sum over ISIC's for a given province is

$$[19] \qquad Y''_{.j} = \sum \left(\frac{y_{ij}}{x_{ij}} . X_{ij} \right);$$

and the total over all ISIC's and provinces is

$$[20] \qquad Y''_{..} = \sum_i \sum_j \left(\frac{y_{ij}}{x_{ij}} . X_{ij} \right) \cdot$$

Normally, in place of [18] and [19], totals by ISIC and province are estimated using 'combined' ratios:

[21] $$Y'_{i.} = \left(\frac{\Sigma_j y_{ij}}{\Sigma_j x_{ij}}\right).X_{i.}$$ and

[22] $$Y'_{.j} = \left(\frac{\Sigma_i y_{ij}}{\Sigma_i x_{ij}}\right).X_{.j},$$

which are not exactly the same as [18] and [19].
Similarly, [21] and [22] summed to the whole country

[23] $$Y'_{(I)} = \Sigma_i Y'_{i.} = \sum\left(\frac{\Sigma_j y_{ij}}{\Sigma_j x_{ij}}.X_{i.}\right)$$

[24] $$Y'_{(J)} = \Sigma_j Y'_{.j} = \sum_j\left(\frac{\Sigma_i y_{ij}}{\Sigma_i x_{ij}}.X_{.j}\right)$$

may differ from each other, from [20], and also from the proper 'combined' overall ratio estimate for the country:

[25] $$Y = \left(\frac{\Sigma_i \Sigma_j y_{ij}}{\Sigma_i \Sigma_j x_{ij}}\right).X_{..}$$

In short, estimates for individual ISIC's may not add up exactly to the 'combined ratio' estimate for all ISIC's combined, or those for individual provinces may not add up exactly to the estimate for the whole country. The discrepancies, even if normally small or even trivial, can be irritating and confusing to the non-statistical user.

How to make these estimates exactly consistent?

1. First make the marginal estimates [21] and [22] to exactly add up to [25]. This is achieved by replacing estimator [21] by

[21.a] $$Y_{i.} = Y'_{i.}.\left(\frac{Y}{\Sigma_i Y'_{i.}}\right) = Y'_{i.}.\left(\frac{Y}{Y'_{(I)}}\right),$$

which summed over i give the correct total Y.

Similarly, replace [22] by

[22a] $$Y_{.j} = Y'_{.j} \cdot \left(\frac{Y}{\Sigma_i Y'_{.j}} \right) = Y'_{.j} \cdot \left(\frac{Y}{Y'_{(J)}} \right),$$

which summed over j also give the correct total Y.

2. Next, for each row in turn, the cell estimates [17] can be re-scaled to add-up to the corrected row total [21.a]:

[17.a] $$Y_{ij}^{(1)} = Y''_{ij} \cdot \left(\frac{Y_{i.}}{\Sigma_j Y''_{ij}} \right) = Y''_{ij} \cdot \left(\frac{Y_{i.}}{Y''_{i.}} \right).$$

Using the results of [17.a], these cell estimates can be re-scaled, for each column in turn, to add-up to the corrected column total [21.b]:

[21.b] $$Y_{ij}^{(2)} = Y_{ij}^{(1)} \cdot \left(\frac{Y_{.j}}{\Sigma_i Y_{ij}^{(1)}} \right).$$

3. The above is the first cycle of an *iterative procedure*. In general for cycle t = 2, 3, … :

[23.a] $$Y_{ij}^{(2t+1)} = Y_{ij}^{(2t)} \cdot \left(\frac{Y_{i.}}{\Sigma_j Y_{ij}^{(2t)}} \right).$$

[23.b] $$Y_{ij}^{(2t+2)} = Y_{ij}^{(2t+1)} \cdot \left(\frac{Y_{.j}}{\Sigma_i Y_{ij}^{(2t+1)}} \right).$$

The iterative procedures may be continued until all terms in parentheses in [23.a, 23.b] become close to 1.0 within a small margin, such as 0.05%. Normally, only a very small number of iterations are needed.

If a single main variable Y and a particular classification of the reporting domains only is of interest, the above adjustments can be incorporated into the sample weights at the micro level, so that strict additivity is automatically ensured when the data are tabulated. However, generally

it is more convenient to apply the above manipulations to the tabulated results.

VI. MODIFICATION OF WEIGHTS TO DIRECTLY ACHIEVE ADDITIVE RATIO ESTIMATES

It can be seen that if the weights of all units in domain (i,j) are multiplied by the factor which is the ratio of the final estimates [27a, b] after the iterative procedure to the original ratio estimates [17], the resulting estimates will automatically have the desired property of additivity across the (i,j) domains, i.e. cells of the classification ISICxProvince.

VII. SUSI ESTIMATION FOR OTHER CHARACTERISTICS FOR ANY DOMAIN D

In the preceding sections, sample weights are defined which automatically yield additive ratio estimates. The additivity property applies strictly only to the type of domains used in the adjustment of the weights: in this case cross-classification by province and ISIC. This restriction is not expected to be a serious one, in the sense that with the new weights, ratio estimates for any set of domains are expected to be additive.

The second point to note is that the weights are slightly different depending on whether we are estimating employment or the number of establishments. For extending the use of these weights to other characteristics, the following procedure is proposed.

All characteristics other than employment and the number of establishments, which are measured in SUSI but not available (or available reliably) from the census, may be divided into two groups: (i) characteristics which are more closely related to employment, and (ii) characteristics more closely related to the number of establishments. The first group is likely to be much larger and more important, and include measures such as output, value-added, turnover etc. The second group may include characteristics which are meaningfully defined only at the establishment level, such as ownership, age, distribution by size, etc. Statistical as well as substantive arguments may be invoked for this classification. There should be no 'residual' category: if in doubt include the characteristics concerned in the 'employment dependent' group.

- For group (i) compute the ratio to employment, using weights as modified in relation to employment, using equation [14], with further modification as described in section VI.

- For group (ii) compute the ratio to number of establishments,

using weights as modified in relation to the number of establishments, using equation [14'], with further modification as described in section VI.