Intertemporal Female Labor Force Behavior in a Developing Country: What Can We Learn from a Limited Panel?[1]

Peter Glick[2] and David Sahn
Cornell University

Abstract

We analyze intertemporal labor market behavior of women in urban Guinea, West Africa using two distinct methodologies applicable to a short (two-year) panel. A multi-period multinomial logit model with random effects provides evidence of unobserved individual heterogeneity as a factor strongly affecting labor market sector choices over time. Results from simpler single period models that condition on prior sector choices are consistent with either heterogeneity or state dependence. Both approaches perform equally well in predicting individual labor market behavior conditional on past choices. In terms of observable characteristics, the estimates confirm the heterogeneous structure of the urban labor market: informal and formal employment appear to differ significantly in terms of skill requirements, compatibility with child care, and costs of entry.

JEL Classification: C33; J13; J22

Keywords: Female employment, fertility, panel data, random effects.

---

[1] The authors thank two anonymous referees for their helpful suggestions and Chad Meyerhoefer and Julie Anderson Schaffner for comments on earlier versions of this paper.

[2] Corresponding author: Peter Glick, Cornell University, 3M02 MVR Hall, Ithaca, NY 14853. Phone 607/254-8782; Fax 607/255-0178; E-mail: pjg4@cornell.edu. The data used in this analysis can be obtained from the corresponding author at this address. This research is supported by the SAGA project, funded by USAID cooperative agreement #HFM-A-00-01-00132-00. For more information, see http://www.saga.cornell.edu.

1. Introduction

As a consequence of the increasing availability of micro-level data, there now exists a large body of research on women's labor force behavior in developing countries[1]. As in many industrialized country studies, this research has explored the effects of factors such as income, education, and the presence of children on female labor force participation and hours of work. However, one important area in which developing country research lags behind that for developed countries is the study of intertemporal aspects of women's labor force behavior, a limitation that reflects the relative paucity of panel data for developing countries. For understanding female work behavior, however, an understanding of labor force dynamics is clearly important. For example, due to childbearing and childcare responsibilities that fall primarily upon women, they are more likely than men to transition between working and non-working states. On the other hand, evidence from industrialized countries suggests that there is nevertheless a high degree of continuity in the work status of individual women. This may be because of heterogeneity among women with respect to time-invariant unobserved preferences for market-oriented activities (or, conversely, home-oriented activities), such that those who chose to work in one period will also do so in later periods. An alternative explanation is state dependence: employment in the current period changes the constraints, incentives, or preferences regarding work, and therefore has a direct positive effect on the probability of work in later periods.[2]

Focusing on industrialized economies, Nakamura and Nakamura (1985,1994) cogently argue for the policy relevance of such year-to-year employment persistence. For example, based on low (relative to men) average cross-sectional participation rates

---

[1]Schultz (1991) summarizes research on male and female employment patterns in LDCs.

[2] A third possibility is serial correlation in transitory unobservable factors influencing the work decision (Maddala 1987).

for women (and specifically women with children), employers may assume that potential female hires will have only a weak attachment to the labor force, making hiring them costly if investments in training or hiring costs are significant. However, this expectation will be wrong if women who do participate tend to do so continuously. This is very much an issue for developing countries as well. Numerous case studies in the developing world indicate that employers' beliefs about women's lack of long-term commitment to the workforce make them reluctant to hire women for formal sector jobs (Anker and Hein 1986). To the extent that women do experience greater interruptions in employment than men, this can negatively affect their rates of pay through reductions in overall work experience and seniority, depreciation of job-related skills, or by forcing them to avoid formal employment entirely in favor of poorly-paid informal work where costs of entry and exit may be lower. Understanding the factors that lead women to withdraw from or enter the labor market therefore may provide insights into women's ability to acquire human capital and to achieve economic parity with men.

To investigate these phenomena empirically requires panel data, that is, repeated observations on the same individuals over time, or else detailed labor force histories collected at a single point in time. In this paper we analyze female work behavior using a two-year panel of household survey data from Conakry, the capital of the West African nation of Guinea. We incorporate a distinctive feature of many developing country labor markets, namely the presence of informal and formal sectors (defined below), which may differ significantly with respect to skill requirements and the costs of entry and exit. Having just two waves a year apart falls well short of what some much longer multiperiod panels in industrialized settings can offer researchers and imposes limits on our ability to model intertemporal factors. Yet, in the

developing country context—and especially in Africa--panel data are rare as is (if less so than in the past) and two waves are in most cases all that is available, given cost considerations or the specific objectives of survey planners.[3] Therefore it is of significant practical interest to determine what kinds of techniques are possible and how much insight into labor force behavior can be gained using a short panel.

We apply and compare two distinct methodologies that are applicable to our data. The first is a multi-period reduced form multinomial logit model of labor market sector choice that accounts for unobserved time-invariant individual heterogeneity. One of the advantages of this model, also known as mixed logit, is that unlike the standard multinomial logit (MNL) model it allows for non-zero correlations of the error terms for different alternatives. We also use the estimates to generate predictions of behavior conditional on past labor market choices, which are of potential significance to policymakers and employers in that they incorporate the factors leading to continuity in individual labor market behavior.

The second method, associated with Nakamura and Nakamura among others, directly estimates period *t* employment sector outcomes conditional on prior (*t*-1) choices, which implicitly controls for heterogeneity as well as capturing state dependence. This approach is much less demanding computationally as well as with respect to data requirements, since it makes use only of indicators of lagged employment state rather than the full set of lagged regressors. Subject to certain caveats, the parameter estimates indicate the impacts of various factors on women's transitions between labor market states. Like the first model, this method yields

[3] In Glewwe and Jacoby's (2000) review of panel data collection in developing countries, most of the cases consist of just two data points per household or individual.

employment predictions conditional on past behavior and indeed is advocated largely for this reason.

The rest of this paper is organized as follows. The next section describes in detail the two estimation approaches we use. Section 3 discusses the data and descriptively analyzes labor force transitions among women in the sample. Section 4 presents the model results and compares the predictive accuracy of the different approaches. Section 5 concludes with a summary and a discussion of methodological as well as policy implications of the results.

2. Model Specifications

2.1. Multinomial logit model with unobserved heterogeneity

The first model we estimate is a reduced form multi-period multinomial logit with random effects. Utility for individual *i* from sector *j* (non-participation, self-employment, wage employment, indexed j=0,1,2)[4] in time *t* (*t*=1,2) is expressed as

$$V_{ijt} = X_{it}'\beta_j + \alpha_{ij} + \varepsilon_{ijt} \tag{1}$$

X_{it} is a vector of explanatory variables including individual characteristics such as income and education as well as year dummies. The ε_{ijt} are time-varying i.i.d error terms while α_{ij} is an individual and sector-specific, time-invariant random effect. The individual chooses the sector for which utility is highest. If the ε_{ijt} follow the Type I extreme value distribution, the probability of choosing sector *j* at time *t* conditional on X_{it} and the random effects takes the multinomial logit form:

$$P(j,t|X_{it},\alpha_{i1},\alpha_{i2}) = \frac{exp(X_{it}'\beta_j + \alpha_{ij})}{\sum_{s=0}^{2} exp(X_{it}'\beta_s + \alpha_{is})} \tag{2}$$

[4] This division of the labor market is discussed in detail in the next section.

For identification, α_{i0} and β_0 are normalized to zero; that is, we make non-participation the base choice. The α_{ij} are assumed to follow a multivariate normal distribution. This is implemented by specifying the vector $\alpha_i = [\alpha_{i1},\alpha_{i2}]'$ to be linear combinations of *J-1* independent standard normal variables: $\alpha_i = A\eta,\ \eta_i \sim N_2(0,I_2)$.[5] A is a 2 x 2 lower triangular matrix, estimated along with β. The matrix of covariances for the heterogeneity terms α_i is then AA'. If the random effects were observed, the contribution to the likelihood of individual *i* with observed sector outcomes y_{i1}, y_{i2} would be the sequence of multinomial logit probabilities

$$L_i(\eta_i) = P(y_{i1}|X_{i1},\alpha_i)P(y_{i2}|X_{i2},\alpha_i) \tag{3}$$

Since the α_i are not observed, to get the unconditional likelihoods the conditional likelihoods must be integrated over all possible values of η_i (hence of α_i).

$$\iint L_i(\eta_i)f(\eta_i)d\eta_{i1}, d\eta_{i2} \tag{4}$$

This involves *J-1* dimensions of integration, which in the present case is 2. We approximate the integral through simulation. For each individual, *R* values of the η_i are drawn from the distribution $N_{j-1}(0, I_{j-1})$ and the likelihood conditional on each set of values is calculated. We replace the integral by the average of the *R* conditional likelihoods:

$$\frac{1}{R}\sum_{r=1}^{R} L_i(\eta_i^r) \tag{5}$$

[5]Although the normality assumption for the heterogeneity terms in such models is standard, an anonymous referee has pointed out that a distribution that allows for bimodality would be more appropriate to capture heterogeneity if women tend to fall into two distinct groups ('high' and 'low') with respect to preferences for work in a sector. While beyond the scope of this paper, this would make an interesting exploration for future work.

Because it incorporates the multinomial logit formula for sector choice probabilities, (6) is a smooth function of the parameters, making the application of simulated maximum likelihood relatively straightforward. The simulator is consistent but its accuracy will be a function of the number of replications (Brownstone and Train 1999). The model was estimated in GAUSS, using 250 draws in the estimation. Because the choice probabilities combine the logit form with a different distributional assumption for the heterogeneity terms (normality in most cases), models of this type are often referred to as "mixed" or "heterogeneous" logit models. Note that the mixed logit nests the standard logit as a special case (α_i=0), so it is possible to compare the two statistically using a likelihood ratio test.

An attractive feature of the mixed logit is that, unlike a standard multinomial logit (whether estimated on a single or pooled cross sections), the error terms of the utility functions for different choices are not assumed to be independent. In the standard model independence leads to the restrictive Independence of Irrelevant Alternatives (IIA) property. Denoting the composite error term for alternative *j* in the random effects model as (suppressing the individual *i* subscript) $\mu_{jt} = \alpha_j + \varepsilon_{jt}$, we have $\text{cov}(\mu_{jt}, \mu_{kt}) = E([\alpha_j + \varepsilon_{jt}][\alpha_k + \varepsilon_{kt}]) = \sigma_{\alpha j,\alpha k}$. Hence the unobserved portions of utility for alternatives *j* and *k* are related through the correlation of their heterogeneity terms, so the IIA property does not hold. By providing greater flexibility in the pattern of error correlations among choices, the random effects specification offers advantages over the standard MNL beyond the usual gains in efficiency associated with modeling an error components structure.

It would be desirable to allow the α_j to be related to, say, the number of young children or household income and test for this correlation. However, identification in such correlated random effects models usually relies on the use of leads and lags in

X_{it}, which is only practical if there is sufficient variation over time in the regressors (see Hyslop 1999 for applications using a binary probit model). Here we are at a disadvantage due to the shortness of our panel, as there is relatively little variation in the data over the two years.[6] Therefore we assume, as in most applications, independence of X_{it} and α_{ij}.

To generate predicted sector probabilities and marginal effects (derivatives of probabilities with respect to the independent variables) for the random effects multinomial logit model, we evaluate for each individual in the sample the sector probabilities and derivatives conditional on a given draw of the η_i and take the average over all (300) draws. In the empirical literatures on marketing and transportation mode choice, in which mixed logit models have been most prominently applied, interest has centered on using the model to forecast individual consumers' demands for new products based on their prior demand behavior (see Revelt and Train 1999; Brownstone and Train 1999). In the present context, we would like to know how well the model forecasts the labor market behavior of particular women or certain groups of women—e.g., those in formal sector wage work—based on their observed prior behavior. The expected probability of employment in sector *j* in period *T+1*, conditional on the sequence of past choices, can be simulated as:

$$P(j,T+1 \mid y_1..y_T) = \frac{\sum_r P(j,T+1 \mid \beta,\eta^r)P(y_1 \mid \beta,\eta^r)...P(y_T \mid \beta,\eta^r)}{\sum_r P(y_1 \mid \beta,\eta^r)...P(y_T \mid \beta,\eta^r)} \qquad (6)$$

The denominator is the simulated likelihood for the sequence of past choices (through *T*) while the numerator is the likelihood of the sequence of choices including for period *T+1* if *j* were chosen for *T+1* (see Revelt and Train 1999). With just two

[6] For example, the inter-year correlation of the number of children under five is 0.8.

periods, this amounts to calculating (6) to generate second period predicted outcomes conditional on first period outcomes (i.e. T=1, $T+1$=2). We compare these calculations to actual second year choices to assess the predictive accuracy of the model.

2.2 *Estimating sector choice conditioning on prior choice*

Consider adding to the model of equation (2) a vector of dummy variables Z_{it} for lagged labor market sector:

$$P(j,t|X_{it},\alpha_{i1},\alpha_{i2}) = \frac{exp(X_{it}'\beta_j + Z_{it}'\gamma_j + \alpha_{ij})}{\sum_{s=0}^{2} exp(X_{it}'\beta_s + Z_{it}'\gamma_s + \alpha_{is})} \tag{7}$$

This yields a dynamic multinomial logit model that incorporates state dependence through the coefficients on the lagged state dummies as well as capturing impacts of unobserved heterogeneity on period t employment state. Estimation of this general model is not feasible with our two period panel for two reasons. First, there is the 'initial conditions' problem that arises because we are unable to model the initial state (sector in period 1), given the lack of information from prior periods. This renders the structural coefficients on the lagged dependent variables inconsistent through the association of these variables with unobserved preferences for work or sector. The most common solution, following Heckman (1981), would be to use reduced forms for the initial period sector utility functions, i.e., excluding the lagged sector choice variables. This could be done in a two wave panel, but identification of the state dependence effect still hinges on the presence of truly time-varying regressors X_{it} (Chamberlain 1984), or else pre-sample information that can plausibly be excluded from the second period equations. These conditions by and large are not met in our data.

The second problem involves the more complex covariance structure of the dynamic model. With Heckman's suggested method, the heterogeneity terms in the initial period reduced form would not be the same as (though they would be correlated with) those in the subsequent structural equations. With just a two period panel, however, this would essentially leave only one wave of the panel to estimate the heterogeneity covariances for the structural model, whereas identification of these parameters requires repeated observations on individuals over time. Hence the fully specified dynamic model is not feasible in our short panel.[7]

Instead, we can simply treat the lagged sector variables as fixed in the second period and estimate (7) as a straight multinomial logit for period 2 choices (dropping the heterogeneity terms). This leads us directly to the conditional approach advocated by Nakamura and Nakamura, following the suggestion of Heckman (1978). As just noted, lagged employment status is likely to be endogenous to current choices, so the coefficients will pick up the effects of unobservables that are correlated with lagged as well as current outcomes. In this sense the model is not an appropriately specified dynamic model for identifying true state dependence, i.e., the direct impacts of work in one period on work in the next. However, while we are not in general able to obtain unbiased coefficient estimates of such causal effects, the endogenous lagged labor supply information is potentially very useful for generating more accurate *predictions* of labor market behavior for specific groups of women (Nakamura and Nakamura 1994). By capturing the effects of both state dependence and time-invariant preferences that are associated with observed work outcomes, the lagged work dummies should be powerful predictors of current or future work states. [8]

[7] Gong, van Soest, Villgomez (2000) were able to estimate such a model using panels from Mexico City consisting of five quarterly observations.

[8] This is not to argue that predictive accuracy rather than unbiased causal estimates

A more flexible parameterization would interact all or some of the covariates in (7) with the lagged employment sector dummies, thus allowing behavioral responses to depend on the prior state. With complete interaction, we end up in effect directly estimating year-to-year transition probabilities between all pairs of labor market states. These estimates would be of value to those who are interested, for example, in understanding the determinants of women's shifts out of non-employment into self or wage employment, or of labor force exits of women who are employed in a given sector. We estimate current (second) period choices using both forms of conditional models in this paper. Using the estimates we will calculate conditional current (year 2) choice probabilities and compare the accuracy to the conditional predictions from the logit model with unobserved heterogeneity derived using equation (7).

We just noted that the coefficients on the uninstrumented lagged work variables are expected to pick up the effects of unobservable tastes. However, prior work status remains an incomplete proxy for such tastes and this should inform our interpretation of the estimates on the other covariates in the models. Variables such as the number of children may be correlated with preferences for sector of work even after conditioning on prior sector status.[9] Since any such remaining association will lead to biases in the coefficient estimates, caution is needed when attempting to make causal inferences from these conditional models.

---

should be the main goal of econometric analysis. Rather, it is that in some cases (such as obtaining estimates of employment persistence) prediction is also of direct policy or analytical interest.

[9] The point is brought out in the formal presentation in Nakamura and Nakamura (1994, p. 323).

3. Data and Context

This study uses two years of data from the Conakry Household Welfare Survey, collected in 1990 and 1991. Guinea shares a number of key features with the economies of other African countries. The country switched from a rigidly controlled state dominated economy to a significantly liberalized one after 1984. Economic growth in the years following reform has been variable but on balance weak. The formal sector of the Conakry labor market remains small, and for women especially, is dominated by public employment. As elsewhere in urban Africa, self-employment is very significant in Conakry, accounting for more than a third of men's employment and about three quarters of women's employment.[10]

The first year of the survey involved a random sample of 1,725 households. In the second year an attempt was made to re-interview all of the first year households, with individuals matched across years by id number in the household rosters. It was possible to re-interview about 80% of all working age (15-65) women appearing in the initial survey. 2,469 women aged 15-65 were used in the analysis. The year-to-year rate of individual attrition from the sample of approximately 20% is in line with reported experiences with panel data collection in other developing countries (see the summary provided by Alderman et. al. 2000, Table 1). Attriters differ statistically from stayers in our data in some respects, though the differences are typically not large: attriters are younger (27 vs. 31 years old), slightly better schooled (4 years vs. 3.7 years), and slightly less likely to be married (60% vs. 67%), as well as being in households with higher non-labor income. They are also less likely to have been self-employed (but not wage-employed) in the first year (21% vs. 29% for stayers).

[10] See Glick and Sahn (1997) for more information on the labor market in Guinea.

It is important to recognize that neither these differences or the overall extent of attrition mean that parameter estimates of labor force behavior based on the sample of stayers will be biased. Some assessment of the potential for such bias is possible using methods suggested by Fitzgerald et. al. (1998) and Becketti et. al. (1988). First, we estimated a probit for attrition on the set of covariates from the first period as well as the year 1 labor market sector dummies.[11] In this multivariate context we still see a significant positive association of attrition with age, income, and a negative association with initial period self-employment. Next we estimated a sector choice model on the first year data including interactions of each regressor with a dummy variable for attriting, to test whether relevant behavior in the first period differs for the two groups. For the determinants of first year self-employment, we cannot reject joint equality of the complete vector of coefficients including the intercept for the two groups (p=.13) and for slopes alone non-rejection is unambiguous (p=.36). The coefficient vectors for leavers and stayers do not differ for wage employment either (p=.66 and .59 with and without constant terms, respectively). Considering individual covariates, significant interactions with the attrition dummy are found only among self-employment determinants, for two variables: non-labor income and (at 10% level) spouse unemployment.

These results suggest that by and large initial year labor force behavior is not different for the two groups. We therefore can feel fairly confident that despite some differences in measured characteristics of leavers and stayers, our parameter estimates

[11] To conserve space we only summarize the results here. The full set of estimates is available from the authors.

using the panel sample will not be unduly biased from selective attrition. Other studies that have applied these methods typically come to similar conclusions. [12]

The survey collected information on various types of income generating activities, including wage employment and all types self-employment activity, whether in family or individual enterprises. We also count as self-employed women who engaged in activities such as sewing or food preparation in the home that was partly for income as well as for household consumption, as long as the estimated time related to income generation (determined by multiplying total hours in the activity by the ratio of estimated value of sales to total value produced) was 10 hours or more. This somewhat arbitrary rule, used to avoid counting as 'in the labor force' women who may have worked no more than an hour or two for revenue, made little difference to the estimates. Our work status indicators are based on reported activity over the past year to avoid counting as non-participants women whose work was seasonal or who were otherwise temporarily inactive; however, use of prior week information yielded essentially the same results.

Means and standard deviations of the explanatory variables for the entire sample and conditional on first year employment state are given in Table 1. Our division into self and wage employment as representing informal and formal labor market sectors is in part borne of necessity: we lack information on employment conditions, e.g., the availability of health or vacation benefits or enterprise size, that could be used to distinguish formal from informal employment. However, the data we have suggest that our division is sensible. As the table shows, female wage employees have much higher schooling than the self-employed (11 years vs. 2 years). In addition, most (two-thirds)

[12]See Alderman et. al. for examples using developing country data and the special issue of *The Journal of Human Resources* (1998 v.33 no.2) for experiences with US data.

female wage employees work in the public sector, and whether in the public or private sector they are found overwhelmingly in clerical or white-collar occupations such as secretarial work, teaching, and nursing (Glick and Sahn 1997). Self-employed women, on the other hand, typically work in very small (usually one-person) retail enterprises using small amounts of capital. By any definition, these enterprises would qualify as belonging to the informal sector.

Table 2 shows the transitions of women among labor market states between 1990 and 1991. Note first that the share of women working in Conakry is quite low by urban African standards--about 34% in 1990[13]--possibly reflecting the weakness of the Guinean economy as well as the legacy of restrictive government practices toward private enterprise in the previous regime. There is evidence of significant movement of women in and out of the labor force overall. Fully 25% of women who worked in the first period in either sector reported no work in the second, while entrants in the second year were equivalent to 26% of the first year working sample.[14] For comparison we calculated the equivalent figures for men and found them to be much lower, about 5% in each case. Altogether about 18% of the total sample of women experienced a transition between any two sectors in 1990-1991, given by the total of the off-diagonal elements of the transition matrix. Since only 12 individuals switched between wage and self-employment, this figure is essentially the same as the share of women changing their overall participation (work/no work) status. These transition rates are high compared with women in developed economies. [15]

---

[13] The 'non-employed' group includes genuine non-participants as well as the unemployed. However, only about 5% of women who were not working were searching for work, i.e., would be defined as unemployed.

[14] The first figure is derived as follows: 209 and 16 women exiting self and wage employment, respectively, in 1991 (first column) over 684 plus 180 1990 self- and wage employed (last column).

[15] For example, for the U.S., Shaw (1994) reports that, depending on the age range

Table 2 also makes clear the need to consider transitions by sector. Year-to-year persistence in participation is much stronger in wage employment than self-employment, where turnover is quite high. 31% of women reporting self-employment activity in 1990 reported no such activity in 1991. Virtually all of these had exited the labor force altogether (only 5 women switched from self- to wage employment). A similar number of women made a transition from non-employment to self-employment. In contrast, only 13 % of the 1990 wage earner sample was not also working in the following year.

Thus is it the presence of the large self-employment sector that accounts for the high rates of female labor force turnover in this sample compared with rates in developed countries. A higher turnover in self- as compared to wage employment is consistent with differences by sector in either state dependence or preferences. There may be substantial costs associated with entry and exit from formal sector employment arising from significant (worker financed) investments in schooling and on the job training, and from skills depreciation or loss of seniority resulting from interrupted employment. These costs are probably much lower for women's self-employment, which is usually very small scale (implying low start-up costs) as well as less skill-intensive. Alternatively, greater period-to-period persistence in formal wage work may occur because wage employed women are particularly strongly oriented toward work and career. Both of these interpretations are consistent with (though they do not prove) the presence of competitive labor markets. In particular, note that differential costs of exit and entry will occur in a heterogeneous yet competitively functioning labor market characterized by differences in technology--

---

used, 12 to14% of married women in her PSID sample changed their labor force status during a recent 2-year interval. Looking similarly only at married women in our sample, we find that 23% entered or exited employment between 1990 and 1991.

one sector having high fixed costs of training and low turnover and the other having little training and high turnover.

Alternatively, however, the sectoral differences in worker flows may reflect the presence of non-market barriers to entry into one or another sectors (in this case, formal wage employment), that is, labor market segmentation. Rates of entry and exit in the wage sector may be low because the supply of jobs is limited, making search costs high. This would limit entry while also making temporary withdrawals from the labor force costly to those already in the wage sector.[16] We will return to this issue after we have considered the multivariate results.

4. Estimation Results

4.1. Multinomial logit model with unobserved heterogeneity

Table 3 presents the estimates for two panel multinomial logit models with random effects: one including the complete series of demographic covariates and the other a reduced specification including only age, schooling, income, and year dummies. We ran the latter specification in view of the fact that demographic variables are potentially endogenous to labor force outcomes; if this is the case, the second model will correspond to the correct reduced form model.[17] We focus first on

[16] Maloney (1999), finding symmetry in the numbers of workers moving between informal and formal employment in his urban Mexican panel, argues against the standard picture of a dualistic labor market in which workers use informal employment as a staging ground for one-way transitions to the restricted formal sector. In our data we also see equal numbers of individuals moving in each direction between sectors, but the absolute number of such transitions is so small (12) that making the same inference as Maloney does is not warranted. The small number itself is compatible with either highly heterogeneous but competitive markets or with segmentation.

[17] We estimated the models excluding demographic covariates on the suggestion of an anonymous referee. A qualification in needed to the statement in the text: this will be a true reduced form only if there exist no additional (unobserved) exogenous determinants of fertility and household structure that are correlated with the included regressors (see

the full specification. The results provide strong evidence of unobserved heterogeneity among women with regard to labor market choices. The variances of the heterogeneity terms for both self-employment and (especially) wage employment are large and highly significant. The addition of these terms leads to a very large increase in the log likelihood value over that for a standard multi-period multinomial logit without random effects, and a likelihood ratio test easily rejects the standard model. Using the fact that the idiosyncratic errors (the ε_{jt}) are each constrained as extreme values to have variances of $\pi^2/6$, we can calculate the share of the total variation in the unobserved portion of utility for each work alternative that is due to time-invariant individual heterogeneity rather than idiosyncratic errors. For self-employment this share is 0.83 while for wage employment it is close to unity (0.97).[18] This result is in line with previous applications of mixed logit models, which by and large have found that unmeasured heterogeneity accounts for the majority of the variation due to unobservables. The particularly high ratio for wage employment is consistent with the fact that we observe little movement in or out of this sector over time. We could conclude, plausibly, that women who choose wage occupations are particularly highly career motivated relative to women on average. Note, though, that in addition to true time-invariant heterogeneity, the heterogeneity terms will also pick up state dependence effects as well any autocorrelation in the errors due to serially dependent unmeasured shocks.

Turning to the parameter estimates for the non-stochastic portion of utility, the determinants of participation are quite different for self-employment and wage

---

Browning 1992).

[18]These ratios can be interpreted equivalently as the intertemporal correlation coefficients of the sector utility disturbances, i.e., $\text{corr}(\mu_{jt}, \mu_{jt-1}) = E([\alpha_j + \varepsilon_{jt}][\alpha_j + \varepsilon_{jt-1}]) = \sigma^2_{\alpha j}/(\sigma^2_{\alpha j} + \sigma^2_{\varepsilon jt})$. Intuitively, the correlation over time must come solely from the time-invariant component of the errors.

employment, confirming the heterogeneous structure of the urban labor market in Guinea. Years of education has a negative effect on utility from self-employment but a strongly positive effect on wage employment.[19] Among self-employment determinants, we observe the expected negative impact of income, represented in the model by household non-labor income. Reflecting the large and complex structure of extended Guinean households, the explanatory variables include fairly detailed breakdowns of household demographics. For self-employment, the effects of own children are non-linear in the number of children though generally positive. For one child under 5 the coefficient is positive though only significant at the 15% level. For two children the effect is larger and highly significant. Only for three or more children does the estimate become negative in sign (the very small share of women with 3 or more young children should be kept in mind). Also for self-employment, there is a positive effect of children under 5 of other women.

These positive impacts of young children on (self-employment) participation are the opposite of what is usually found in industrialized county studies of female labor force participation. However, in this sample of mostly very poor women the need to generate additional income to meet children's needs is likely to be particularly pressing and may overwhelm the need for childcare in the home, at least until the number of young children exceeds two. Further, many forms of self-employment activity are likely to be compatible with child supervision and care. In this regard it is noteworthy that the positive effect of children (one's own or others') is completely absent for wage employment, where compatibility is likely to be much lower.

[19] Due to normalization, the parameters are to be interpreted as showing the effect of the covariate on utility from working in the given sector relative to the base choice (non-employment).

Older daughters—both age 5 to 14 and 15 to 20—also have significantly positive impacts on self-employment. The fact that we do not observe similarly significant impacts of sons suggests that daughters are seen as potential substitutes for working mothers in household activities. The number of adult men (age 21 and older) has negative effects on wage employment, possibly reflecting an income effect since a greater number of male adults means there are more potential income earners. Being married is strongly associated with working in self-employment, despite the fact that access to the income or assets of the husband should raise a woman's reservation wage. However, this assumes that spouses pool their incomes, whereas non-pooling seems more typical in the West African context (Fapohunda 1988; Hoddinott and Haddad 1995). An alternative and possibly conflicting interpretation is that married women are more likely to enter self-employment because their spouse's capital, expertise or connections makes it easier for them to set up small enterprises.

Finally, we include an indicator of whether the women's spouse was reported as unemployed at the time of the survey. This variable, which can be thought of as capturing a negative resource shock to the household, has a large and strongly significant positive effect on participation in self-employment. Comparative statics calculations show that current period spouse unemployment raises a woman's self-employment probability by about 12 percentage points, which is equivalent to almost a 50% proportional increase. Since we would expect short-term labor supply responses to negative income shocks to occur where barriers to entry are lower, the fact that they are seen only in self-employment and not also wage employment is an additional sign that there are significant costs in terms of training or search to entering the formal sector.

It can be seen from the table that the model excluding demographic covariates yields parameter results that are qualitatively and even (especially for wage employment) quantitatively very similar to the full specification. This applies as well to the estimates of the heterogeneity variances and covariance. We also estimated all subsequent models discussed in this paper using both specifications. The choice of specification had essentially no qualitative effect on any results reported in this paper—the signs and significance of the remaining covariates, predictive capabilities (which remained largely unchanged in absolute terms for all models) and comparisons of model predictive accuracy. In view of this finding, we will henceforth present only the results for the models including the demographic variables.

Despite the unambiguous rejection of the standard MNL model in favor of the random effects specification, the economic implications with respect to the non-stochastic determinants of behavior are quite similar. Calculated marginal effects (derivatives of sector probabilities with respect to the X_{it}, calculated from the estimates and the data) proved to be generally very close for the two models (results are available from the authors). This can be interpreted with reference to the discussion in section 2, where we noted that a significant difference in the two specifications is that the mixed logit permits correlations in errors across choices. Since for our data we cannot reject the null that the covariance of the errors (equal to $\text{cov}(\alpha_1,\alpha_2)$ in Table 3) is zero,[20] we would not expect substantive differences in the implications of the two models. The same holds for predictive capability. Table 4 examines this in two ways. The first calculates for each year and alternative *j* the mean predicted probability of *j* for the subsample actually choosing it in that year. The second way (shown in square brackets) calculates the percent of successful

[20]The correlation of the errors of self- and wage employment is just 0.07.

predictions by year and actual alternative chosen, with success defined to occur when the alternative chosen by the individual in the period is also the one with the highest predicted probability. Note that these are unconditional predictions, that is, not conditional on previous employment states; the latter will be examined below. There is very little difference in predictive accuracy for the two models. Both yield the correct sector assignment in exactly 68% of the total cases in each year. As is typical when the choice outcomes are unbalanced in the data, there is a tendency to underpredict participation in the lower frequency states of self- and wage employment.

4.2 Estimation conditioning on prior choice

Next we present estimates from multinomial logit models of sector choice for 1991 that incorporate information on 1990 sector status. The first model (shown in Table 5) adds 0-1 indicators for previous year self- and wage employment to the same set of regressors used above (non-employment is the excluded category). Lagged self-employment has a very large and significant impact on current period self-employment and the same is true for lagged wage employment and current wage employment. This model therefore provides evidence, in a rather different form from the mixed logit model, of strong period-to-period persistence in employment status. As discussed above, the lagged dummies in our case will capture both true state dependence and individual heterogeneity. Not very many of the other covariates have statistically significant effects once we condition on past choice. However, this is to be expected. It occurs in part because for covariates that change little over time the bulk of their effects come though the sorting into first period labor market states. In addition, conditioning purges, at least partially, the effects of the associations of the

covariates with tastes for work or sector. Still, even controlling for prior year status, both being married and having an unemployed spouse increase the likelihood of self-employment.[21] Also noteworthy is the continued strong positive effect of schooling on wage employment.

Table 6 presents the results for the more flexible model allowing interaction of prior sector status with all covariates. As noted above, with this approach we are in effect directly estimating transition probabilities between different pairs of labor market states. Because of the very small number of individuals moving from self-employment into wage employment or the reverse, we were not able to estimate the determinants of these transitions, so these observations were dropped. Thus the model is limited to estimating the determinants of second period self-employment for first period non-participants and self-employed (parameter estimates shown in columns 1 and 2), and the determinants of second period wage-employment for first period non-participants and wage-employed (columns 3 and 4).[22] Consistent with the results of the previous models, having an unemployed spouse in the current period causes women who were initially not working to enter self-employment, as does being married (column 1). Income effects are also observed: an increase in household non-labor income reduces the likelihood that a non-participant will enter self-employment while increasing the likelihood that a self-employed woman will stop working (column 2). Note that for the group of year 1 self- employed the model is estimating the determinants of remaining self-employed, which is to say we are modeling exits from self-employment into non-employment with signs reversed.

---

[21] Although the numbers with an unemployed spouse are small, this variable does show significant variation over time: of the 158 women with an unemployed spouse in either year, 95 had spouse unemployed in one year but not the other

[22] An implication of this is that the interacted model does not nest the lagged sector dummy model of Table 5.

Despite the relatively small numbers of women moving into or out of the wage sector between years, we find several significant determinants of these transitions. Schooling and age have positive effects on transitions from non-employment to wage work (column 3). Among the first year wage employed, there are negative effects on second year wage employment of having one child under 5 and two children under 5, though only the latter coefficient is statistically significant (column 4). That is, women with young children are more likely than those with no young children to leave their formal wage jobs. In view of the discussion in Section 2, a strictly causal interpretation of this result is difficult to make. Even conditioning on prior selection into wage work, fertility and preferences for work may be (inversely) related. Nevertheless, these estimates are consistent with expectations regarding the difficulty of combining formal work with childcare, as well as the likely inflexibility of formal work in the sense of offering fewer part time opportunities to working mothers. Note also that while spouse's unemployment induces non-participating women to enter self-employment, it is not associated with entry of non-participants into the wage sector, consistent with higher costs or barriers to entry in the short term in formal employment.[23]

As noted earlier, an important advantage of the conditional approach is said to lie in its use of information on past behavior to obtain accurate predictions of future employment status. We assess this in Table 7, which calculates predicted 1991 sector probabilities using the estimates from the model including lagged sector dummies. The table also shows the conditional predictions from the random effects multinomial

[23] This claim may seem odd given that the spouse unemployment variable does not appear in the equation for non-participation to wage work transitions. However, the variable is dropped because of collinearity arising from the fact that *no* non-participating women with current spouse unemployment enter the wage sector. The logit estimation breaks down because spouse unemployment is a perfect (negative)

logit model of the previous section, calculated using equation (6). Before comparing these two models, recall that in Table 4 we used the random effects estimates to generate *un*conditional predictions of employment sector. Comparing those predictions for 1991 with the conditional ones for the same model in Table 7, it is clear that conditioning on information about past choices yields a substantial improvement in predictive accuracy for the last (second period) choice. The strongest gain is for observations choosing wage work in 1991: the mean predicted probability of wage work for this group rises from 0.37 (next to last column, 3rd row of Table 4) to 0.74 (last column, first row of Table 7).

Comparing the conditional 1991 predictions for the random effects logit and the simpler lagged sector model, we observe from Table 7 that the two yield very similar results. For the percent correctly predicted criterion, the results are almost identical. What makes this similarity in conditional predictive accuracy especially noteworthy is that the multi-period random effects model makes much greater use of the panel data, since all the first year data are used in the estimation rather than simply the first year choice outcomes. Apparently, as Heckman originally conjectured, lagged outcome variables contain a lot of information, capturing both state dependence and individual heterogeneity.

Finally, to compare the implications of the conditional and unconditional approaches, Figures 1a and 1b present simulated second period wage and self-employment probabilities by age and number of young children using estimates from the following two models for 1991 employment sector: a standard multinomial logit model that does not condition on prior choice, and the model of Table 7, which conditions on prior choice by interacting all regressors with previous year sector

---

predictor of wage work among this group.

indicators. The top panel shows the wage sector probabilities. When calculating these probabilities, covariates other than the age and children variables are set equal to the means for the 1990 wage worker sample. Comparison of the curves from the two models vividly demonstrates how expectations of current or future behavior change when we condition on past behavior. 1991 wage employment probabilities conditional on having been wage employed in 1990 are much higher than we would predict for the same group of women without information on their prior status. Having more young children reduces the likelihood a wage-earning woman will remain in her job, but 'commitment' to work remains very high for all wage workers: the second period conditional probabilities are very high in absolute terms and always much larger than the unconditional probabilities, no matter what child status is assumed for the latter. [24]

For self-employment too, predicted probabilities conditional on prior participation in the sector lie well above the unconditional sector probabilities (Figure 1b; all the self-employment probabilities are evaluated at the sample means for the first year self-employed).[25] However, in absolute terms they are not as high as the wage work probabilities conditional on prior wage work, and the proportional difference between the conditional and the unconditional probabilities is generally

[24] With regard to the much flatter participation–age profile for the conditional estimates, note that this does not show true life-cycle patterns for this group. Since only the prior year work information is used, conditioning necessarily drops some older women who had left wage employment before the previous year as well as younger women who have not yet entered, which serves to flatten the wage work-age profile relative to what it would be for all those who will be wage employed during their working lives.

[25] Although the unconditional self-employment probabilities suggest significantly less concavity in age than the analogous predictions for wage employment, the comparison is misleading. Different groups of women are used in the two cases, and the change in probability with respect to age depends on the mean subsample values of the covariates, which (especially education and income) are very different for wage and self-employed women.

smaller for self-employment, reinforcing the idea that selection based on attachment to work operates particularly strongly in wage employment. A more general lesson to be taken away from these figures is that inferences about the future work behavior of specific groups of women will be misleading if they are based on analysis (descriptive or econometric) of all women. Of course, in other cases—in fact in most cases—we are interested in making inferences about the population as a whole (conditioning on exogenous factors such as education and age) rather than about specific groups defined on endogenous prior outcomes. For this purpose the unconditional reduced form models retain their value.

5. Summary and Discussion

This paper has explored two distinct methodologies for analyzing labor market choices that are applicable to a short—two year—panel. These methods attempt in different ways to incorporate factors linking an individual's labor market behavior across periods. A multi-period random effects multinomial logit model of sector choices provides strong evidence of time-invariant heterogeneity in preferences (or possibly, ability) as an explanation for the period-to-period persistence in women's labor market outcomes. The alternative and simpler approach estimates a single period choice model including lagged sector indicators or more flexibly directly estimates transitions between sectors. Both approaches confirm that women who work (or who work in specific sectors) differ in terms of preferences or constraints, or both, from other women. Therefore, expectations about the future labor market behavior of such women will be highly misleading if based only on cross sectional data for all women. As Nakamura and Nakamura (1994) have noted, the use by employers of such incorrect inferences may result in a form of statistical

discrimination since it underestimates the commitment to work of those women who do choose to participate.

Both approaches can use information on past choices to make more accurate predictions of future labor market choices of individuals or of specific groups, e.g., formally employed women, than is possible using just cross-section data. The more complex random effects logit model did not perform better on these conditional predictions than a simpler one-period sector choice model including lagged labor market status. Therefore strictly from the point of view of obtaining such predictions, the simpler approach can be recommended. On the other hand, from the perspective of obtaining unbiased reduced form *un*conditional (population) parameters, the multinomial logit model incorporating unobserved heterogeneity is in principle to be preferred to the standard multinomial logit, primarily because it is able to model correlations in errors across choices.

The results with respect to the impacts of observable factors confirm the heterogeneous structure of the labor market in this urban African setting. The determinants of participation in the informal (self-employment) and formal (wage employment) sectors differ in important ways that are apparently related to skill requirements, aspects of work such as childcare compatibility, and costs of entry and exit. Formal work is more difficult for women with young children: the presence of children under 5 is positively associated with self-employment but not wage employment, and makes women who are already wage employed more likely to stop working. High costs of entry/exit for formal employment are suggested by low worker flows in and out of the sector and the econometric estimates showing an association of a negative income shock (proxied by spouse unemployment) with entry into self-employment but not wage employment.

The econometric results as well as the descriptive data thus point to constraints on women's access to formal sector work. But even with our panel data we are not able to say whether our results are evidence of differences in job characteristics (and individual preferences) in a diverse but competitive labor market or instead reflect institutional factors that act to limit entry to formal sector employment. This is unfortunate since the appropriate policy stance may depend on the answer. However, other sources of information can be brought to bear on the question and these suggest that institutional factors play some role. Hiring in the public sector had slowed considerably by the time of the surveys, and in fact many workers were being retrenched as part of the economic reform program (Mills and Sahn 1995). Therefore some rationing of entry into formal employment was likely to have been operative at the time of the surveys.[26] Glick and Sahn (1997) note that a different form of rationing—gender bias in hiring—is also likely to be occurring, especially for low skill wage jobs, most of which are in the private sector. In sharp contrast to men, extremely few women with no or little schooling are found in wage employment.

For well-educated women, who would have a chance at formal sector employment even given these institutional constraints, the results suggest that entry into (or permanent employment in) this sector is inhibited by childcare responsibilities. Therefore while year-to-year continuity of employment among wage-earning women is already high, it might be even higher if such women had access to affordable childcare services, which is something that public policy can provide. Further, while raising employment persistence among women currently involved in wage employment, such policies would also work on the extensive

[26] Inferential evidence for this is provided by the very low share of men age 21-30 reporting being in paid work (27%), usually a good indicator of a lack of formal sector job opportunities (Glick and Sahn 1997).

margin: they would draw more women into the sector, subject, of course, to the supply of such jobs. Currently only women with particularly strong career aspirations are likely to find that the benefits of wage employment exceed the cost in terms of foregone time in child care. Publicly subsidized childcare would make it possible for women with less inherent work 'commitment' to participate in formal employment with high levels of period-to-period continuity. Since formal sector work is usually associated with greater pay and possibilities for advancement than most self-employment activities, such policies may serve to improve women's economic status.

Although the results of this study point to the potential benefits for labor market analysis of just a two-year panel, we have also emphasized that such a short panel imposes limits on the analysis of intertemporal behavior. More elaborate dynamic models that could distinguish the alternative sources of serial persistence in women's employment status require more observations per individual. Such long panels have provided important insights into dynamic labor force behavior in industrialized settings and would be equally beneficial in developing countries.

References

Alderman, H., J. Behrman, H.-P. Kohler, J. Maluccio and S. Watkins, 2000, Attrition in longitudinal household survey data: Some tests for developing country samples, Policy Research Working Paper Series #2447, World Bank, Washington D.C.

Anker, R. and C. Hein, 1985, Why do third world employers usually prefer men?, International Labor Review 124, 73-90.

Becketti, S., W. Gould, L Lillard, and F. Welch, 1988, The Panel Study of Income Dynamics after Fourteen Years: An Evaluation, Journal of Labor Economics 6, 472-92.

Browning, M., 1992, Children and Economic Behavior, Journal of Economic Literature 30, 1434-1475.

Brownstone, D. and K. Train, 1999, Forecasting new product penetration with flexible substitution patterns, Journal of Econometrics 89, 109-129.

Chamberlain, G., 1984, Panel data, in: Z. Griliches and M.D. Intrilligator, eds., Handbook of Econometrics, Vol. 2 (Elsevier Scientific Publishers, Amsterdam).

Fapohunda, E., 1988, The nonpooling household: A challenge to theory, in: D. Dwyer and J. Bruce, eds., A Home Divided: Women and Income in the Third World (Stanford University Press, Stanford).

Fitzgerald, J., P. Gottschalk, and R. Moffitt, 1998, An Analysis of Sample Attrition in Panel Data, The Journal of Human Resources 33(2), 251-99.

Glewwe, P. and H. Jacoby, 2000, Recommendations for collecting panel data, in: Grosh and Glewwe, eds., Designing Household Surveys for Developing Countries: Lessons from 15 Years of World Bank Experience (World Bank, Washington, D.C.).

Glick, P. and D. Sahn, 1997, Gender and education impacts on employment and earnings in a developing country: the case of Guinea, Economic Development and Cultural Change 45, 793-823.

Gong, X., A. van Soest, and E. Villagomez, 2000, Mobility in the urban labor market: a panel data analysis for Mexico, IZA discussion paper No. 213, Institute for the Study of Labor, Bonn.

Heckman, J., 1981, Heterogeneity and state dependence, in: S. Rosen, ed., Studies in Labor Markets (University of Chicago Press, Chicago) 91-139.

Hoddinott, J. and L. Haddad,1995, Does female income share influence household expenditure patterns: Evidence from Côte d'Ivoire, Oxford Bulletin of Economics and Statistics 57, 77-96.

Hyslop, D., 1999, State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women, Econometrica 67, 1255-1294.

Maddala, G.S., 1987, Limited dependent variable models using panel data, The Journal of Human Resources 22, 307-338.

Maloney, W. 1999. Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico, The World Bank Economic Review 13(2), 275-302.

Mills, B. and D. Sahn, 1995, Reducing the size of the public sector workforce: Institutional constraints and human consequences in Guinea, The Journal of Development Studies 31, 505-528.

Nakamura, A. and M. Nakamura, 1985, Dynamic models of the labor force behavior of married women which can be estimated using limited amounts of past information, Journal of Econometrics 27, 273-298.

Nakamura, A. and M. Nakamura, 1994, Predicting female labor supply: effects of children and recent work experience, The Journal of Human Resources 29, 304-327.

Revelt, D. and K. Train, 1999, Customer-Specific Taste Parameters and Mixed Logit, University of California, Berkeley.

Shaw, K, 1994, The growing persistence of female labor supply and wages, The Journal of Human Resources 29, 348-378.

Schultz, P., 1991, International differences in labor force participation in families and firms, Economic Growth Center, Yale University.

Table 1

Women 15-65: Variable means and standard deviations by initial year sector

	First year employment sector							
	Non-employed		Self-employed		Wage employed		All Women	
Variable	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Age	28.34	11.98	35.17	10.69	35.24	7.59	30.73	11.82
Years schooling	3.72	4.58	1.79	3.57	10.71	5.27	3.69	4.87
Non-labor income[a]								
(Guinean Francs)/10000	17.54	48.31	11.74	25.35	20.44	73.75	16.15	45.79
One child<5	0.23	0.42	0.31	0.46	0.32	0.47	0.26	0.44
Two children<5	0.11	0.31	0.18	0.39	0.16	0.36	0.13	0.34
3 plus children<5	0.01	0.10	0.01	0.11	0.02	0.15	0.01	0.11
# other kids<5	1.08	1.33	0.98	1.30	0.56	0.85	1.01	1.30
# sons 5-14	0.27	0.63	0.55	0.79	0.54	0.84	0.36	0.71
# daughters 5-14	0.21	0.53	0.49	0.73	0.46	0.71	0.31	0.62
# sons 15-20	0.10	0.38	0.23	0.53	0.19	0.50	0.15	0.44
# daughters 15-20	0.08	0.33	0.20	0.49	0.23	0.51	0.13	0.40
# other kids 5-14	2.04	1.98	1.55	1.69	1.36	1.62	1.85	1.90
# other males 15-20	0.70	1.03	0.50	0.80	0.59	1.16	0.64	0.99
# other females 15-20	0.91	1.02	0.57	0.83	0.46	0.70	0.78	0.97
# men 21 plus	2.12	1.60	2.09	1.59	1.69	1.15	2.08	1.57
# women 21 plus	2.43	1.66	2.26	1.52	2.26	1.43	2.37	1.61
Married	0.57	0.50	0.88	0.32	0.78	0.42	0.67	0.47
Spouse unemployed	0.03	0.17	0.08	0.27	0.03	0.18	0.04	0.21
No. of observations	1605		684		180		2469	

Notes:

"Other" refers to children in the household other than the woman's own children

[a]Income received by the household from pensions, social security, insurance, interest earnings, and remittances

Table 2
Women 15-65: Transitions among Employment States, 1990-1991

		1991 sector			
1990 sector		Non-employed	Self-employed	Wage-employed	All
Non-employed	number	1389	192	24	1605
	% of sector[a]	86.5	12.0	1.5	100.0
	% of all women	56.3	7.8	1.0	65.0
Self-employed	number	209	470	5	684
	% of sector[a]	30.6	68.7	0.7	100.0
	% of all women	8.5	19.0	0.2	27.7
Wage-employed	number	16	7	157	180
	% of sector[a]	8.9	3.9	87.2	100.0
	% of all women	0.6	0.3	6.4	7.3
All	number	1614	669	186	2469
	% of all women	65.4	27.1	7.5	100.0

Notes:
[a] number in the indicated transition category divided by number employed in the originating sector in 1990

Table 3
Multi-period multinomial logit models of sector choice with random effects

	Excluding demographic covariates		Including demographic covariates	
Variable	Self-employment	Wage employment	Self-employment	Wage employment
Intercept	-12.9182	-55.3090	-10.7584	-53.3276
	(14.22) ***	(7.75) ***	(10.90) ***	(7.04) ***
Age	0.6515	1.8479	0.4536	1.8524
	(12.77) ***	(6.87) ***	(7.81) ***	(6.02) ***
$Age^2/100$	-0.7862	-2.0146	-0.5319	-2.0229
	(11.73) ***	(6.34) ***	(7.04) ***	(5.47) ***
Yrs. Schooling	-0.1625	1.1629	-0.1173	1.1291
	(6.88) ***	(7.27) ***	(4.74) ***	(7.14) ***
Non-labor income[a]	-0.0080	0.0016	-0.0084	0.0033
	(3.53) ***	(0.47)	(3.51) ***	(0.95)
One child<5	—	—	0.2803	0.1049
			(1.54)	(0.15)
Two children<5	—	—	0.7440	-0.0909
			(3.15) ***	(0.10)
3 plus children<5	—	—	-0.3251	-0.6436
			(0.44)	(0.13)
other kids<5	—	—	0.2137	-0.5907
			(2.98) ***	(1.32)
sons 5-14	—	—	-0.0481	0.2912
			(0.41)	(0.65)
daughters 5-14	—	—	0.2406	0.0019
			(2.09) **	(0.00)
sons 15-20	—	—	0.2402	-0.0076
			(1.47)	(0.01)
daughters 15-20	—	—	0.3644	-0.1536
			(2.03) **	(0.21)
other kids 5-14	—	—	-0.0028	0.2382
			(0.06)	(1.16)
other males 15-20	—	—	-0.1258	-0.1318
			(1.34)	(0.32)
other females 15-20	—	—	-0.1378	0.1819
			(1.29)	(0.49)
men 21 plus	—	—	0.0368	-0.7443
			(0.63)	(2.44) **
women 21 plus	—	—	-0.1211	-0.0408
			(1.79) *	(0.15)
Married	—	—	1.1293	-0.3418
			(4.36) ***	(0.37)
Spouse unemployed	—	—	1.5054	-0.6040
			(5.12) ***	(0.31)
Year 2	-0.2075	-0.2867	-0.1773	-0.2746
	(2.08) **	(0.47)	(1.73) *	(0.82)
Heterogeneity covariances:				
$var(a_1)$	8.13 (4.34) ***		7.97 (5.60) ***	
$var(a_2)$	67.60 (3.66) ***		62.73 (3.52) ***	
$cov(a_1,a_2)$	3.03 (0.359)		1.81 (0.21)	
Log-likelihood	-2789.8		-2725.6	
No. of observations	2469		2469	

Notes: Aysmptotic t-statistics in parentheses.
Standard errors of heterogeneity terms calculated using delta method
[a] Divided by 10,000 for the estimation.
***significant at 1%; **significant at 5%; *significant at 10%

Table 4

Predictive accuracy of multi-period sector choice models

	Standard logit				Logit with random effects[a]			
Year	Non-empl.	Self-empl.	Wage empl.	All	Non-empl.	Self-empl.	Wage empl.	All
1990	0.72	0.39	0.39	—	0.72	0.39	0.36	—
	[0.86]	[0.29]	[0.46]	[0.68]	[0.88]	[0.29]	[0.40]	[0.68]
1991	0.72	0.39	0.40	—	0.72	0.38	0.37	—
	[0.88]	[0.27]	[0.48]	[0.68]	[0.89]	[0.25]	[0.41]	[0.68]

Notes:

Each cell shows mean predicted probabilty of sector j in time t for the subsample of women actually choosing j in time t. Figures in brackets give the percent of successful predictions for each subsample, with success defined to occur if the alternative actually chosen in time t is also the one with the highest predicted probability.

[a]From model in Table 3, columns 3 and 4. Predictions for random effects model are based on 300 replications. See text for details.

Table 5

Multinomial logit model of 1991 sector choice including lagged sector status

	Self-employment			Wage employment		
Variable	Estimate	T-statistic		Estimate	T-statistic	
Intercept	-5.420	-8.34	***	-11.993	-5.70	***
Age	0.179	4.60	***	0.379	3.16	***
$Age^2/100$	-0.219	-4.32	***	-0.418	-2.73	***
Yrs. Schooling	-0.028	-1.77	*	0.191	6.02	***
Non-labor income[a]	-0.005	-2.24	**	0.003	1.01	
One child<5	0.042	0.28		-0.082	-0.19	
Two children<5	0.247	1.27		-0.947	-1.63	
3 plus children<5	-0.587	-1.07		—	—	
Other kids<5	0.163	2.85	***	-0.332	-1.65	*
Sons 5-14	-0.009	-0.10		0.389	1.88	*
Daughters 5-14	0.078	0.87		0.059	0.21	
Sons 15-20	0.130	1.09		-0.414	-1.13	
Daughters 15-20	0.108	0.80		-0.197	-0.53	
Other kids 5-14	0.003	0.07		0.071	0.70	
Other males 15-20	-0.003	-0.05		0.052	0.26	
Other females 15-20	-0.192	-2.24	**	-0.200	-0.84	
Men 21 plus	-0.049	-1.15		-0.074	-0.65	
Women 21 plus	0.011	0.22		0.015	0.12	
Married	0.488	2.62	***	-0.517	-1.12	
Spouse unemployed	0.562	2.24	**	-0.350	-0.41	
1990 sector (excluded=non-employed)						
Self-employment	2.424	20.21	***	0.227	0.43	
Wage employment	0.851	1.78	*	5.522	14.10	***

Notes: Base choice is non-employment. Number of observations =2469

[a] Divided by 10,000 for the estimation.

***significant at 1%; **significant at 5%; *significant at 10%

Table 6
Multinomial logit model of 1991 sector conditioning on 1990 sector status

	Self-employment		Wage employment	
	1990 sector status		1990 sector status	
Variable	Non-employment	Self-employment	Non-employment	Wage employment
Intercept	-7.500	0.037	-14.357	0.623
	(7.64) ***	(0.04)	(4.68) ***	(0.09)
Age	0.323	0.033	0.460	0.108
	(5.37) ***	(0.55)	(2.53) **	(0.24)
$Age^2/100$	-0.418	-0.026	-0.521	-0.006
	(5.14) ***	(0.35)	(2.12) **	(0.01)
Yrs. Schooling	-0.012	-0.018	0.224	0.041
	(0.56)	(0.74)	(4.86) ***	(0.49)
Non-labor income[a]	-0.006	-0.008	0.003	0.012
	(1.81) *	(2.07) **	(1.04)	(0.71)
One child<5	0.189	-0.280	1.018	-1.414
	(0.91)	(1.25)	(1.79) *	(1.51)
Two children<5	0.176	0.190	—	-2.020
	(0.65)	(0.65)		(1.98) **
3 plus children<5	—	-0.095	—	—
		(0.13)		
Other kids<5	0.275	0.080	-0.727	-0.104
	(3.26) ***	(0.89)	(2.07) **	(0.22)
Sons 5-14	-0.180	0.171	-0.069	1.481
	(1.49)	(1.40)	(0.20)	(1.96) **
Daughters 5-14	-0.149	0.346	0.006	-0.102
	(1.14)	(2.54) **	(0.02)	(0.22)
Sons 15-20	0.272	0.044	-0.094	-1.417
	(1.68) *	(0.25)	(0.14)	(2.09) **
Daughters 15-20	0.235	-0.026	-1.215	0.395
	(1.26)	(0.13)	(1.16)	(0.49)
Other kids 5-14	0.001	0.035	0.094	0.044
	(0.02)	(0.58)	(0.62)	(0.21)
Other males 15-20	-0.112	0.124	0.187	0.486
	(0.96)	(0.99)	(0.75)	(0.97)
Other females 15-20	-0.186	-0.122	0.005	-0.445
	(1.42)	(0.95)	(0.02)	(0.93)
Men 21 plus	-0.095	-0.051	-0.212	0.142
	(1.37)	(0.79)	(1.29)	(0.45)
Women 21 plus	-0.133	0.143	0.171	-0.307
	(1.67) *	(1.67) *	(1.06)	(0.88)
Married	0.612	0.259	-0.128	-0.292
	(2.19) **	(0.89)	(0.19)	(0.29)
Spouse unemployed	1.258	-0.087	—	-1.100
	(3.71) ***	(0.27)		(0.82)

Notes: Aysmptotic t-statistics in parentheses. Base choice is 1991 non-employment.
Number of observations = 2457
[a] Divided by 10,000 for the estimation.

'—' indicates that the covariate was excluded due to an absence of variation within the transition subsample.

Table 7

Predictions of second period sector choices conditional on prior choice

Model	Non-employment	Self-employment	Wage employment	All
Multiperiod logit with random effects[a]	0.80	0.54	0.74	—
	[0.87]	[0.67]	[0.84]	[0.82]
1991 sector choice logit including lagged sector status[b]	0.80	0.55	0.77	—
	[0.88]	[0.66]	[0.84]	[0.82]

Notes:

[a]Based on model estimates in table 3, cols. 3,4 and using 300 replications for predictions. See text for details.

[b]Based on model estimates in table 5

Each cell shows mean predicted conditional (on t-1 choice) probabilty of sector j in time t for the subsample of women actually choosing j in time t. Figures in brackets give the percent of successful predictions for each subsample, with success defined to occur if the alternative actually chosen in time t is also the one with the highest predicted probability.

Figure 1a - Simulated Conditional and Unconditional 1991 Wage Employment Probabilities

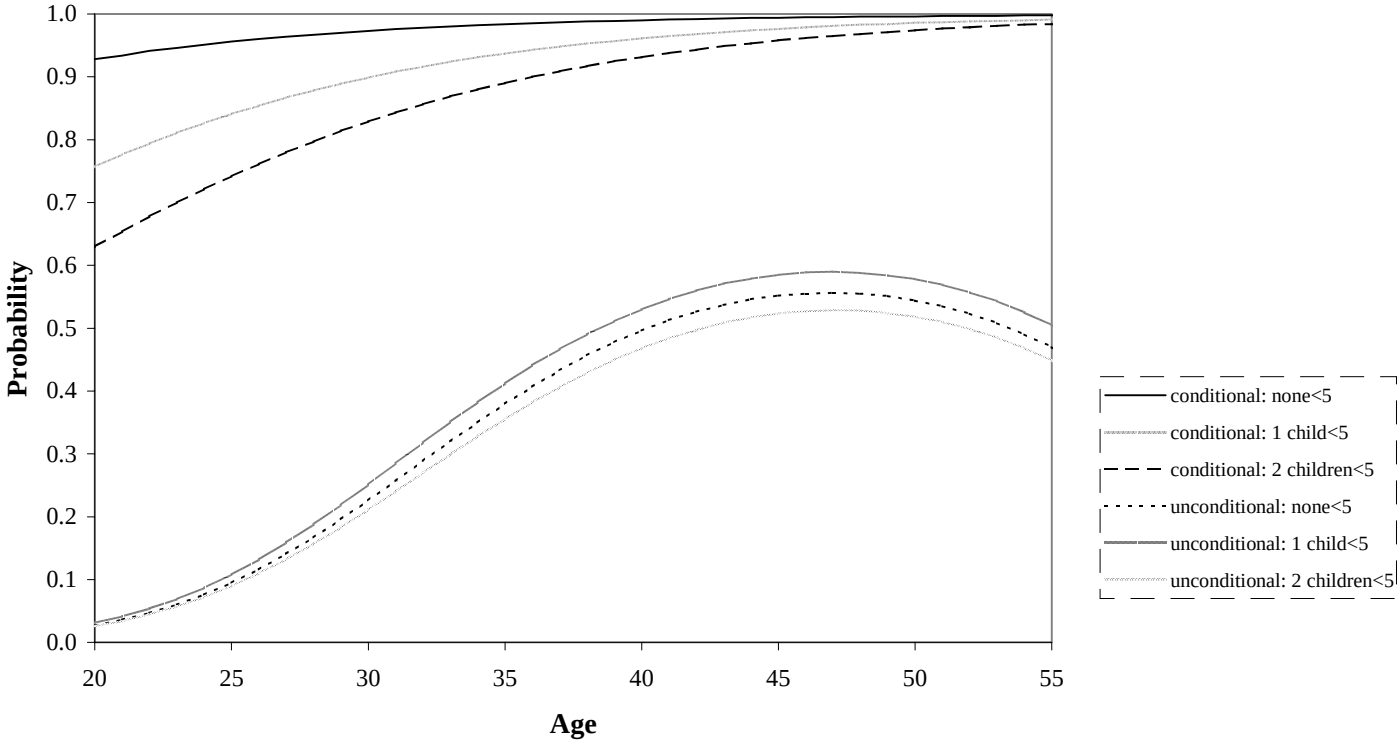
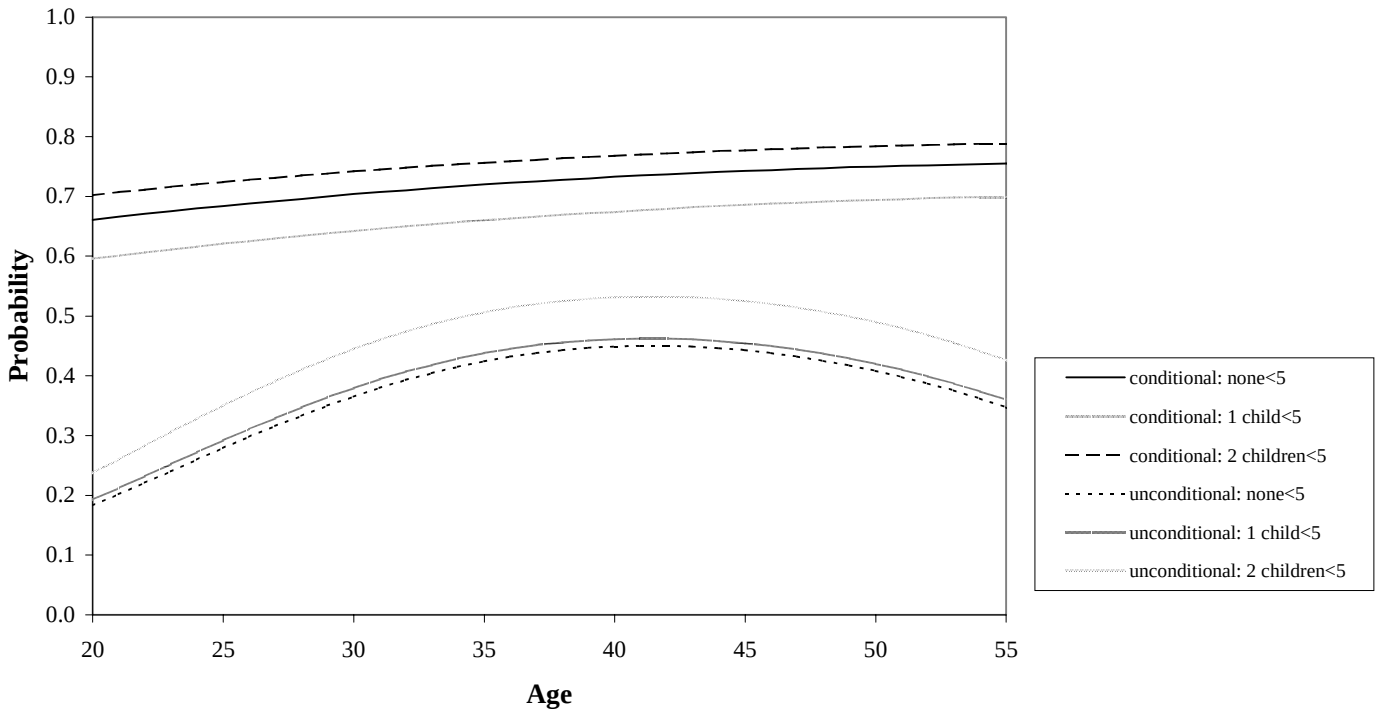


Figure 1b - Simulated Conditional and Unconditional 1991 Self-employment Probabilities



Notes:
Unconditional probabilities are predicted from a standard multinomial logit model of 1991 sector choice. Conditional probabilities are predicted from a multinomial logit model of 1991 sector choice conditioning on 1990 sector and show wage (self-) employment probabilities conditional on previous wage (self-) employment (see text for details). For wage (self-) employment predictions, covariates other than age and child status are set equal to subsample means of wage (self-) employed women.