The SPSS Project Model of Population, Labor Force, and Pensions in Armenia

Armenia Social Protection Systems Strengthening Project

Landis MacKellar

SO 3.4

USAID Armenia 111-C-00-06-00067-00

AECOM International Development
December 5, 2007

USAID
FROM THE AMERICAN PEOPLE

USAID ARMENIA, SOCIAL PROTECTION SYSTEMS STRENGTHENING PROJECT
²ØÜ Ø¼¶ Ð³Û³ëï³Ý, êáóÇ³É³Ï³Ý å³ßïå³ÝáõÃÛ³Ý Ñ³Ù³Ï³ñ·»ñÇ Ñ½áñ³óÙ³Ý Íñ³·Çñ

The SPSS Project Model of Population, Labor Force, and Pensions in Armenia

Prepared by SPSS Labor Modeller
Landis MacKellar

This paper was made possible with the support of the American People through the U.S. Agency for International Development. The contents of this paper are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the US Government.

Yerevan, December 2007

Introduction and basic concepts

The SPPS Armenia pension model consists of a set of linked Excel workbooks (linked in the sense that changing an assumption in one workbook may lead to a change in numbers calculated in another). What we call a model is, in fact, a series of *modules* consisting of individual workbooks.

The model has two main purposes:

- To project the likely evolution of the current pension system given the most likely evolution of fertility, mortality, migration, labor force participation and unemployment, the number of persons complying with mandatory social contribution rules, indexation of pensions, and all other relevant variables. By convention, a "most likely" scenario of this type is often called a *baseline scenario*.
- To make it possible to do "if – then" experiments in which one (or more) parameters or assumptions is changed. For example, if compliance were higher than in a baseline scenario, by how much would the balance of the SIF pension system move towards surplus? The counterfactual is commonly referred to as an *alternative scenario*. Alternative scenarios can amount to of simple *sensitivity tests*, in which only one assumption is changed (from the baseline), to broader depictions of a very different state of the world, in which many assumptions are different than they are in the baseline.

An extended version of the SPSS model takes the baseline-alternative scenario process one step further by making it possible to compare two different pension system structures – one in which the current zero- and first-pillar systems remain unchanged, and one in which they are gradually supplemented or replaced by a fully-funded second pillar.

Whether used for analyzing future trends, quantifying the impact of alternative sets of assumptions, or studying different structural approaches to pension provision, the purpose of the model is the same: to serve as a transparent tool generating quantitative estimates that can serve as a basis for policy discussions. The transparency aspect is the most important, as any model is credible only to the extent that its workings are accessible to and understood by its users.

A few general rules

First, copy all files into a directory named SPSSModBackup or something similar.

When working with a model as flexible and user-friendly as this one, it is a good idea to have one directory corresponding to each scenario – say, "Baseline," "Low Fertility," and "Improved Compliance" – that contains all model workbooks. The reason for this is to make sure that changes in assumptions made for one scenario do not "contaminate" another. Every time you want to do a different scenario, just make a copy of the directory Baseline, save it under a different name, and make assumption changes in the new directory.

The fact that the SPSS model contains links between different worksheets requires the user to adhere to a certain amount of discipline:

- When files are saved in a new directory, links are automatically changed, i.e. the model "knows" that it is no longer looking for files in (say) the directory "Baseline" but in the directory "Low Fertility."

- When a workbook is loaded, you will be asked whether it should be updated using external links to other workbooks. Unless there is a good reason to do otherwise, just say, "Yes." When you close a workbook, you will be asked if you want to save the changes you have made. Again, unless there is a good reason to do otherwise, say, "Yes."
- If you make the slightest structural change to a workbook – say deleting or adding a line -- files that are not open at the time will have no way of "knowing" that you have made this change. When you load another file and update, it will go to the old, pre-change cell references in the workbook you have altered, typically with serious consequences.
- The same goes for changes in the names of workbook or worksheets (the "sections" of workbooks). Any file not open at the time you make the change will have no way of "knowing" what has happened.
- The SPSS model has many dimensions – age, sex, economic sector, formal / informal sub-sector, and rural / urban residence. As a result, workbooks are large; they cannot all be loaded simultaneously. It is not feasible to simply load all workbooks at the same time and rely on Excel's automatic update features to make required model-wide changes when any single module is changed is not feasible.
- Workbook cells are colour-coded:
    - Green cells contain data which you, the user, must enter. Sometimes these represent historical data and there is no reason why you would want to change them. Other times, they represent assumptions for future years that you must make in order for the model to solve.
    - Turquoise blue cells contain parameters or assumptions imported from another worksheet or workbook. There is no reason for you to touch them. Often, these data are imported from the worksheet "Notes," which is the first worksheet in every workbook. "Notes" contains assumptions as well as the calculation of basic parameters that are required by other worksheets in the workbook. It often also contains some information on the sources of data.
    - Yellow cells are automatically calculated by the spreadsheet. There is no reason for you to touch them. Note that sometimes a line of cells is yellow for the historical period but green or turquoise for the projection period. Similarly, cells may be green for the historical period but yellow or turquoise for the projection period.

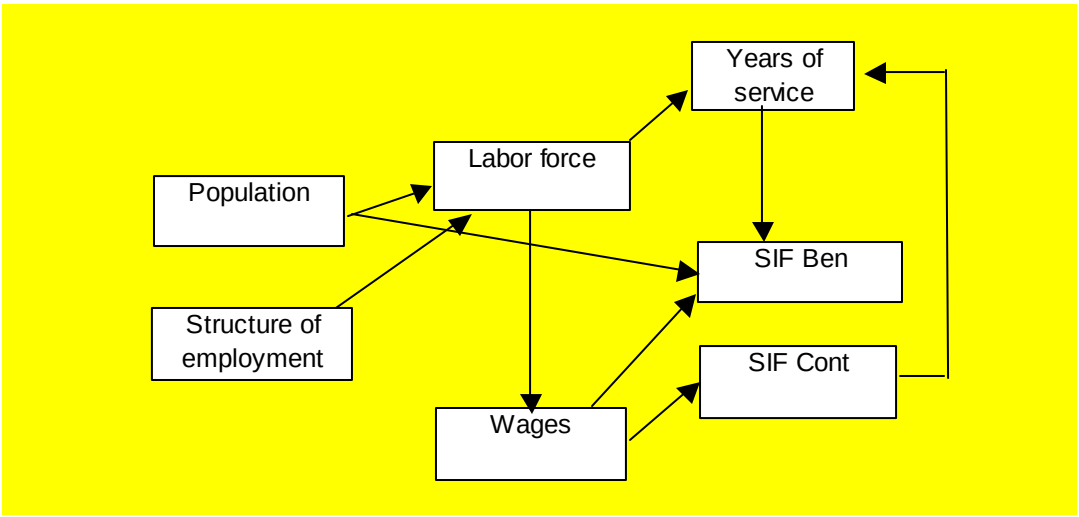
We often refer to "worksheets" inside a workbook. To switch between these, click on the tabs at the bottom of your screen.

Model summary

The workbooks comprising the model are contained in the directory "Version 0.0 Current," where "current refers to the fact that these workbooks model the pension system in its present form. The files in a different directory, "Version 0.0 Reform" permit the user to simulate the effects of introducing a second fully funded pillar.

The following schematic roughly summarizes the structure of the model:

The SPPS model: an overall schematic



- The workbook "Structure of employment" breaks down all employment, whether formal or informal, into economic sectors.
- "Population" performs a cohort-component projection of the population, on the basis of the 2001 Census, by single-year age group, sex, and urban-rural residence. In "Population," separate worksheets deal with the urban and rural population.
- Based on the population projection, "Labor force" – actually two workbooks, "Labor force urban" and "Labor force rural" -- calculates labor force and employment (thus, unemployment as well) by single-year age group, sex, and urban-rural residence. It applies the distribution from "Structure of employment" to allocate employed workers across economic sectors. Based on assumed "informality rates," i.e. the proportion of employed persons in a given economic sector who are informally employed), it splits employment into formal and informal employment.
- "Wages," which like "Labor force" consists of separate urban and rural workbooks, estimates nominal rates based on recent data; it also imposes assumed male-female and formal-informal (but not urban-rural) wage differentials if desired. Assumptions regarding increase in wage rates are described below, but basically the evolution of wages is taken to follow the evolution of GDP per capita. Wage rates times number of persons employed gives the wage bill.
- The formal sector wage bill is the basis for mandatory social contributions, which are calculated in "SIF Cont." In order to estimate these, we make assumptions on the compliance rate, i.e. the proportion of formal sector workers who actually make the required contributions. We also adjust downward for under-reporting of income. Separate worksheets within the workbook deal with urban and rural areas.
- Each year that a worker is employed in the formal sector and making required social contributions, he or she accrues one year of contributory service in the SIF system. The workbooks "Years of service rural" and "Years of service urban" estimate the number of years of contributory service accumulated by single year age group and sex.

- When a person reaches age 63 and retires, first, either he or she is or is not entitled to an old-age insurance benefit or not (almost all are). We estimate the number of persons qualifying, i.e. the number of new entrants into the old-age insurance beneficiary pool. The size of the benefit that those qualifying will receive in a given year (say 2010) will depend on the basic benefit prevailing in 2010, the number of years of SIF contributory service accrued when they retired (which may be some years prior to 2010) and the value attached in 2010 to each one of those contributory years in 2010. In "SIF Ben," we perform all the required calculations, estimating the number of persons reaching 63 each year who are eligible for a SIF old-age insurance pension and the size of the pension they are entitled to. We then "survive" them forward in time, each year adjusting the pension they receive based on the changing value of the basic pension and a year of accrued pension rights. As we assume that the both components of the pension are indexed to wages, the workbooks "Wages" are also relevant to these calcualtions. In "SIF Ben," we also calculate, based on "Population" and recent SIF data, the number of persons entitled to other types of pensions. These include labor and social disability pensions, special pensions (privileged, partial privileged and long service), insurance and social survivors pensions, and finally the old-age social pensions.

That, in a nutshell, is the SPSS model as it applies to the first- and zero-pillar pension systems in Armenia. Further modules are used to simulate transition to a reformed regime in which fully-funded second-pillar system plays an important part. However, those modules are still being modified to take into account new elements of the proposed reform as they become available. [For the flavour of these, however, see the document "Towards second-pillar pension reform in Armenia: transition cost estimates."] We do not discuss them in this manual.

The remainder of this manual goes through the SPSS model workbook by workbook.

Population and demography

The workbook "Population" performs a standard cohort-component projection of population by sex, rural-urban residence, and single-year age group (0-1 to 100+).

"Population" projects population using a standard cohort-component methodology. Here is how.

Fertility. The worksheets "Fertility_Urb" and "Fertility_Rur" contain assumptions on age-specific fertility rates (ASFRs), i.e. the number of babies born per woman, by age. Following convention, we use five-year age specific fertility rates for women aged 15-19, 20-24, and so on up to 45-49 (see Lines 137-143 of the worksheets).

For 2001-2005, we rely on age-specific fertility rates used in the Government of Armenia Population Strategy. In the worksheet "Notes," these are given for 2001 in Lines 130-36, for 2002 in Lines 140-46, and so on. The Population Strategy gives age specific fertility rates for all of Armenia. In order to disaggregate urban and rural, we use differentials calculated from age specific fertility rates presented in the *Demographic Yearbook 2006* (see Lines 182-90). Note that the fertility rates published in the Yearbook are much higher than those used in the Population Strategy.

In the worksheets "Fertility_Urb" and "Fertlity_Rur," age specific fertility rates by five-year age group for 2001-2005 are shaded turquoise to indicate that these have

been copied in from "Notes." The total fertility rate (TFR) in Line 145 of the two worksheets is calculated based on age-specific fertility rates. The TFR is a synthetic index which tells us the number of babies that a hypothetical woman would have if she lived from age 15 to 49 experiencing childbirth at the age-specific fertility rates observed in that year. The TFR cells for 2001-2005 are shaded yellow to indicate that these are calculated based on the values in other cells.

For the years 2006-2035, we rely on projections of the TFR and calculate age-specific fertility rates assuming that the age pattern of fertility observed in 2001 remains constant over time. The path of the TFR (see "Notes," Lines 197-217) is consistent with a return to replacement level fertility in 2035. After 2035, the TFR is assumed to remain constant at replacement level.

Having estimates and projections for age-specific fertility rates at 15-19, 20-24, etc., we now want analogous rates by single year age group. In "Fertility_Urb" and "Fertility_Rur" Lines 5-44 we follow the expedient of assuming that the five-year age-specific rate (which is the female population-weighted average of single-year age specific rates) applies to each single-year age group within the five-year range. EG, the age-specific rate at 20 is equal to the age specific rate at 21, 22, 23, and 24; all of them being set equal to the age-specific rate at 20-24 estimated as described above.

Births to women by single-year age group are calculated as the number of women in that age group multiplied by the age-specific fertility rate; these are then split into male and female births using an assumed constant sex ratio at birth. These calculations are made in Lines 48-131 of the worksheets "Fertility_Urb" and "Fertility_Rur."

Mortality. The worksheets "Mortality_Urb" and "Mortality_Rur" contain estimates and projections for age-specific death rates (ASDRs) by sex. The *Demographic Yearbook 2006* gives number of deaths by sex and age for the age groups 0-4, 5-9, and so on up to 85+. Since these are not disaggregated by urban-rural residence, we have made the assumption that urban and rural mortality rates are the same, while retaining the same urban-rural worksheet structure in case data later become available.

The implied age-specific death rates, calculated in Lines 41-78 of "Notes," are copied into copied into Lines 427-67 of "Mortality_Urb" and "Mortality_Rur." As in the case of fertility, we assume that the five-year age-specific death rates obtain at each of the single-year ages in the relevant range (e.g., death rates at 20, 21, 22, 23, and 24 are all assumed to be equal to the observed death rate for persons aged 20-24), except for in the terminal 85+ age group. For age 85, 86, …, 100+ the ad hoc estimates in Lines 83-98 of "Notes" were used. As a check, these were applied to the estimated population by single year age group over 85 and it was confirmed that the resulting ASDR(85+) was close to the one reported in the Demographic Yearbook.

Just as we base projections of ASFRs on assumed changes in the TFR index, we base projections of changes in ASDRs on assumed changes in the life expectancy at birth. This synthetic index, often written E0, is the number of years that a hypothetical person would live if he or she were born and aged experiencing the age-specfic death rates observed in a given year. In 2001, E0 in Armenia was 73.3 for both sexes combined (69.6 for men ands 77.0 for women). If E0 increases by 5 percent by 2050, this would bring it to 75.7 (71.6 for men and 79.9 for women). To implement this assumption, we assume consistent equiproportional declines in all ASDRs out to 2050, after which assume them to remain constant. Five-year ASDRs are given in

Lines 427-67 of "Motality_Urb" and "Mortality_Rur"; the breakdown into single-year age groups is in Lines 5-210. In Lines 217-422, urban and rural deaths by sex and single-year age group are calculated. The male and female life expectancies implied by ASDRS are calculated (for checking purposes only) in Lines 474-678.

Internal Migration. The worksheet "Rur-Urb-Migration," which has exactly the same structure as the mortality worksheet, contains sex- and age-specific urban-rural migration rates (net rural-urban migration in a given year expressed as a proportion of the rural population). These are currently set equal to zero.

International migration. The worksheets "Int_Migration_Urb" and "Int_Migration_Rur" have the same basic structure and calculate the number of net emigrants (persons leaving minus persons arriving) by age and sex. For males aged 15-59, the net international migration rate (number leaving minus number arriving as a proportion of population) is set equal to –4.4 per 1,000 in both urban and rural regions. International migration of women and of males aged under 15 and over 59 is assumed to be zero.

Projecting population. The base-year population by single-year age group is taken from the 2001 population census of Armenia. In the 2006 Demographic Yearbook, the terminal age category is 85+. The ad hoc distribution of population over 85 shown in the worksheet "Notes" (Lines 12-27) is used to fill out population by single-year age groups up to the terminal category of 100+.

The worksheets that project population forward in time are "Pop_Urb" dnd "Pop_Rur." In a given year, the population aged 0-1 is taken as the number of births (calculated in "Fertlity_Urb" and "Fertility_Rur"). For all age single-year age groups 1 and above, taking the male urban population as an example:

```
Pop_Urb_Male(a,t)  = Pop_Urb_Male(a-1,t-1)
        - Deaths_Urb_Male(a-1,t-1)
        + Rur_Urb_Migration_Male(a-1,t-1)
        + Int_Migration_Urb_Male(a-1,t-1)
```

The worksheet "Pop_Tot" simply aggregates urban and rural into nationwide population.

Labor force

Labor force lies at the heart of the SPSS model, not only in its own right, but because the nature of the labor force determines the number of persons making mandatory social contributions, the number eventually qualifying for social insurance benefits, and the level of benefit to which they are entitled.

In the workbooks "Labor force urban" and "Labor force rural," active labor force, unemployment, employment, and employment by major economic sector are calculated by age and sex. Base data from the 2001 Census are given in the worksheet "Notes." The five-year age-specific labor force participation rates (ASLFPRs) for men and women in columns H and I are copied into Lines 425-485 of the worksheets "LF_Urb" and "LF_Rur." The initial age group for labor force participation is 15-19 and the terminal age category in

published Census data is 65+. We make the simplifying assumption that labor force participation rates are constant above age 65. Lines 4-209 of the "LF_Urb" and "LF_Rur" worksheets contain ASLFPRS by single-year age groups, estimated by assuming that single-year rates within broad five-year age groups are the same (and similarly for the terminal 65+ age category).

We make the assumption that age- and sex-specific labor force participation rates remain constant moving forward in time, an assumption that can easily be varied. Lines 215-420 calculate labor force as the product of single-year ASLFPRs and population from the worksheet "Population."

The worksheets "Unempl_Urb" and "Unempl_Rur" are analogous to "LF_Urb" and "LF_Rur." Base-year (2001) unemployment rates for men and women by five-year age groups are copied in from "Notes" and single-year rates are estimated based upon them. In the case of published unemployment data, the terminal age category is 70+.

Whereas we assumed labor force participation rates to be constant over time, we know from recent history that unemployment rates have declined, leading us to believe that they will probably be yet lower in the future. We implement this by making assumptions regarding aggregate unemployment rates for men and women (i.e., total unemployment as a proportion of total active labor force) and then applying the trend to age-specific rates. This is similar to the approach that we took in dealing with fertility and mortality.

In cells B445 and B469 of the worksheets, aggregate unemployment rates for men and women in 2001 are calculated by dividing the total number of unemployed by the total active labor force. For years 2002 and beyond, the worksheet calls for assumptions to be made on the aggregate unemployment rate – we have assumed declines through 2007 and then constant levels thereafter. The proportional year-on-year changes in the unemployment rate are then applied to each age-specific rate in Lines 429-443 (men) and 453-67 (women). This is equivalent to assuming that, whereas the total unemployment rate may change, age structure of the unemployed population will remain the same.

The active labor force minus the number of unemployed persons gives the number of employed persons. We divide employment into eight economic sectors, given in the accompanying table. The workbook "Structure of employment" contains calculations on the sectoral breakdown of employment based on the 2001 population census. In the respective sector worksheets of "Labor force urban" and "Labor force rural," these shares are applied to employment by age and sex to calculate employment by economic sector by age and sex. Economic sector shares, assumed to be constant over time, are copied into Lines 4 and 5. In Lines 13-219, these shares are multiplied by total employment to calculate the number of persons employed in a given economic sector by age and sex. Lines 7 and 8 contain assumptions (also constant over time, although this can be changed) on the share of employees in a given sector who are in formal employment – i.e., who have a labor contract and are therefore subject to compulsory social contributions. These sub-sector shares are subjective, but when all are aggregated together, they result in a total formal-sector share that matches estimates made by other analysts for the urban and rural economies taken as a whole. In Lines 255-644, male and female informal and formal sub-sector employment in each economic sector is calculated. The worksheets "Empl_Urb_Tot" and "Empl_Rur_Tot" sum across sectors to give total urban and rural employment by age, sex, and formal / informal sub-sector.

Sector	Abbreviation
Agriculture	AG
Industry	ID
Construction	CO
Transport	TR
Wholesale and retail trade	WR
Social services (health, education)	SS
Public administration	PA
Other services	SV

Wages

Wages, differentiated by sex, economic sector of employment, formal / informal sub-sector, and urban / rural residence, are estimated in the workbooks "Wages urban.xls" and "Wages rural.xls." These are the largest spreadsheets in the model.

The foundation for wage estimation consists of assumptions made about real GDP growth and inflation as measured by the GDP deflator. The GDP deflator (2000 : 100) is the only price term in the model; there is no wage deflator or consumer price index (CPI). Assumptions for growth of real per capita GDP and the GDP deflator are entered in Lines 18 and 20 of the worksheet "Notes" in the workbooks "Wages urban.xls" and "Wages rural.xls." The rate of population growth, copied in from the worksheet "Population," is added to the assumed per capita GDP growth rate to calculate the rate of growth of aggregate GDP (Line 17).

A key assumption, entered in Line 12, is the difference between the rate of growth of the average (economy-wide) real wage rate and the rate of growth of real per capita GDP. We have assumed that the average real wage rate grows 0.5 percent per year faster than real GDP per capita through 2025, after which, it grows at the same rate.

Finally, Line 24 of "Notes" contains an assumption on the full-time equivalent factor. While we have set this equal to 100 percent throughout the time period covered, the model allows the flexibility to build in the fact that not all employed workers work full time.

The overall assumptions in "Notes" are then built into each economic sector in the AG, ID, … SV worksheets of the workbooks "Wages urban" and "Wages rural." The average nominal wage rate in each economic sector is available from data in the *Statistical Yearbook 2006* through 2005. We do not have separate data for urban and rural regions, so we assume that there is no wage differential, i.e., we use the same average nominal wage rate for urban and rural regions. In Line 6 of each economic sector worksheet, we project the sectoral average nominal wage rate forward using the growth rate of the economy-wide average nominal wage rate copied in from "Notes." This means that we assume that economic sector

wage differentials observed in 2005 (the last year of hard data) remain the same in proportional terms.

The model allows for building in assumptions on two wage differentials in each economic sector – between men and women and between the formal and informal sub-sectors. These differentials are expressed as "multipliers" in Lines 11 and 12, where we (arbitrarily) take the female nominal wage rate and the informal nominal wage rate as numéraires. The interpretation of Lines 11 and 12 is that we assume males in the economic sector earn 20 percent more than females and that the average nominal wage rate in the formal sub-sector of the economic sector is 10 percent higher than it is in the informal sub-sector. We make the same assumptions regarding these differentials in all eight economic sectors, although we are not bound by this simplifying assumption.

At this point, we have an accounting consistency problem to solve. In "Labor force urban" and "Labor force rural" we have estimated the number of male and female, informal and formal sub-sector employees in each economic sector. The weighted average of these labor force totals and the corresponding nominal wage rates must, by definition, be equal to the average nominal wage rate for the sector. But we have calculated the average nominal wage rate for the sector separately, based on the growth rate of the average economy-wide nominal wage rate. In order to force consistency, we invoke a standard modelling trick in the form of a "squasher." In Lines 17-20 of each economic sector worksheet, we calculate male and female, formal and informal sub-sector nominal wage rates according to the sector male-female, formal-informal wage differentials assumed. Line 17 calculates the sector-wide average nominal wage rate given these wage rates and the structure of the labor force in the sector that emerges from "Labor force urban" and "Labor force rural." In Lines 24-27, we "squash" these wage rates in Line 17-20 to enforce consistency.

We do *not* build in an assumed age-wage curve according to which wages earned vary by age and, implicitly, years of experience. The reason is simple: we already have four wage rates in each sector (male and female, formal and informal); building in an age-wage curve would multiply this to about 350, each needing to be squashed to ensure consistency.

The many remaining lines in each economic sector worksheet of the workbooks "Wages urban" and "Wages rural" calculate total wages paid to male and female, formal and informal sub-sector workers in the sector. Finally, the worksheets "Wages_Urb_Total" and "Wages_Rur_Total" calculate total wage bills (across economic sectors) for males and females in the formal and informal sectors.

First-pillar pension system contributions

In the workbook "SIF Contr" the worksheets "SIF_Contributors_Urb" and "SIF_Contributors_Rur" calculate the number of persons making mandatory social contributions to the SIF system. In "Notes," the number of contributors in 2005 (from SIF) is divided by the number of persons employed in the formal sector (imported from the worksheets "Labor force urban" and "Labor force rural" to calculate average male and female compliance rate – average over all age groups (15+), over urban and rural regions, and over sectors of employment. The (roughly first 200 lines of "SIF_Contributors_Urb" and "SIF_Contributors_Rur" calculate total formal sector employment by sex. The next (roughly) 200 lines copy in or calculate age- and sex- specific compliance rates. For 2001-2005, these simply replicate the estmates based on the calculation made in "Notes." For 2006 and beyond, the user enters a number for the average sex-specific compliance rate (in Lines 319 and 424). Age-specific rates are calculating using the implied proportional change from

year to year; all this does is to enforce the assumption that all age specific rates are equal t the assumed average. Note that compliance rates under age 15 are assumed to be zero; there is no reason for the user to change this assumption. The worksheet "SIF_contributors_tot" simply sums urban and rural contributors by age and sex.

In Line 3 ad 10 (corresponding to men and women, respectively) of the worksheets "SIF_Contributions_Urb" and "SIF_ Contributions_Rur" we calculate the average SIF contribution. From 201-2005, we do this by applying the following SIF contribution formulae:

Employer: Contribution = AMD 7,000 + [0.15 * (GrossWage – AMD 20,000)]

Employee: Contribution = 0.03 * GrossWage

The 3% assumption for the employee contribution is an assumption entered in Line 6 for men and Line 13 for women. For gross wage, we import the average gross formal sector wage rates calculated in "Wages urban" and "Wages rural."

In Lines 4 and 11, for 2001-2005 we calculate employers' implied average mandatory social contribution rate, dividing the calculated average employer nominal contribution by the average nominal gross wage. In 2006 and thereafter, we invert the approach: we make an assumption regarding the average employer contribution rate and calculate the average nominal employer contribution (Line 5 for men and Line 12 for wmen) as the assumed rate multiplied by the average gross wage. The main reason for this inversion is that the current employer contribution calculation formula given above is unlikely to be the same over the long term, or even in a few years – pension authorities will no doubt adjust it in line with inflation and the growth of real wages.[1] To repeat and emphasise: over the projection period, the average employer and employee contribution rates represent assumptions.

It is no secret that, not only do many persons evade making any mandatory social contributions at all, but many of those who do under-report income. Lines 7 and 14 allow the user to make an assumption of the extent of this by entering an "evasion factor" defined as the proportion of owed contributions that are actually made (by those who have not evaded the system altogether; we have already built that assumption into the worksheets "SIF_Contributors_Urb" and "SIF_Contributors_Rur").

Lines 8 and 13 sum the employer and employee contribution rate, then multiply by the evasion factor, to calculate average effective male and female contribution rates, again for those who are not evading mandatory social contributions the system altogether.

The remaining lines of the worksheets "SIF_Contributions_Urb" and "SIF_ Contributions_Rur" calculate contributions by age and sex, and these are summed in "SIF_Contributions_Tot."

First-pillar (and zero-) penson system benefits

As benefits are calculated according to years of contributory service, the first worksheets we will be concerned with are "Years of service urban" and "Years of service rural."

[1] We do, however, assume that the 3 percent employee contribution rate remains the same, although the assumption can of course be changed.

It would be lovely if there were a simple, straightforward way of simplifying the task of calculating years of service. Unfortunately, there is no real substitute for tracking each single-year age cohort through its lifetime labor market history, cumulating for each year the SIF accrual credit gained. In order to do this, we need to "backcast" far into the past using rule-of-thumb estimates. Fortunately, results moving forward are not terribly sensitive to the assumptions that we make about the past.

For example, consider a woman reaching 63 in 2001. Her estimated years of contributory service would be

Pr (the woman was employed in formal sector aged 15 in 1953) x first-pillar compliance rate in 1953

plus

Pr (the woman was employed in formal sector aged 16 in 1954) x first pillar compliance rate in 1954)

…

plus

Pr (the woman was employed in formal sector aged 62 in 2000) x first-pillar compliance rate in 2000

plus

Years credited for military, maternity, and higher education

For a woman reaching 63 in 2010, the cumulation would begin in 1964 rather than 1953, and so on. Note that we deal with single year age cohorts rather than individuals, so we are speaking of the *average woman* reaching 63 in the given year.

In order to apply this approach, we need to "backcast" formal sector employment rates and first-pillar compliance rates, that is, the proportion of formal sector workers who actually contribute as required (albeit perhaps not in the amount required). We do this according to the simple set of rules given in the following box. In a nutshell, we assume that levels of informality and unemployment observed in 2001 applied throughout the period 1989-2000 and that in prior years there was no unemployment and no informality.

These calculations are done in the workbooks "Years of service urban" and "Years of service rural," eac of which consists of only one worksheet. Lines 2 and 3 of the worksheets contain assumptions on the proportion of male and female employees who work in the formal sector. From 2001 forward, these are copied in from the workbooks "Labor force urban" and "Labor force rural," hence, these cells are shaded in turquoise. For 1989-2000, assumptions are set equal to their level in 2001, hence these cells are shaded yellow. From 1926-1988, assumptions must be entered manually, so these cells are shaded green.

Lines 5 and 6 contain assumptions on men's credit for military service and higher education and women's credit for maternity and higher education. We assume 2 years' credit for both sexes throughout the time period covered.

Assumptions underlying estimates of years of contributory service			
	1988 and prior	1989-2000	2001 -
Labor force participation rate	As in 2001	As in 2001.	2001 Census data by age, sex, and rural-urban residence; assumed constant moving forward.
Unemployment rate	*Nil*	As in 2001.	2001 Census data by age, sex, and urban-rural residence; adjusted to reflect Labor Force Survey data 2002-2005 and adjusted downward moving forward.
Proportion of employed in formal sector	*Nil*	As in 2001.	Authors' estimates by economic sector (agriculture, industry, etc.); same for all ages, men and women, urban and rural.
Proportion of formal sector employed complying	All	As in 2001.	2005 SIF data for males and females (urban and rural combined, all ages taken together).
Years of credit for military, maternity, and higher education	2 years credit for all men and all woman, whatever year they retire.		

Lines 10 through 110 give, year by year and by single-year age group, the number of years of formal sector employment accumulated by men. For 2001 and forward, this is calculated based on data imported from "Labor force urban" and "Labor force rural." In Year t the number of years spent in formal sector employment accumulated by a person aged a is

Labor force participation rate (A,t) x [(1 – Unemployment rate (A,t)] x Proportion of employment in formal sector in year A

where the labor force participation rate and unemployment rate are age-indexed but the proportion of all employment that is in the formal sector is not.

For 1989-2000, formal-sector employment accrual is assumed to be the same as observed in 2001. For 1926-1988, we assumed zero unemployment and no informality (so that the

number of years of formal-sector employment accrued is equal to the labor force participation rate).

Lines 111 and 217 give the number of years of formal sector employment that would be accrued by an average person who survived to 100 according to the labor force participation, unemployment, and formal – informal sector breakdown of employment in a given year. Clearly very few people survive to 100, and this is for informational and tracking purposes only. Of more interest are Lines 221-226 and 230-235, which give the average number of years of formal sector employment accrued by the average man and woman aged 60, 61, 62, 63, 64, and 65.

Years of formal sector employment history is only half the story; we need to know how many years of contributory SIF history were accrued. In Lines 239-339 (man) and 344-445 (women), we multiply years of formal sector employment by SIF compliance rates imported from the workbook "SIF Contr"

(where they were assumptions). From 1989-2000 we assume that compliance rates were equal to their 2001 level; previous to that, we assume that there was universal compliance (as well as coverage).

Lines 340 and 446, for tracking purposes only, show the number of years of SIF contributory history that would be accrued by a person who survived to 100 experiencing the labor force participation, unemployment, informality, and compliance rates estimated for that year.

Of more importance are lines 450-455 (men) and 460-465. These give the result we are after; the average number of years of SIF contributory history accrued by a member of the single year age cohort retiring in that year. This is a "backward looking" variable; it builds in the entire labor market history of the cohort between the ages of 15 and retirement. While we focus on retirement at 63, we provide estimates for ages 60-65. Lines 456-457 and 466-467 are of interest later when we simulate the transition to a different pension system, because these divide years of accumulated SIF contributory history into years accumulated under the pre-reform SIF and years accumulated under the post-reform SIF.

Having estimated years of contributory history, we are in a position to estimate benefits according to the prevailing formula. This is done in the workbook "SIF Ben." Since this workbook is constantly using population date, we put a worksheet at the end, "Pop_Tot" that simply imports population data from the workbook "Population." That way, other worksheets can import population data from within the workbook.

Information from SIF on the level and structure of benefits as estimated (budgeted, actually) for 2007 are contained in the worksheet "Notes." Lines 6-11 give population in various age groups relevant for SIF benefit programs. In the following table, we specify these. Some are approximate, for example, there may be a few persons qualifying for special pensions before age 50. By and large, however, the table gives a reasonable picture of the target populations for various first- and zero-pillar pension programs.

Age group	Relevant programs	
	Insurance	Social
15+	Labor disability	Social disability, Social survivors
63+	Old age insurance, Old age survivors	
65+		Old age social
50-62	Special pensions: privileged, partial privileged, long service	

Lines 31-38 give the proportion of eligible populations receiving the benefit, as reported by SIF; Lines 42-48 give the total benefits paid out, and Lines 55-61 give the average benefit level. Note the importance of always distinguishing between annual and monthly figures; these are always explicitly marked in the various spreadsheets but it is easy to get confused nonetheless! Lines 66 –72 annualize the monthly figures calculated above. Finally, Lines 79-90 contain the calculation for an adjustment factor to years of service that we will discuss later.

Now we will go through the calculation of various benefit levels and payments. By far the most important of these from the standpoint of SIF finances is old-age insurance pensions. These are dealt with in the worksheet "OldAgeIns."

A crucial assumption is to be entered for 2008 and beyond in Lines 4-5 – the proportion of the population turning 63 that qualifies for an old-age insurance pension. We have assumed, in our baseline, that this proportion is 81.1 percent for men and 95.6 percent for women – i.e. that the proportion of the eligible population (63+) receiving the benefit in 2007 is also the proportion of the population turning 63 that will qualify. We also assume, for the sake of historical calculations, that the same assumptions apply for 2002-2006. The estimate for 2001 is copied in from "Notes," hence the turquoise shading.

In a special study, *Eligibility for the old-age insurance pension and implications for the social pension*, we analyzed eligibility data in detail. One point of view is that, with average years of SIF contributory history set to decline rapidly (a trend which we document and analyze in the same paper) the proportion of 63 year olds qualifying for an old age insurance pension will also decline. Another point of view holds that, with the required minimum contributory period so low (5 years, towards which maternity, military service, and higher education count), almost everyone reaching retirement age will qualify, as they do now. In our baseline, we have adopted the assumption of continued high eligibility, however, we have also done sensitivity analyzes.

Lines 7 and 8 import average years of SIF contributory service from "Years of service urban" and "Years of service rural." Since we do not have rural-urban disaggregated on benefits or

contributions from SIF, we aggregate the rural and urban years of service estimates into a weighted average using the shares 2/3 urban and 1/3 rural.

Line 9 contains an adjustment factor used to reconcile average years of service from "Years of service urban" and "Years of service rural" with average years of service estimated base on SIF data. These are the calculations in Lines 80-90 of "Notes" referred to earlier. According to SIF, in 2006 a little under 300,000 persons began to receive old age insurance pensions (we make the assumption that this was at age 63) with 26 or more years of contributory service while a little more than 50,000 retired with less than 26 years of contributory service. The first group retired with a pension of about AMD 13,000 per month while the second group retired with an initial benefit of about AMD 7,500. Because we have the SIF old age pension benefit formula, we can use this to assume that the first group retired with an average of 37.9 years of contributory history while the second group retired with an average of14.1 years of contributory history. Because we have the shares, we can "back out" that the average person who began receiving an old-age pension benefit in 2006 did so with 34 years of contributory history. "Years of service" estimates an average contributory period of about 27.5 years, suggesting that we should impose an upward adjustment of 6.5 years in 2007 in order to line up. All of these calculations are approximate and, for simplicity, we apply an upward adjustment of 5 years over the entire period of model calculation. It is important to note that, given the fact that the basic pension is so high as compared to accrual component, the size of the adjustment will not make much difference to the calculated pension. Lines 10 and 11 contain the final calculated years of service at retirement, starting at about 35 for men and 30 for women in 2007.

Lines 13 and 14, like Lines 10 and 11, are crucial to the calculation of pensions. Line 10 contains the basic pension, i.e. the component of the old-age insurance pension that is not related to years of contributory service. Line 11 contains the accrual value of a single year of contributory history.

For 2001-2007, these figures are historical. For 2008 and beyond, they are calculated by multiplying the previous year's parameter by an indexation factor, given in Line 19 for the basic pension and 20 for the service-related component. We discuss the calculation of the indexation factor in a moment; for the moment, what is important to keep in mind is that the indexation factor, and the basic pension / accrual value of a year's service are both in nominal terms.

Lines 16 and 17 simply replicate, for tracking purposes, the average monthly benefit for a person retiring in year t at age 63 as calculated in Lines 328 and 435.

The nominal indexation factors in Lines 19 and 20 are calculated with reference to the growth of the nominal formal sector wage rate, calculated in Line 25 based on the average (over men and women, urban and rural) formal sector wage rate. This is calculated, in turn, based on data imported from the workbooks "Wages urban," "Wages rural," "Labor force urban," and "Labor force rural"; hence the turquoise shading.

Lines 27calculates the monthly total old age insurance pension benefits paid out to men and women; Line 28 multiplies by 12 to annualize this. Lines 30-31 calculate average male and female monthly pensions by dividing the monthly total by the number of pensioners.

We now have a line-up or consistency issue. By the mechanisms described above, we will have calculated average male and female old-age pensions in 2007. But SIF has reported its budgeted estimates of the average male and female old-age insurance pension in 2007. To force consistency, we compare the average pension calculated with the model in 2007 to the average pension observed, calculate a "squashier," and apply this to every age-specific

pension benefit as described below. In theory, every time the model is rerun with any changes affecting 2007the "squashier" must be re-estimated; in point of fact, so long as a reasonably close line-up with 2007 is acceptable, it is alright to leave the squashier alone. If it is desired to re-calculate it, do via an iterative process:

SquasherMale (i) = SquasherMale(I-1) x [Average old age insurance pension male 2007 SIF / Average old age insurance pension 2007 calculated from model(i)]

SquasherFemale (i) = SquasherFemale(I-1) x [Average old age insurance pension female 2007 SIF / Average old age insurance pension 2007 calculated from model(i)]

where i indexes the first, second, third, etc. iteration Three or four iterations should be sufficient to line up for 2007. Note, of course, that if the required squashier is assuming large proportions, one must question whether there is not something erroneous or unreasonable about the assumptions that have been made for 2007. In this way, the squashier serves not only as a convenient line-up device but as a reality check as well.

Having calculated the squashier for 2007 and applied it to age-specific pension benefits, we apply it to all other years of the model solution, as well. This is required to avoid jump-off problems.

Lines 39-43 contain various variables calculated for tracking purposes. These include the average real pension (deflated by the GDP deflator imported from "Wages urban") and the ratio of the average monthly pension to the average monthly formal-sector wage.

The remainder of the workbook – over 600 lines – calculates number of beneficiaries, average pension received, and total pensions paid out by sex and age. Lines 52-257 calculate the number of beneficiaries. From 2001-2007, this is simply the ratio of pension recipients to population aged 63+ observed in 2007 multiplied by population. From 2008 on, however, we implement a more structural, survival-based, approach. New retirees at age 63 are calculated as the assumed proportion of 63 year-olds eligible (Line 4 for men and Line 5 for women multiplied by the population aged 63. Remember that we made the simplifying assumption that, for 2007, the proportion of persons 63 qualifying for the old age insurance pension could be approximated by the proportion of the 63+ population actually receiving one; this is our "high eligibility" baseline assumption which we tested by also running a "low eligibility" scenario. In every year 2008, 2009, 2010 … the number of new entrants into the old-age insurance pension beneficiary pool is calculated based on assumed eligibility and the population aged 63.

As each single year cohort of new beneficiaries ages, we adjust population down to account for mortality and net international out-migration (although we assumed in "Population" that there was no migration at older ages, we keep the migration rte in this formula in case we ever want to change the simplifying assumption).

Thus

Male old-age insurance pension beneficiaries(64,2009) =
Male old-age insurance pension beneficiaries(63,2008)
x [1 – Mortality rate(63, 2008) + NetIntMigRate(63,2008)]

Male old-age insurance pension beneficiaries(65,2010) =
Male old-age insurance pension beneficiaries(64,2009)
x [1 – Mortality rate(64, 2009) + NetIntMigRate(64,2009)

… and so on

Lines 265-472 calculate average pension benefits by single-year age cohort. For a person entering the beneficiary pool at age 63, the average benefit received will depend on the basic pension and the accrual value of a year of contributory service, both measured in the year of retirement. For example,

Average monthly benefit(63,2007) =
Index-adjusted basic pension (2007) +
[Estimated years of contributory history (63,2007) x Index-adjusted worth of a year of contributory service (2007)]

Average monthly benefit(63,2008) =
Index-adjusted basic pension (2008) +
[Estimated years of contributory history (63,2008) x Index-adjusted worth of a year of contributory service (2008)]

… and so on

As beneficiaries age, we need to look backwards because, for example, if our 63 year-old who retires in 2007 survives to be 64 in 2008, his/her average pension will be

Average monthly benefit(64,2008) =
Index-adjusted basic pensions (2008) +
[Estimated years of contributory history (63,2007) x Index-adjusted worth of a year of contributory service (2008)]

Average monthly benefit(65,2009) =
Index-adjusted basic pension (2009) +
[Estimated years of contributory history (63,2007) x Index-adjusted worth of a year of contributory service (2009)]

… and so on

Remember two things:

- "average" here does not mean average over all members of the single-year age cohort, it means average over all members of the single-year age cohort who qualify for the old-age insurance pension.
- Not shown in the boxes, but described above, the average pension in each age group in 2007 is "squashed" to ensure consistency with the average pension over all age groups observed in that year and the "squashier" is applied to each year going forward (as well as going back in history).

Lines 478 and above of the worksheet "OldAgeIns" simply sum up male and female benefits to calculate totals paid out.

Calculation of other types of pension are much simpler. The worksheet "DisIns" deals with labor disability pensions. Column H, Lines 4 and 5 contains the proportion of the population aged 15+ eligible, copied in from "Notes." Lines 7 and 8 of the same column contain the average pension level, from the same source. For 2008 and beyond, we assume that the proportion of the population eligible remains constant; hence the green shading. For 2001-06 we assume that the rate is the same as in 2007. For 2008 going forward, we adjust the average benefit level by the indexation factor in Line 10; this is, in turn, assumed to be same as the indexation factor for the basic pension and is copied in from the worksheet "OldAgeIns." For history, we assumed on an entirely ad hoc basis that the level of the labor disability pension increase at 5 percent per year.

The remaining lines of the spreadsheet calculate benefits paid out by age and sex.

The approach implemented for survivor's insurance for survivors insurance in the worksheet "SurvIns" is exactly the same.

The approach implemented for special pensions (privileged, partial privileged, and long service) in the worksheet "Special" is exactly the same, except that the eligible population is aged 50-62 and the proportion benefited is reduced in 2009 and beyond. Eventually these are eliminated altogether (although the user can easily change this assumption to estimate the savings made by eliminating special pensions).

The approach implemented for social disability pensions in the worksheet "SocDis" is the same; the eligible population group being the population aged 15+.

The approach implemented for social survivor's pensions in the worksheet "SocSurv" is the same; again, the eligible population group is the population aged 15+.

The approach implemented for old-age social pensions in the worksheet "OldAgeSoc" is slightly different. As elsewhere, 2007 data on the proportion of the eligible population (65+) benefiting and the average monthly pension received are imported from "Notes." Again, we make the ad hoc assumption that the social pension increased 5 percent per year between 2001 and 2007 and that after 2007, the social old age pension is indexed t the basic pension. We estimate the proportion of the eligible population benefiting, however as, taking 2010 as an example:

Proportion population 65+ receiving old-age social pension (2010) =
Proportion population 63+ receiving old-age insurancel pension (2010) –
Proportion population 63+ receiving old-age insurance pension (2009)

This is a purely ad hoc means of implementing the assumption that, if eligibility for the old-age insurance pension declines, the proportion of the eligible population receiving the old-age social pension rises.

Tables and charts

The workbook "Tables and Charts" extracts highlights from the model solution and presents a series of tables and charts for use in report writing. All are in printer-friendly format.

The worksheet "Major Assumptions" contains a summary table in which major assumptions are summarised for 2005 and 2010-60 in ten year intervals. GDP growth and inflation figures (Lines 17) are imported from "Wages_Urban."

The worksheet "SIF_Rev_Exp" presents SIF accounts consistent with the model solution. Line 4 contains revenues in the form of mandatory social contributions. Lines 6-10 import insurance pension expenditures from "SIF Ben." In Line 11, we estimate that funerals, internal payment costs, and SIF administration cost ten percent of total insurance pension expenditure. In making this estimate we have been guided by the need to approximately line up on SIF accounts for recent years and budgeted expenditure in 2007. Note that the line up is approximate, however -- we make no claims to precisely reproduce SIF's historical budget (but we come close). In Line 13, we add another 10 percent to insurance pension expenditure to account for maternity and unemployment benefits and the purchase of external payment services. The balance in line 15 refers to those payments that are financed by mandatory social contributions.

Lines 19-21 display SIF social pension expenditures, which are financed by the State budget. Payment costs, administration, etc. are estimated at 10 percent. We make no attempt to account for military pensions.

Line 25 calculates the SF balance (which leaves out social pension programs, these being financed by the State budget) as a share of GDP.

Lines 27-29 calculates SIF insurance program, SIF social program, and total expenditures as shares of GDP.

The worksheet "Key ratios" implements the decomposition of trends in the ratio of the average pension to the average wage into trends in various constituent ratios following the decomposition analysis described in the SPSS Project "Status Paper" on the Armenian pension system.

Conclusion

This note has described a model that can be used to project the state of the Armenian zero- and first-pillar pension systems in their current state. The impact of a wide variety of alternative assumptions, demography, the labour market, and the basic parameters of the pension system, can be estimated.

AN expanded version of this note will describe how an expanded version of this model can be used to estimate the impacts of introducing a fully-funded second pillar.