A microsimulation framework for projecting public pensions' sustainability. Simulating labour-market transitions and contribution years in Argentina.

Leonardo Calcagno.

Univ. Orléans, CNRS, LEO[1]

Abstract

Using a standard microsimulation framework, this paper intends to draw perspectives on Argentina's pension system from a household-level quarterly survey data, the EPH[2]. After describing the country's current public Pay-As-You-Go pension system, this paper uses the country's household survey's labour market related data to describe and estimate transitions between labour market states for individuals of working age. Finally, together with demographic data provided by Argentina's statistical body INDEC, we describe how we intend to use these empirical labour market transition probabilities as inputs for simulating forward, up to 2040, and backwards, labour careers of our surveyed population. We should hence be able to reconstitute the labour career of each individual, i.e. the total amount of contributive years each individual has validated once he reaches retirement age. This would lay the basic framework for future projections on Argentina's pension system's mid to long-term sustainability.

[1](Univ. Orléans, CNRS, LEO, UMR 7322, F45067, Orléans, France). e-mail: leonardo.calcagno@univ-oleans.fr

[2]"Encuesta Permanente de Hogares", which stands for "Permanent Household Survey". In its quarterly form, covers the 2003-2013 period.

Introduction.

This paper is part of a broader project : to provide a reliable projection of Argentina's Unified Social Security System (SUSS), run by the "National Social Security Administration" (ANSES). We intend to simulate ANSES future income, expenses and ultimately financing needs, based on a set of different demographic, social and economic scenarios. This public agency manages social security benefits[3] as well as family and unemployment benefits. Its total non-figurative[4] expenses amounted to 9,7% of GDP in 2012 [Calcagno (2013), p 18]. In the context of this project, this paper aims to provide the general outlines of future simulations of pension rights of a sample population, which would mean reconstituting backwards its likely contribution years and forward its future contributions, given a set of demographic hypothesis and labour-market transition probabilities. Not only are contribution years essential to calculate contributive benefits paid by ANSES, they are also central for calculating its future contributive revenue. But before further expanding on this subject, let us briefly present why we chose to study Argentina's social security system.

In the past twenty years, Argentina's social security system has undergone various reforms. Starting from 1993, it has been run by ANSES, an autonomous agency that began functioning on that year in the midst of a constitutional reform. It was first meant to run a unified Pay-As-You-Go (PAYG)[5] pension system, as part of the two-pillars system that was implemented at that time, where the second pillar was a new fully funded pension system[6]. This PAYG pension system paid pension benefits to people that had retired in one of the previous existing PAYG systems and collected the contributions of the minority of workers that had explicitly opted for the public PAYG system. Apart from the predictable huge social security deficit this reform generated, the adequacy of retirement pension benefits decreased sharply for a number of reasons. On the one hand, since 1991 all indexation on inflation was banned as part of the "convertibility plan" that was launched in that year, in a move to establish a currency board to end with chronic hyperinflation outbursts. This was maintained throughout the 1990's, and meant for instance that the minimum pension was frozen at 150 pesos (150 US$) a month from December 1995 to December 2001[DNPE (2012), p 28]. On the other hand, private pensions had a defined contribution scheme that made benefits rely on the profits of the funds. And finally, the requirements to benefit either from a private or a public pension were toughened, leading to a drop in pension coverage, which was also due to an unstable and, in the late 1990's, recessive economy. At the same time, private fully funded pensions had on average very high operative costs and although

[3]Mostly retirement benefits and survivors pensions, but also unemployment benefits.

[4]Figurative expenses stem from programs run on behalf of other government agencies, and are balanced out by figurative contributions, which are transfers from the Treasury.

[5]In a PAYG pension system, today's contributions to the system pay for today's pensions. By contributing, a worker acquires rights for future pensions, paid by future workers.

[6]Here, each person contributes to its individual account, that is invested by private pension funds. The return of the invested funds pays for that fund's current retirees, depending on the amount these had previously contributed.

they had fewer retirees than ANSES, starting from 2002[7] the federal government had to subsidise some of the fully funded pensioners since their pensions did not reach the minimum retirement benefit.

ANSES' role became however increasingly important starting from the 2000's, first with the 2005 pension moratorium Law 25994 that came into effect in 2007 and raised coverage rates from approximately 50% of retirement-age population to almost full coverage in 2010. Together with a previous independent workers' moratorium Law 24476, "these two moratoriums now have a bigger coverage than PAYG pension plan's. In 2011, 45,35%[8] of retiring age population had a retirement pension by moratorium, while 32,36%[9] of this population had a retirement pension through the regular PAYG pension plan." [Calcagno (2013), 19]. The Pension mobility Law 26417 was adopted on October 2008, and it indexed pensions and social security benefits on an average of the evolution of wages and ANSES revenue. But at the same time, following the outbreak of the global financial crisis that hit hard Argentinian private pension funds, the nationalisation of private pension funds was decided. The two pillars of the system were fused into a single PAYG pension system, the Argentinian Integrated Pension System (SIPA). This meant on the one hand that ANSES, through its national pension fund, the FGS, which was created the previous year, began administrating almost 10 GDP points worth of individual pension accounts and receiving most of retirement contributions; but on the other hand it meant the federal PAYG system had to fully pay most pension benefits[10]. However, as it has been pointed out in the literature by the likes of Mesa-Lago, this reform "[was] not supported by actuarial studies and apparently they have not been done after the re-reform" [Mesa Lago (2014), p 20]. If we add to these considerable institutional changes a continued demographic transition, with a gradual ageing of Argentina's population, a middle to long-term projection of future Argentinian social security system's financing needs becomes all the more important.

The purpose of this paper is to lay the foundations of a forthcoming microsimulation of ANSES mid to long-term financing needs by establishing the methodology for simulating contributed years at retirement. To do so, we will describe the methodology we intend to use, as well as some descriptive statistics and econometric analysis of our household survey data set.

[7] Source: CGN's Evaluation of results by program and project, social security agencies, 2002.

[8] Source: own calculus from CGN's physical-financial follow-up and CELADE's data.

[9] Ibid.

[10] There are Provincial and municipal PAYG pension funds for some provincial workers, where around 15% of total contributing workers are affiliated.

Literature review.

Dynamic microsimulation literature.

The microsimulation literature in economics is quite old, as it traces back to Orcutt's[Orcutt (1957)] seminal paper. As Li and O'Donoghue explain, "[[Orcutt et al. (1961)]] described the first dynamic microsimulation model following the inspiration of Orcutt's (1957) article. Most dynamic microsimulation models that have developed in the following decades trace a direct or indirect link back to this model." [Li and O'Donoghue (2013), p 4]. Basically, a microsimulation model simulates unit-level behaviour, accounting for the diversity of characteristics and of possible responses among the population. In the fields of economics and sociology, it is mostly used to analyse social economic policies. Static microsimulation models simulate an immediate response to a shock, without a time dimension[11] . Dynamic microsimulation models, on the other hand, model this behaviour over time. As explained by Li and O'Donoghue, "dynamic models (...) extend the static model by allowing individuals to change their characteristics due to endogenous factors within the model (...) and let individual units to progress over time"[Li and O'Donoghue (2013), p 4]. These are the most common in the economic field, since they make it possible to generate synthetic micro-level data useful for prospective analysis (typically, social security and pension plans [12], but also for instance future income distribution[13]). In our case, we intend to carry out a dynamic microsimulation model about retirement and social security contributions and validated contributive years. We will hence study more closely the literature on social security and pension plans microsimulation models.

However, one of the major problems of the field has been its fragmentation. As Li and O'Donoghue explain, "microsimulation models are mostly developed in governmental or policy institutions, where developing a literature on which a wider group of scientists has built has been a lesser objective. Furthermore, the documents are mainly spread with limited books and conference presentations, which may not be easily available for researchers outside of the network. (...)Thus a significant proportion of the extensive methods used in the field are not formally codified, meaning that to a large extent new models have had to reinvent the wheel and redevelop existing methods over and over again. This has made it very difficult to work in the field"[Li and O'Donoghue (2013), p 34]. In addition to the intrinsic complexity of microsimulation models and of their

[11] Some examples of static microsimulation models are the EUROMOD[Sutherland (2007)] model and the IZAΨMOD[Peichl et al. (2010)] model.

[12] "Models such as LIAM[O'Donoghue et al. (2009)], PRISM[Kennell and Sheils (1990)], the Belgian dynamic model[Joyeux et al. (1996)], the SfB3 population model[Galler and Wagner (1986)], LIFEMOD[Falkingham and Johnson (1993)], SESIM[Flood (2007), Klevmarken (2010)] and Belgium MIDAS[Dekkers et al. (2010), Dekkers and Belloni (2009)] have all been used to look at pension reform"[Li and O'Donoghue (2013), p 6]

[13] [Li and O'Donoghue (2013), p 7] mention, among others, the DESTINIE1/2[Bonnet and Mahieu (2000), Blanchet (2009)], APPSIM[Harding (2007)], SESIM[Klevmarken and Lindgren (2008)] or SADNAP[Van Sonsbeek (2010)] models.

high computing power needs, these aspects hamper the transfer of knowledge and know-how in the field and make it very difficult to replicate previous results using the same methodology. It also meant entry costs in the field were for a long time particularly high. Moreover since a microsimulation model cannot be properly explained in a standard paper, most of the available literature are working papers or reports from organisations describing their microsimulation model. As such, this work is based mostly in working papers about existing microsimulation models for given pension systems.

Our main inspiration for the general outlines of our project comes from various pension-related dynamic microsimulation models. Direct references are INSEE's model Destinie 2[Blanchet et al. (2011)], CNAV's dynamic microsimulation model PRISME[Poubelle et al. (2006), Albert et al. (2009)] and DREES' TRAJECTOIRE[Duc et al. (2013)] model, among others. These are the leading models for microsimulation of pension policy in France, and although there are many others that currently exist [14]we chose to focus on these, as the basis of our microsimulation framework. Since none of these models are open source nor directly applicable to Argentina's pension system, our aim is to follow some of their methodological choices in our study of Argentina's social security system. This aspect is further discussed in our methodological section.

Literature on labour market transitions.

Our work has required we compute transition probabilities between labour market states for the individuals in our data set. Indeed, having a dynamic microsimulation model implies simulating transitions over time of single individuals between different states. So be it regarding demographic or labour-market characteristics, we have to specify for the states we will study a transition rule between them. Our demographic transition sources come from recent demographic projections from the INDEC. [15]

Regarding labour-market transitions, we have mobilised a wide array of economic literature to compute these from the EPH data set. On the one hand, a part of the literature calculates empirical transition probabilities between labour-market states, usually from individuals followed in household surveys. Most of these papers compute empirical transition probabilities at least for the sake of descriptive statistics. Some use a quarterly step, for instance, Silva and Vázquez-Grenno "obtain the gross labor market flows by calculating the quarter-on-quarter transitions made by individuals workers between different labor market states"[**?**, p 162]. Others have a yearly or bigger step, such as [Theodossiou and Zangelidis (2009)], [Zissimopoulos and Karoly (2007)] or [Bradley et al. (2003)] . Finally, others compare quarterly and yearly tran-

[14]Li and O'Donoghue [Li and O'Donoghue (2013), p 9] provide an exhaustive list of existing dynamic microsimulation models and of their uses.

[15]Both conditional probabilities of dying and probabilities of giving birth have been computed by INDEC from 2008-2010 life tables computed from the 2010 national Census. The data is available at INDEC's website http://www.indec.gov.ar, in the "projections and estimations" section of the "population" thumbnail.

sition rates, such as [Güell and Petrongolo (2007)]and [Gomes (2012)]. Usually these papers use these transition rates as the input for some econometric analyses, such as cross-correlation analysis [?], multinomial logit models [Zissimopoulos and Karoly (2007)], multinomial probit models [Theodossiou and Zangelidis (2009)] or duration models[Güell and Petrongolo (2007)]. In general, this literature focuses on some specific transitions, such as transitions in and out of unemployment, into retirement or between types of employment. It provides interesting methodological possibilities, which we will further analyse in our methodological section.

Methodology.

The choice of a dynamic microsimulation framework.

We have made two major methodological choices in order to carry out the retirement and social security projections of ANSES. The first is to opt for a microsimulation approach, starting from available individual data, instead of simulating a model at the macro level. Because this approach builds upon an existing population sample, we have chosen as our main data set the quarterly "Permanent Household Survey" (EPH), which is available at the micro level and covers the 2003-2013 period. This survey studies a randomised sample of the urban population of 31 Argentinian cities and metropolitan areas. These represent most of Argentina's urban population, and hence most of its total population, since about 91% of Argentinians live in towns of 2000 and more inhabitants[16]. Each household is surveyed four times all in all: two consecutive quarters first, then after a break of six months two other consecutive quarters. This allows for a rotation in the studied population while still making it possible to study quarterly as well as yearly transitions. An average of around 50000 individuals are interviewed each quarter.

The choice of a microsimulation framework follows a number of reasons, the main one being that "microsimulation models (...) differ from (semi-) aggregate budgetary models in that they simulate the impact of policy measures and schemes on real people, and not on averages or representative agents. (...) Where macro simulation considers averages, a micro simulation model attempts to take into account the heterogeneity behind the average."[Dekkers et al. (2012), p 4]. Since a number of conditions come into play when considering both benefits granted and the amount a given individual must contribute, it was only by going down to the individual level that we deemed possible to precisely study and project ANSES benefits and contributions. Moreover,while these projections will take into account aggregated expected evolutions such as growth rates, productivity gains or activity rates in their scenarios, they should also account for the effects of expected structural changes on Argentina's population due to ageing and to changes in the productive structure of the country. Given the

[16]Source: own calculus from INDEC, 2010 Argentinian Census (Censo Nacional de Población, Hogares y Viviendas, 2010).

significant economic changes the country has experienced since it defaulted on its debt in early 2002, it is not possible to assume that the labour career of generations about to retire today will be similar to younger generations' when they retire some decades from now.

The descriptive statistics on the EPH all show unemployment but also informal labour have sharply decreased since 2003, implying a rise in contributing workers as we can see in Table 1. Also, prior to the 1998-2002 crisis, there had been throughout the 1980's and 1990's a very unstable economic situation, with huge fluctuations in employment and formal employment that have not faced younger working generations. Wages have also risen sharply in the last ten years, although it is difficult to measure exactly how much they have since official inflation figures have been mired in controversy since 2007[17]. Given this, we can expect that cohorts that will retire in the coming years will be structurally different from cohorts that have recently retired, will less often benefit from moratorium retirement or family benefits but will on the other hand be entitled to higher benefits. It is hence necessary to take into account this change in structure in working-age population in our social security financing needs' projections, meaning for instance that it is highly likely that in the future most pension benefits will no longer be moratorium or minimum retirement benefits as it is the case today in Argentina.

Here is where the methodological advantages of this technique for our particular problem are relevant. As described by Li and O'Donoghue: " In order to evaluate certain impacts of public policies (...) it is necessary to utilise a long panel data set. In general, such data sets are not available, either because the analysis relates to the future, as in the case of pension forecasts, or because collected data sets do not cover sufficiently long time periods; therefore, analysts use dynamic microsimulation models to assist in their analysis (...). Essentially, microsimulation is a tool to generate synthetic micro-unit based data, which can then be used to answer many "what-if" questions that, otherwise, cannot be answered." [Li and O'Donoghue (2013), p 4]. This synthetic population can hence give detailed consequences on the adoption of different hypothesis for our scenarios, both demographic and macroeconomic. With this population, we can precisely gauge the structural changes that would derive from an ageing of present-day young cohorts.

One of the methodological challenges this method however entails is the handling of weights. Indeed, as is the case of the EPH dataset, most household surveys are weighted so that the interviewed population is representative of a larger partly surveyed population, whose overall characteristics (age group and gender structure for instance) are known from external sources. Without these, the dataset is simply no longer representative of this broader population that interests the modeler in the first place. The simplest solution from a

[17] This has resulted for some years in gaps of ten percentage points between the consumer price index and the GDP deflator, that before then did not exist with such a magnitude. The implementation, starting from February 2014, of a new National Price Index developed with the support and the monitoring of the IMF however solved this problem, returning significantly higher monthly inflation figures.

methodological point of view would be to expand the dataset: if an individual represents n individuals, then we clone it n times and assign a weight of 1 to each of these clones. However, larger datasets imply longer computation times. Although there is an incipient literature of alternative solutions to expanding the dataset[18], in this paper we have chosen to expand the dataset in a limited way. Instead of reaching a weight of 1 for each clone, we create a limited amount of clones with lower weights[19]. In the end we get an expanded dataset, where clones with lower weights behave independently from one another, but which does not require large computation capacity or time.

The software that has been chosen for this is LIAM2. This is an open-source and free microsimulation package recently developed by the Belgian Federal Planning Bureau. They state in their user manual that "the goal of the project is to let modelers concentrate on what is strictly specific to their model without having to worry about the technical details. This is achieved by providing a generic microsimulation toolbox which is not tied to a particular model. By making it available for free, our hope is to greatly reduce the development costs (in term of both time and money) of microsimulation models" [Bryon et al. (2013), 1]. This choice was hence driven for practical reasons.

The choice of empirical transition rates for modeling transition between labour-market states.

Among different available methods, we have decided to use the simplest and least theoretically demanding one of inputting transition matrices. As Li and O'Donoghue explain, "microsimulation models could use structural behavioural models, reduced form statistical model or simple transition matrix to simulate changes. Behavioural models are grounded in economic theory, in the sense that changes to institutional or market characteristics result in a change in the behaviour of agents within the model. In contrast, reduced form statistical models aim to model the transition probabilities of individual characteristics using related variables. (...) Reduced form models usually (...) assume a stable policy environment implicitly. (...) It is the easiest way to model potential changes with least theoretical considerations"[Li and O'Donoghue (2013), p 25]. We should keep in mind though that "[as] reduced form models and transition matrices (...) usually do not depend on policy parameters, they are not suitable for reform analysis, and are often restricted to simulating status quo only. " (Li, 2013, p. 26). The first version of our model will hence be with an unchanged legislative framework, since we will not for now include individual decision making concerning labour states, the most interesting of which could be retirement decision modeling or transitioning in and out of the informal workforce.

We have defined five distinct activity and inactivity states, in which all our working-age population can be categorised. These are contributing salary work-

[18]For instance,[Dekkers and Cumpston (2012)] propose a method to assign weights to newborns and immigrants.

[19]We warmly thank M. Dekkers, who provided us the code we used for expanding our dataset.

ers, contributing independent workers[20], undeclared workers, unemployed, and inactive people. Unfortunately, the EPH does not contain information about registered independent workers: indeed, the questions "for this work, do you have a retirement contribution" or "do you on your own contribute to any retirement system"[INDEC (2003), p 22] are only asked to salary workers. In order to avoid having all independent workers be counted as undeclared workers, we have used methodology from [Maurizio (2012)] to measure informality among independent workers. Indeed, based on "ILO's 15th and 17th International Conference of Labour Statisticians" recommendations, she assumes that "in the case of independent workers, only those with no professional skills are considered as part of the [informal sector], as an operational way to leave only independent workers with low productivity in this sector"[Maurizio (2012), p 6]. Furthermore, "Also on the ILO's recommendation, given the lack of enough information from household surveys, in the case of independent workers, their formal/informal character is directly determined by the characteristics of their enterprises: informal own-account and employers are those working in enterprises of the [Informal Sector]."[Maurizio (2012), p 7]. She applies this method to the EPH data set for Argentina and finds "4,4%" of the urban labour market in 2006-2007 was comprised of formal non-wage earners [Maurizio (2012), p 11], a figure close to the one we got for that period. We have hence defined independent contributing workers as those working in firms with five or less individuals of the private sector and whose job was not qualified or required only an operative qualification[21].

The computation of these transition matrices will be first made, for descriptive statistic's sake, across different sub-populations categorised by gender, age and education, as seen in [Duc et al. (2013)] concerning the TRAJECTOIRE microsimulation model. Although we have failed to reject the non-stationarity null hypothesis of the Phillips-Perron test for quarterly transition rates between labour market states between the years 2003 and 2013, after controlling for seasonality, we have reached the conclusion that we must however compute our transition probabilities using exclusively the second half of our database. Indeed, we got unrealistic labour-market states frequencies when simulating with the means of probabilities from the whole observed period. As will be seen in more depth in section 2.3, there are huge changes in the labour-market structure of our population in the 2003-2008 period. This entailed a structural mobility of the workforce out from unemployment and unregistered employment to registered dependent employment, which affected transition probabilities out from these states and into formal employment. We will thus exclude these temporarily higher transition probabilities from our prospective analysis, using the 2008-2013 as our benchmark in order to skim out this structural mobility from our transition probabilities. However, even when simulating within the 2003-2013 period with empirical quarterly transition probabilities, we get different simu-

[20]These pay comparatively less contributions than salary-men and women, but also receive less generous retirement pensions.

[21]As defined in Argentina's 2001 "Clasificador Nacional de Ocupaciones", or National Occupation Classifier.

lated population structures than the observed ones. There is hence a need to calibrate our transition probabilities so that the resulting simulated population has a similar structure than the surveyed population. We will present further in this paper our calibration methodology, as well as the resulting transition probabilities and simulated population, which we will compare to the surveyed one.

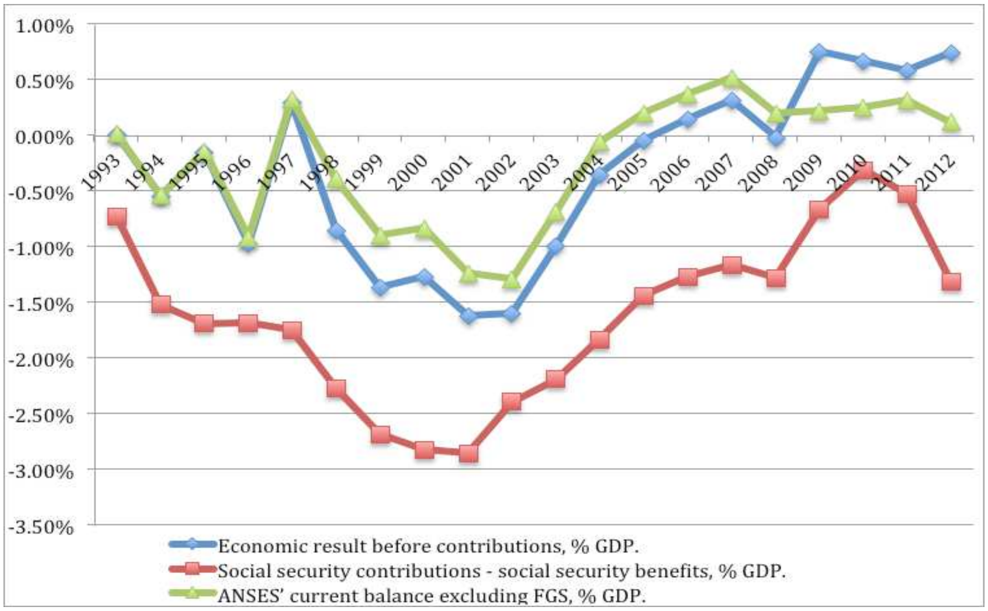
1 Transitions in Argentina's labour market: descriptive statistics and logistical regressions.

1.1 A brief description of Argentina's social security system.

As we have already said in the introduction, Argentina has a federal Pay-As-You-Go pension system, the SIPA, run by ANSES.

Its story is summarised in Figure 1, which expresses different measures of ANSES economic results. We have a purely social-security balance (that is social security contributions minus social security expenses), ANSES' economic result before contributions (that is, non-figurative revenues minus non-figurative expenses) and finally this same economic result excluding capital expenses and resources, mainly FGS revenues starting from 2008 (which are currently reinvested in the Fund, and as such do not finance ANSES current expenditures).

Figure 1: ANSES economic result in GDP points. [Calcagno (2013), p 21]



In the beginning ANSES was the 'poor parent' of a two-pillars system where the fully-funded pillar got most of the contributions and paid the least retirement benefits, with chronic deficits that substantially worsened when the economy entered in a recession in 1998. When the recession ended in late 2002, and a substantial economic recovery started only to be interrupted by the 2008 financial crisis, this deficit was gradually curbed, reaching sizeable surpluses of up to 0,5% of GDP in 2007. With the onset of the crisis in 2008 together with an acceleration of social security expenses[22] this surplus was reduced until it nearly reached the equilibrium in 2012, excluding FGS revenues. An important point is that ANSES revenues are not all explicitly allocated for the payment

[22] This stems from a combination of factors. The private pension funds nationalisation transferred to ANSES the payment of pensions from the fully-funded system, but mostly the indexation of pensions on wages and ANSES income variation in 2008, as well as the pension moratorium that came into place the same year, highly increased retirement expenses. This was however offset by the rise in contributions stemming from the end of the fully-funded system in 2008.

of a given benefit, since in any case total contributions to social security have never been enough to pay for social security benefits[23]. That is why it is not possible to exclusively study the sustainability of retirement pensions: it cannot be dissociated from the sustainability of the social security system as a whole.

We believe these figures summarise well the history of ANSES, and help understand better the motivation of the whole project of providing a meaningful projection on this agency's future current balance, given how this balance has strongly fluctuated over time. Now let's focus on the core of our paper: labour-market transition probabilities and the empirical bases for our future projections.

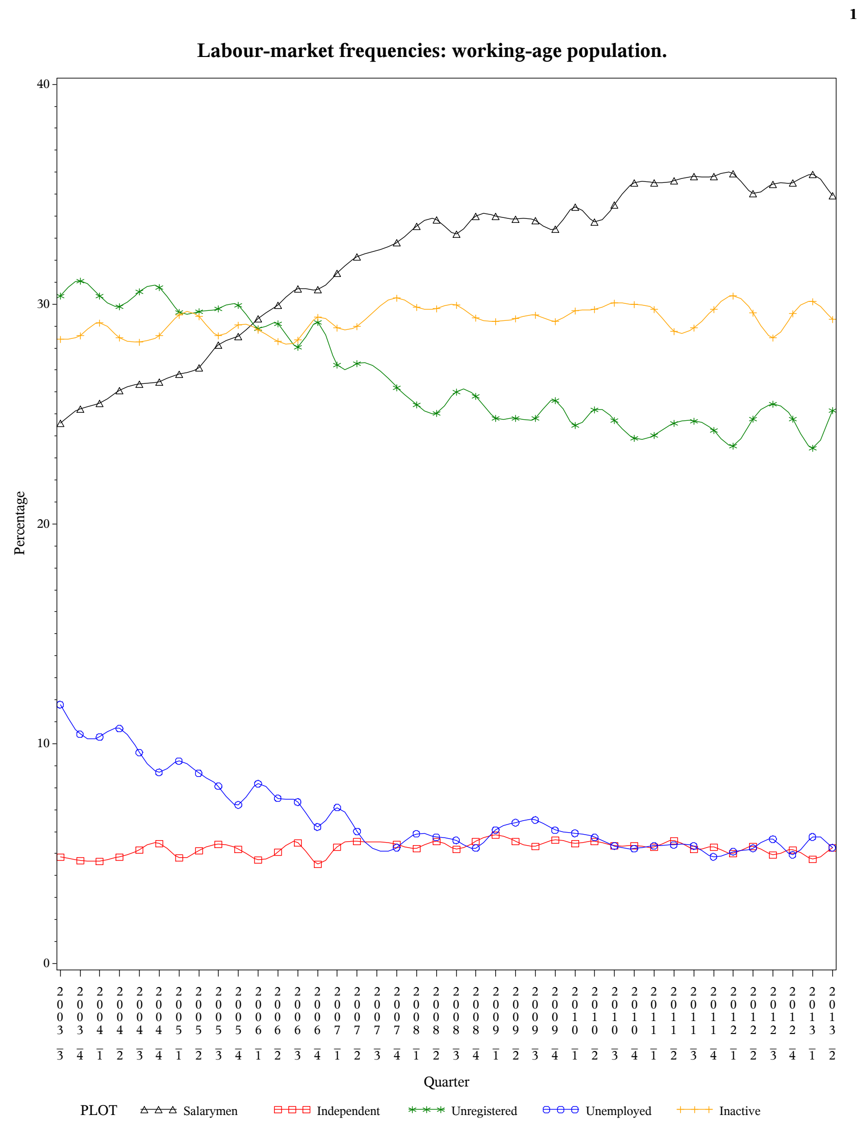
1.2 Empirical vs. calibrated transition probabilities: how do we reproduce the changes in the labour-market structure?

Following the guidelines we established in our methodological section, we first calculated the quarterly labour-market transition rates of the population of working age. A worker can earn a contributive retirement pension if he has reached retirement age (60 for women, 65 for men) and has contributed at least 30 years. He can get a better replacement ratio[24] by contributing more years, until he reaches 35 years of contributions. Since in order to validate a contribution year a worker has to contribute two trimesters in a year, we have chosen to retain a quarterly step in our simulations. After taking into account weightings, we have arrived to the conclusion that these transition rates were surprisingly stable over our 10 year period despite the considerable change in activity and inactivity rates of our working age population. Figure 1 shows the evolution in structure of Argentina's labour market, which has been significant in the 2003-2013 period.

[23]In 2011, 68% of ANSES revenues excluding government transfers came from social security contributions, 55% including these. Source: own calculus from the Savings-Investment-Financing Account of Social Security Organisms (Contaduría General de la Nación, CGN).

[24]That is, the pension / reference wage ratio.

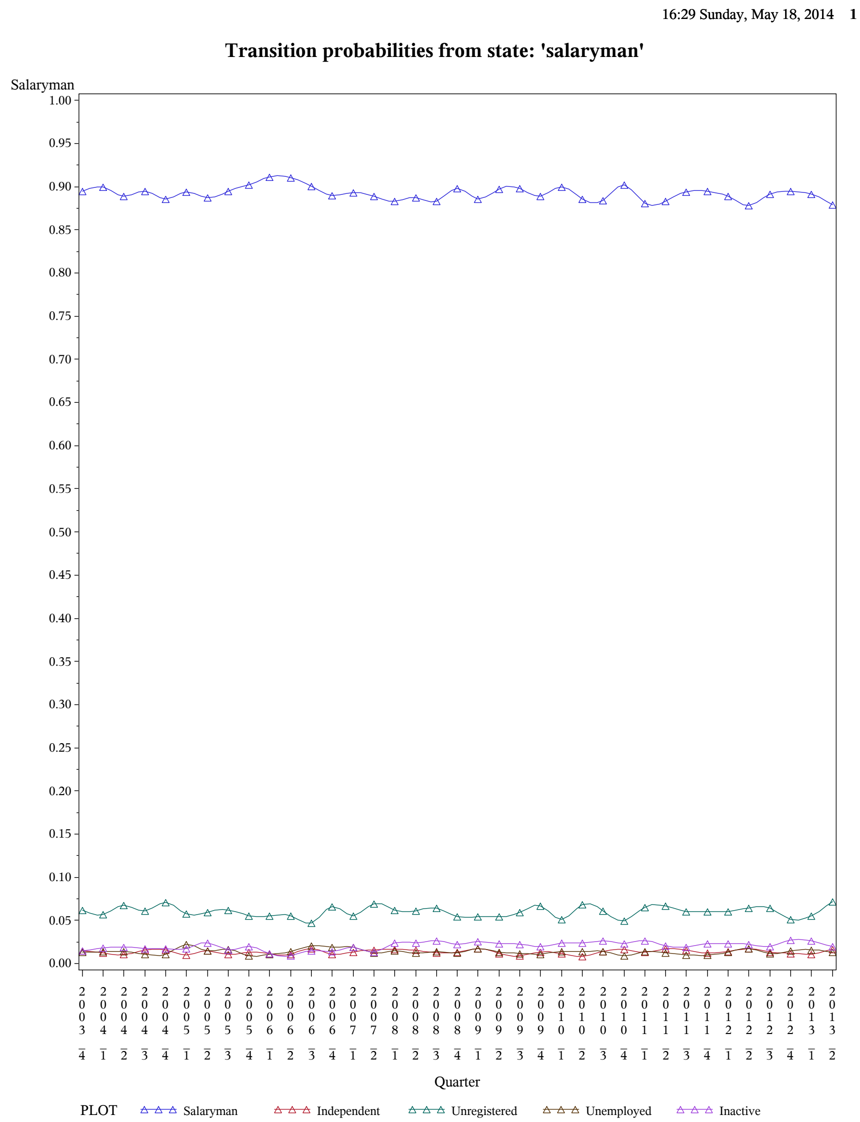
Figure 1: labour-market states as a percentage of working-age population. Argentina, 2003-2013. Source: EPH dataset



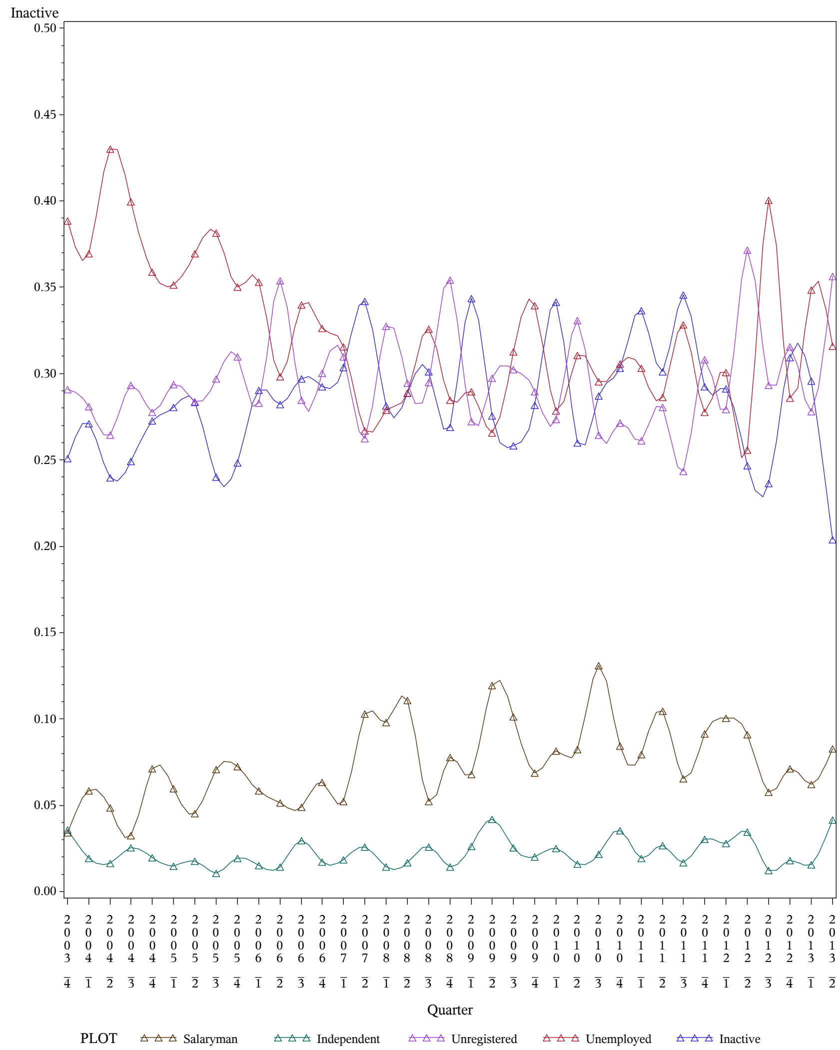
We can see that the inactivity rate remained stable at around 29% of working-age population. However, the structure of the active population changed, and Argentina witnessed a significant increase in contributors to social security. Indeed unemployed and unregistered workers as a percentage of working-age population fell respectively from 11,1%[25] to 5,2% and from 30,6% to 25,1% between 2003 and 2012. Hence, contributors to social security as a percentage of working-age population rose from 30% to 40%, and it increased from 42% to 57% as a percentage of active population in that decade. Despite this considerable change in structure, we can see that empirical transition rates between labour-market states have remained stable throughout the period, as Figure 2 shows.

[25]These figures are taken from Annex 1, which compares the labour market structure of Argentina in the last semester of 2003 and of 2012.

Figure 2: Quarterly transition probabilities from the states "salaryman" and "unemployed" for people of working age, 2003-2013.



Transition probabilities from state: 'unemployed'



We can here see that the transition probabilities are quite stable for all of the possible transitions when the individual was a contributing salary-man in the previous quarter, with seasonal fluctuations. The first probability is to stay a contributing salary-man, roughly around 90%, then it is to become an unregistered worker round 6%, with the remaining probabilities being quite similar. We have computed these probabilities for each sub-sample that resulted from combining gender, education and age group variables. Although some categories had few observations we ruled out the null hypothesis at the 5% threshold on the existence of a trend in these transitions through a Phillips Perron test, and after taking into account seasonality. The only cases where a slight downwards trend can be graphically seen, although not confirmed by a Phillips Perron test, was for transitions from unemployment: we see a decreasing probability of remaining unemployed between two consecutive quarters, and an increasing one of transitioning into inactivity or unregistered work, to the point that at the end of the period these are more or less the same, as can be seen in Figure 4. Nevertheless, the first simulations we ran with transition probabilities averaged across the whole period resulted in unrealistic labour-market states frequencies, with exceedingly low unemployment and unregistered work frequencies, precisely those categories that diminished the most at the beginning of the period. Hence, we found ourselves in a situation similar to [Duc et al. (2013)]'s, where "descriptive statistics by generation show us a quite important evolution in behaviour between generations". We will therefore "avoid using older generations for calibrating the model, since their behaviour does not necessarily mirror more recent generation's" [Duc et al. (2013), p 17], and probably use an average of transition probabilities from the second half of our database, where labour-market behaviour seems to have stabilised since the onset of the 2008 international financial crisis.

Nevertheless, averaging transition probabilities in the whole period still helped us see some characteristics in transitioning between labour-market states, as can be seen with Table 2's labour-market transition probabilities.

Table 2: average labour market transition probabilities by gender, age groups and education: from contributing employee to contributing employee.

Feuille1

Probability of staying a contributing employee, quarterly step

Transition probabilities, men	No high school	High school	University
16-19	0.59	0.65	0
20-24	0.76	0.82	0.78
25-29	0.84	0.89	0.92
30-34	0.89	0.92	0.93
35-39	0.9	0.94	0.93
40-44	0.92	0.93	0.94
45-49	0.9	0.94	0.93
50-55	0.9	0.91	0.92
55-59	0.9	0.9	0.9
60-64	0.89	0.9	0.86
Transition probabilities, women	No high school	High school	University
16-19	0.33	0.65	0
20-24	0.67	0.81	0.85
25-29	0.78	0.87	0.9
30-34	0.81	0.9	0.93
35-39	0.83	0.91	0.94
40-44	0.82	0.91	0.96
45-49	0.87	0.93	0.96
50-55	0.82	0.93	0.95
55-59	0.87	0.91	0.94

Table 2 shows that the higher the formation, the higher the probability to remain a contributing employee. A similar phenomenon was seen when comparing male to female individuals of working age. Finally, we witnessed some differences in levels across age groups. There is a non-linear impact of age on transition probabilities: here, for instance, age increases the probability of staying a contributing employee until ages 35-39 were it decreases this probability.

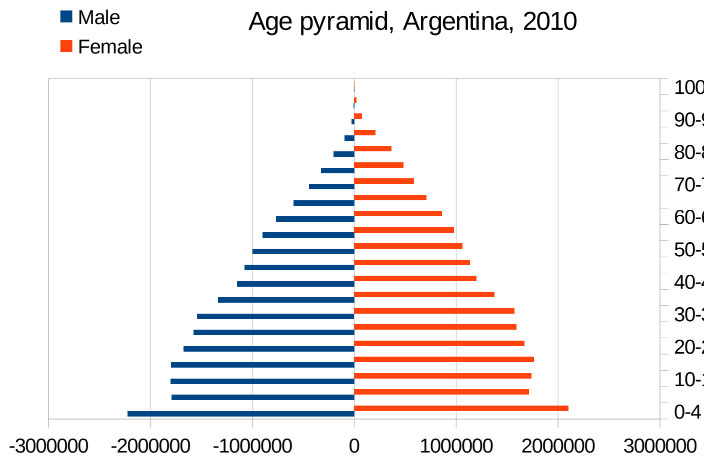
In order to remedy this stability of transition probabilities that does not correctly reproduce the structural changes Argentina's labour market has experienced in the 2003-2013 period, we have decided to recalculate them using the following alignment routine: using these empirical transition probabilities, we simulated each quarter of our dataset for only one period. Then, we changed the weights of each of these simulated quarters in order to make them reproduce the labour-market structure of the corresponding period in our dataset. For instance, we took the population interviewed in the third quarter of 2003, we simulated it for one period and then made it match the labour-market structure of the population interviewed in the fourth quarter of 2003. We then used these two datasets (the population interviewed in the third quarter of 2003 and this population simulated to the fourth quarter of 2003) to calculate new transition probabilities. This is still work in progress: as we will see later on in this article, this calibration technique helped reproduce better the labour market structure of our dataset, but has however failed to do so in a satisfying way. Nevertheless, the ultimate goal of this calibration is to obtain suitable transition probabilities for our prospective, as well as retrospective, simulations. The preparation of which we will describe in the second section of this paper.

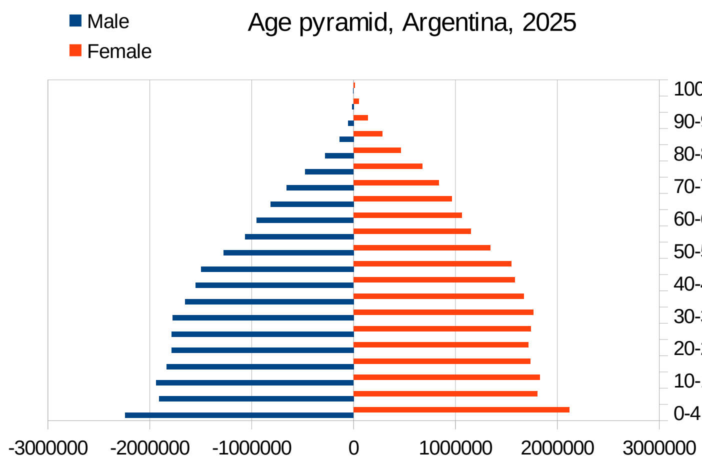
2 Laying the basis for a microsimulation analysis: computing demographic transitions and identifying and correcting discrepancies between the data and some early microsimulations.

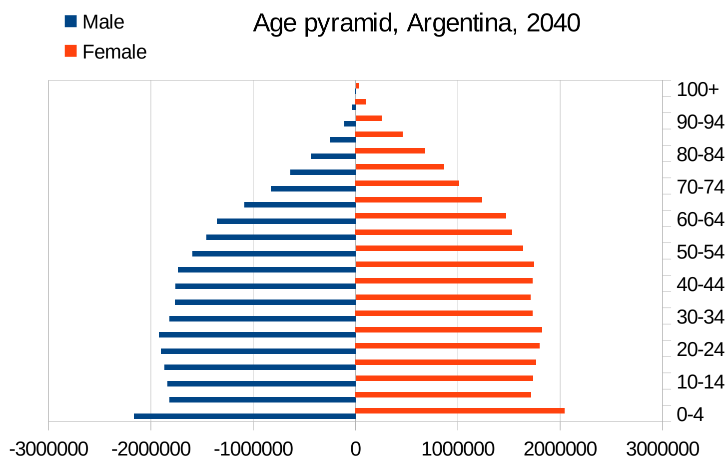
2.1 An ongoing demographic transition: INDEC's projections on Argentina's population.

The other key input for our future microsimulation are demographic transition rates. The most important ones are conditional probabilities of dying, depending on age and gender, and probabilities of giving birth conditional to the age of the mother. We computed the first ones based on demographic projections from INDEC until 2040, reconstituting year by year the death rates with deaths by age data that the INDEC kindly provided us on demand. The second ones were directly available at INDEC's website, although a possible sophistication could be making these probabilities vary depending on the existing number of children and the civil state of each woman. This will be our baseline demographic scenario, which will result in the following demographic structure of the simulated population, as can be seen in figure 3's demographic pyramids.

Figure 3: Demographic pyramids in 2010, 2025 and 2040.







The estimated population will have undergone a demographic transition, with a progressive drop of fertility and increasing longevity. For instance, The INDEC supposes that by 2040 the fertility rate will be slightly below the population replacement level of 2.1 children per woman. The crucial question that remains out of the scope of this work is whether we should expect it to stabilise at that value or continue falling to levels akin to those of Southern Europe. Our demographic scenarios will be derived from this baseline projection, and we expect them to have a huge impact on the estimated overall sustainability of the pension system, probably bigger than the macroeconomic scenarios we will later on develop.

2.2 Testing the methodology in our data: comparing backwards simulation, forward simulation and actual data in the 2003-2013 period.

As a result of these methodological considerations, and before simulating up to the 2040 horizon, we have run some simulations on the period covered by our database. On the one hand, backwards microsimulation is made using the last available survey, that is the 2013 second quarter wave, up until the beginning of our database in the third quarter of 2003, supposing there is no migration. Labour-market transitions are determined by reversed transition probabilities: if we know at time t a given individual is in a given state, these are the probability of having been at time t-1 in a given state. On the other hand, we simulate forward the first available surveyed population, that is the 2003 third quarter wave, up until the end of our database. We have computed life tables using data on population and death by age and gender issued by Argentina's Ministry of Health [26], obtaining conditional probabilities of death for the 2003-2009 period,

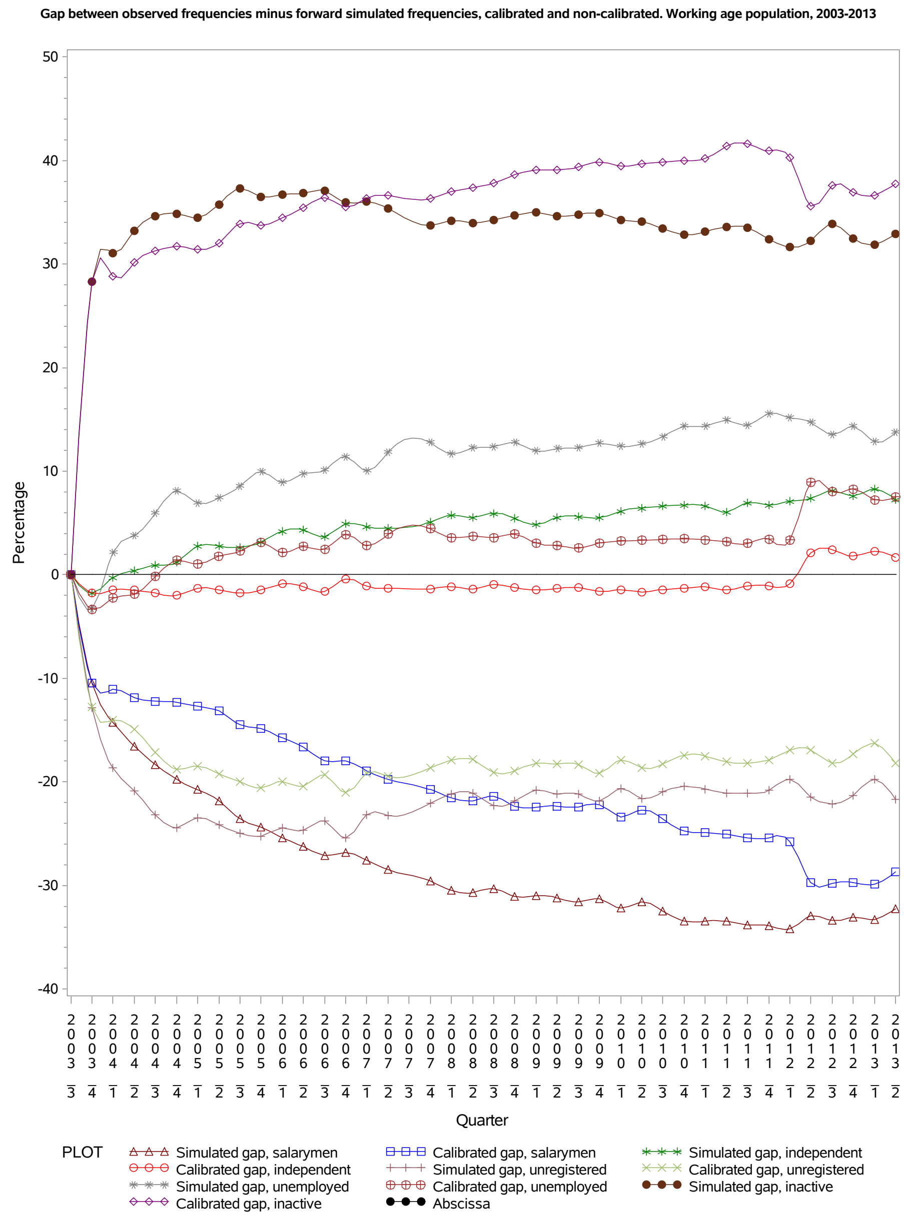
[26] The data is compiled in yearly reports issued by the "Dirección de Estadísticas e Información de Salud" in their "Anuario de Estadísticas vitales" series. These can be downloaded in

which we complete with the ones we computed for the 2010-2040 projection. From these reports we have also obtained conditional probabilities of giving birth by age.

The objective of this is, by fitting our simulations to our available data, to calibrate our model and obtain transition probabilities which reproduce the observed labour-market structure. It is these transition probabilities we intend to use later on to make simulations that exceed the span of our dataset, both in the past and in the future. After having observed that empirical transition probabilities failed to properly reproduce our data, we applied a calibration strategy, which we described earlier in our methodology section. However, although it helped tighten the gap between observed and simulated frequencies, we still have failed to calibrate these to our dataset. In Figure 4, we have graphed the gap between observed frequencies and forward simulated frequencies, both calibrated and non calibrated for population of working age. In Figure 5 we have done the same thing for backward simulation.

the following URL: http://www.bvs.org.ar/php/level.php?lang=es&component=31&item=19 .

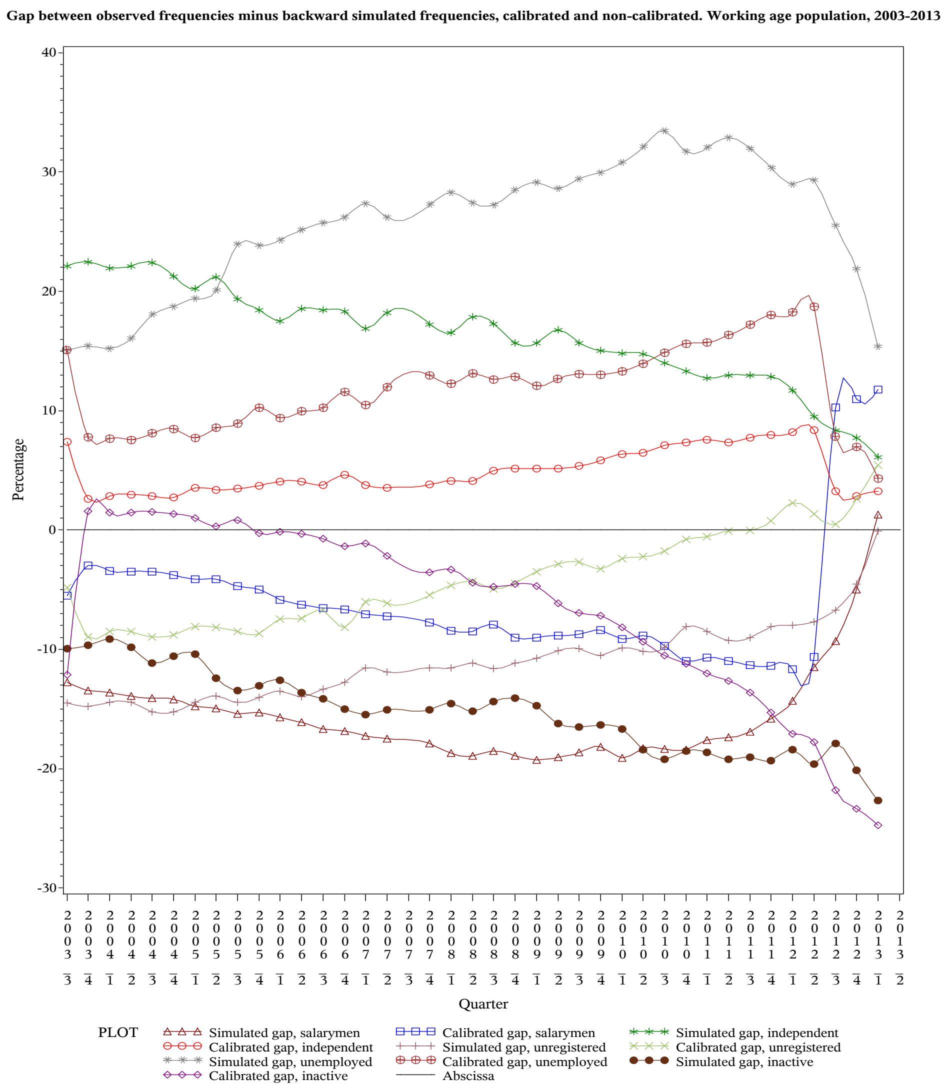
Figure 4: Forward simulation from 2003 compared to 2003-2013 data.



The differences between the simulated frequencies and the actual frequencies are sometimes huge. Both calibrated and non calibrated simulations overestimate inactivity in working-age population, by 30 to 40 percent points throughout the period. Calibration is more efficient for other categories though: for instance, there is a 10 to 15 percent points overestimation of unemployment in the uncalibrated simulation, whereas this gap falls to around 5 percentage points after calibration. Finally, there is a significant underestimation of unregistered workers (a stable 15 to 25 percent points throughout the period) and an increasing underestimation of registered dependent workers (reaching almost 30 percentage points by 2013). We do not understand precisely why we get this phenomenon, nor why the gap got wider when we started expanding our input dataset for taking weights into account, but it is still work in progress and we are working on adjustment techniques to calibrate our prospective transition probabilities for our future simulations up to 2040.

Regarding backwards microsimulation, crucial for reconstituting a contributing record of each individual in our database compute their pension, we get a similar problem, as pictured in Figure 5.

Figure 5: Backward simulation from 2013 compared to 2003-2013 data.



The behaviour of calibrated and non-calibrated gaps differ significantly from our forward projection. Indeed, for our non-calibrated simulation, is very low at the first simulated period (first quarter of 2013) but afterwards increases substantially. For instance, for simulated salary-men without calibration, the gap begins slightly above zero before quickly reaching -20 percentage points. The behaviour of the gap for calibrated salary-men is the opposite: it begins with huge gap (including a jump between the second and the third quarter of

2012, which we do not yet understand) but then this gap steadily diminishes, reaching -5 percentage points in the third quarter of 2003.

Hence, we can see from these figures that our calibration method is still not good enough to fit the data. Our current objective is to enhance these techniques and achieve a satisfactory adjustment routine which would produce narrower gaps between our simulations and our data. We do not understand either why these gaps have widened, both for our forward and our backward simulation, when we started working with expanded datasets to take into account weights. Nevertheless, although we are currently working on improving our calibration methods before simulating further in time and starting to get meaningful prospective simulations of Argentina's labour market and pension system, all in all, we have been able to make our first simulations and set the foundations for a dynamic microsimulation model of Argentina's social security system future income, expenses and hence possible financing needs.

Conclusion

All in all, in the current state of our work, we have obtained all the input necessary for carrying out a basic forward microsimulation in studying the social security system of an emerging country, Argentina. We have estimated empirical transition rates that depend on age, gender and formation, and together with conditional death and birth rates, we were able to run simple simulations on labour-market history. We then calibrated these empirical transition rates with our dataset in the 2003-2013 period and achieved to better reproduce the labour-market structure of our dataset. We still need to hone our calibration methods to better fit our data, although a partial solution could be to only use for our forward simulations transition rates from 2008 onwards, where Argentina's labour-market structure has stabilised as can be seen in Annex 1. Once this will be done, we will be able to make a forward simulation of our surveyed population up to 2040, as well as a backward simulation to reconstitute past contributed years, and hence have a synthetic future generation where we will estimate replacement ratios at retirement.

Future developments will include wage functions to simulate future and past wages in order to calculate the reference wage and hence determine the retirement pension each individual is entitled to at retirement. It will also compute total contributions by individual, helping us simulate total contributions to social security. We may also derive labour market transition rates from multinomial logistical regressions, to further hone the studied transition rates and take into account more variables in transitioning between labour market states. On the other hand, we intend to carry out a civil state modelisation, with matching functions for simulating marriage and divorce, which should help us simulate the beneficiaries of non-retirement benefits, such as survivors pensions, but also contributive and non-contributive family benefits. These objectives are bound to take time to achieve, but we are, nevertheless, quite confident we will ultimately achieve a comprehensive microsimulation of social security expenses and contributions, producing reliable forecasts for Argentina's social security system evolution and future financing needs.

Annex 1:Labour-market states in working-age population in the last semester of 2003 versus the last semester of 2012. Source, EPH data set.

Table of Gender by Workstate: working-age population, 2003

Gender	Workstate					
Percent Row Pct Col Pct	Unregistered worker	Unemployed worker	Inactive	Independent worker	Salaryman	Total
Men	17.23 35.25 56.27	5.53 11.32 50.10	8.19 16.75 28.50	3.13 6.40 65.61	14.81 30.29 59.62	48.89
Women	13.39 26.20 43.73	5.51 10.78 49.90	20.54 40.18 71.50	1.64 3.21 34.39	10.03 19.63 40.38	51.11
Total	30.62	11.05	28.72	4.77	24.84	100.00

Table of Gender by Workstate: working-age population, 2012

Gender	Workstate					
Percent Row Pct Col Pct	Unregistered worker	Unemployed worker	Inactive	Independent worker	Salaryman	Total
Men	14.90 29.94 59.38	2.42 4.87 45.94	8.36 16.80 28.59	3.25 6.53 64.31	20.82 41.85 58.89	49.75
Women	10.19 20.27 40.62	2.85 5.67 54.06	20.88 41.54 71.41	1.80 3.59 35.69	14.53 28.92 41.11	50.25
Total	25.08	5.27	29.23	5.05	35.35	100.00

References

[Albert et al. (2009)] Albert, C., Berteau-Rapin, C., & Di Porto, A. (2009). PRISME: le modèle de microsimulation dynamique de la CNAV, régime français de pension de retraite. CNAV. URL: http://www.statistiques-recherches.cnav.fr/projections-a-moyen-long-terme.html

[Ando et al. (2000)] Ando, A., Brandolini, A., Bruno, G., Cannari, L., Cipollone, P., D'Alessio, G., ... Altimari, N. (2000). The Bank of Italy's DYNAMITE: Recent Developments. Roma: Banca d'Italia.

[Blanchet (2009)] Blanchet, D. (2009). The Destinie 2 microsimulation model: increased flexibility and adaptation to users' needs. Présenté à IMA Conference, Ottawa, Canada.

[Blanchet et al. (2011)] Blanchet, D., Buffeteau, S., Crenner, E., & Le Minez, S. (2011). Le modèle de microsimulation Destinie 2: principales caractéristiques et premiers résultats. Économie et Statistique, (441), 101 121.

[Bonnet and Mahieu (2000)] Bonnet, C., & Mahieu, R. (2000). Public pensions in a dynamic microanalytic framework: the case of France. In Microsimulation in the new millennium: challenges and innovations. Cambridge: Cambridge University Press.

[Bradley et al. (2003)] Bradley, S., Crouchley, R., & Oskrochi, R. (2003). Social exclusion and labour market transitions: a multi-state multi-spell analysis using the BHPS. Labour Economics, (10), 659 – 679.

[Bryon et al. (2013)] Bryon, G., Dekkers, G., & de Menten, G. (2013). LIAM 2 User Guide. Release 0.70. Federal Planning Bureau, Economic analysis and forecasts.

[Calcagno (2013)] Calcagno, L. E. (2013). Argentina's pension system: horizons to 2050. Memoir, University of Paris I, Paris.

[Dekkers and Belloni (2009)] Dekkers, G., & Belloni, M. (2009). Micro simulation, pension adequacy and the dynamic model MIDAS: an introduction. Project AIM -Deliverable 4.10.

[Dekkers et al. (2010)] Dekkers, G., Buslei, H., Cozzolino, M., Desmet, R., Geyer, J., Hoffmann, D., Tedeschi, S. (2010). What are the consequences of the European AWG-projections on the adequacy of pensions? An application of the dynamic micro simulation model MIDAS for Belgium, Germany and Italy. In Life-Cycle Microsimulaion Modelling. Constructing and using Dynamic Microsimulation Models (p. 325 – 357).

[Dekkers and Cumpston (2012)] Dekkers, G., & Cumpston, R. (2012). On weights in dynamic-ageing microsimulation models. International Journal of microsimulation, (5), 59 – 65.

[Dekkers et al. (2012)] Dekkers, G., Inagaki, S., & Desmet, R. (2012). Dynamic Microsimulation Modeling for policy Support: An Application to Belgium and Possibilities for Japan. The review of socionetwork strategies, 31 – 47.

[DNPE (2012)] DNPE. (2012). Boletín Estadístico de la Seguridad Social, Cuarto Trimestre 2011, Primer Trimestre 2012. Ministry of Labour and Social Security, Argentina, Dirección Nacional de Programación Económica. Buenos Aires: Ministry of Labour and Social Security, Argentina.

[Duc et al. (2013)] Duc, C., Leuquien, L., Housset, F., & Plouhinec, C. (2013). Le modèle de microsimulation TRAJECTOIRE. Trajectoire de carrières tous régimes. (Document de travail No. 40).

[Falkingham and Johnson (1993)] Falkingham, J., & Johnson, P. (1993). A unified funded pension scheme (UFPS) for Britain. Welfare State Programme, Suntory-Toyota International Centre for Economics and Related Disciplines, London School of Economics.

[Flood (2007)] Flood, L. (2007). Can we afford the future? An evaluation of the new Swedish pension system. In A. Harding & A. Gupta, Modelling our future: population ageing, social security and taxation (p. 33). Amsterdam: Elsevier.

[Galler and Wagner (1986)] Galler, H. P., & Wagner, G. G. (1986). The Microsimulation Model of the Sfb 3 for the Analysis of Economic and Social Policy. In G. H. Orcutt, J. Merz, & H. Quinke (Éd.), Microanalytic Simulation Models to Support Social and Financial Policy (p. 227 – 247). Amsterdam.

[Gomes (2012)] Gomes, P. (2012). Labour market flows: Facts from the United Kingom. Labour Economics, (19), 165 – 175.

[Güell and Petrongolo (2007)] Guell, M., & Petrongolo, B. (2007). How binding are legal limits? Transitions from temporary to permanent work in Spain. Labour Economics, (14), 153 – 183.

[Harding (2007)] Harding, A. (2007). APPSIM: The Australian dynamic population and policy microsimulation model. Présenté à IMA Conference, Vienna, Austria.

[INDEC (2003)] INDEC. (2003). Diseño de Registro y Estructura para las bases de microdatos. Individual y hogar. Ministerio de Economía y Finanzas, Argentina.

[Joyeux et al. (1996)] Joyeux, C., Plasman, R., & Scholtus, B. (1996). A Model of the Evolution of Pensions Expenditures in Belgium. Présenté à Meeting of the European HCM Network on Socio-Economic Modelling, Cambridge.

[Klevmarken (2010)] Klevmarken, N. A. (2010). Microsimulation for public policy: Experiences from the Swedish model SESIM (Discussion Paper No. 242). Cabinet Office, Tokyo, Japan: Economics and Social Research Institute.

[Klevmarken and Lindgren (2008)] Klevmarken, N. A., & Lindgren, B. (2008). Simulating an ageing population: a microsimulation approach applied to Sweden (Contributions to Economic Analysis No. 285). Bingley, UK.

[Kennell and Sheils (1990)] Kennell, D. L., & Sheils, J. F. (1990). PRISM: Dynamic simulation of pension and retirement income. In Microsimulation Techniques for Tax and Transfer Analysis. (p. 137 – 172). Washington DC: The Urban Institute Press.

[Li and O'Donoghue (2013)] Li, J. O., & O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure and methodology. International Journal of microsimulation, 3 – 55.

[Maurizio (2012)] Maurizio, R. (2012). Labour informality in Latin America: the case of Argentina, Chile, Brazil and Peru (Working paper No. 165). The Brooks World Poverty Institute.

[Mesa Lago (2014)] Mesa Lago, C. (2014). Reversing pension privatization: the experience of Argentina, Bolivia, Chile and Hungary. (No. 44). Geneva: International Labour Office, Social Protection Department.

[O'Donoghue et al. (2009)] O'Donoghue, C., Lennon, J., & Hynes, S. (2009). The Life-cycle Income Analysis Model (LIAM): a study of a flexible dynamic microsimulation modeling computing framework. International Journal of microsimulation, (2), 16 – 31.

[Orcutt (1957)] Orcutt, G. H. (1957). A new type of socio-economic system. The review of Economics and Statistics, (39), 116 – 123.

[Orcutt et al. (1961)] Orcutt, G. H., Greenberger, M., Korbel, J. J., & Rivlin, A. M. (1961). Microanalysis of socioeconomic systems: a simulation study. New York: Harper and Brothers.

[Peichl et al. (2010)] Peichl, A., Schneider, H., & Siegloch, S. (2010). Documentation IZAΨMOD: the IZA POlicy SImulation MODel ., IZA Discussion Paper(4865).

[Poubelle et al. (2006)] Poubelle, V., Albert, C., Beurnier, P., Couhin, J., & Grave, N. (2006). Prisme, le modèle de la Cnav. Retraite et société, (2), 202 – 215.

[Silva and Vázquez-Grenno (2013)] Silva, J. I., & Vazquez-Grenno, J. (2013). The ins and outs of unemployment in a two-tier labor market. Labour Economics, (24), 161 – 169.

[Sutherland (2007)] Sutherland, H. (2007). EUROMOD: the tax-benefit microsimulation model for the European Union. In A. Harding & A. Gupta, Modelling our future: population ageing, health and aged care. (p. 483 – 488). Amsterdam: Elsevier.

[Theodossiou and Zangelidis (2009)] Theodossiou, I., & Zangelidis, A. (2009). Should I stay or should I go? The effect of gender, education and unemployment on labour market transitions. Labour Economics, (16), 566 – 577.

[Van Sonsbeek (2010)] Van Sonsbeek, J. M. (2010). Micro simulations on the effects of ageing-related policy measures. Economic Modelling, (27), 968 – 979.

[Zissimopoulos and Karoly (2007)] Zissimopoulos, J. M., & Karoly, L. A. (2007). Transitions to self-employment at older ages: The role of wealth, health, health insurance and other factors. Labour Economics, (14), 269 – 295.