CAPP_DYN: A Dynamic Microsimulation Model for the Italian Social Security System

Carlo Mazzaferro, Marcello Morciano

CAPP*aper* n. 48

settembre 2008

Dipartimento di Economia Politica – Università di Modena e Reggio Emilia –
www.capp.unimo.it

CAPP_DYN: A Dynamic Microsimulation Model for the Italian Social Security System

Carlo Mazzaferro*

Marcello Morciano**

September 2008

Abstract:

We present the technical structure of CAPP_DYN, a population-based dynamic microsimulation model for the analysis of the long-term redistributive effects of social policies, developed by the CAPP (Centro di Analisi delle Politiche Pubbliche) in order to study the intergenerational and intragenerational redistributive effects of social security system reforms. The model simulates the probabilistic socio-demographic and economic evolution of a representative sample of the Italian population for the period 2005-2050. After a brief review of similar existing models for the Italian economy, we offer a more detailed analysis and discussion of the model's functioning, together with a description of the estimation procedures employed in each of the model's individual modules.

JEL Classification: C51, C52, H55

Keywords: Dynamic microsimulation, lifetime and intragenerational redistribution, social security systems

* Department of Economics, University of Bologna and Capp

** University of East Anglia and Capp

1. Introduction[1]

Despite the fact that dynamic micro simulation models (MSM) have been employed for some time now in various developed countries as instruments with which to evaluate the long-term distributional effects of publicpolicies (O'Donoghue, 2001; Zaidi and Rake, 2002; Klevmarken, 2005), in Italy their use is much more recent, and as such has not been completely perfected. All the same, Italy has a complex, far-reaching welfare state, in particular with regard to pensions, which is currently faced with the rapid demographic aging of the population. Nevertheless, very little research has been conducted to date into the long-term redistributive effects of the reform of the social security system or of the said ageing process.

The first dynamic MSM for the Italian economy, DYNAMITE (Ando and Nicoletti Altimari, 2004) was developed towards the end of the 1990s within the context of a Bank of Italy research project. It was mainly employed to analyze the effects of demographic transition and social security reforms on private savings. Following this work, Vagliasindi (2004) developed MINT, a dynamic population MSM that analyses the medium/long-term distributional effects of pension systems, and the medium-term redistributive impact of changes in the personal income tax system. Neither of these models are currently in use, and to our knowledge, no other dynamic MSMs capable of evaluating the long-term distributive impact of public policy exist in Italy at present.

Like the aforementioned models, CAPP_DYN aims to provide a detailed description of the socio-demographic structure of the Italian population, together with a micro analysis of the evolution in the supply side of the labour market, in the labour income structure and in pension-related choices.

CAPP_DYN evolved from a research project carried out by the CAPP (*Centro di Analisi delle Politiche Pubbliche*), under the auspices of the Italian Department of Employment and Social Policies, designed to assess the distributional effects of those social security reforms implemented during the previous decade (Ministero del Lavoro e delle Politiche Sociali, 2005). The model was subsequently improved and developed further (Mazzaferro and Morciano, 2005; Morciano, 2007; Ministero della Solidarietà Sociale, 2008).

The CAPP_DYN model permits the simulation of the socio-demographic and economic evolution of a representative sample of the Italian population over the period 2005-2050. The base year population (2005) is derived from the 2002 wave of the Bank of Italy's Survey of Households Income and Wealth (SHIW).

The sample is re-weighted in order to align socio-demographic distributions with the Italian population. The dynamic aging of micro-characteristics is probabilistic, and is carried out by means

[1] We wish to thank Simone Tedeschi for technical assistance.

of finite, discrete Markovian processes. Certain behavioural functions have been introduced, the main one being that governing retirement choices.

Once the population structure has been defined, and labour incomes have been generated, the model simulates the main social security benefits in considerable institutional detail, according to the pension scheme provisions in force. The model can then estimate the distributional effects of key social security components, as well as the impact of social security reforms, allowing for the implementation of both cross-sectional (at different point of time) and inter- temporal life-cycle (of individuals living during different periods) analyses. A module that estimates the number of disabled people has been recently embedded in the model, thus allowing the projection, over the whole period, of the number of non self-sufficient individuals and the related long term healthcare expenditure.

An alignment process links CAPP_DYN to the official demographic forecasts provided by ISTAT, and the model is calibrated in order to follow GDP and wage trends in keeping with changes in the number of employed individuals.

As with other MSM models, CAPP_DYN shares certain advantages and drawbacks of this particular technique. For example, it allows for a detailed redistributive analysis of the social security system, which may be conducted from both a cross-sectional and an inter-temporal perspective. On the other hand, given that it is based on a population derived from a survey, considerable care must be taken during the initial selection sample process, with regard to sample representativeness and the difficulty of extracting the effects of un-observables from the data in question.

Moreover, it is important to remember that CAPP_DYN does not simulate the supply side of the Economy, and therefore alignment requires special attention in order to guarantee consistency with external demographic and economic forecasts.

2. General Features

According to the taxonomy offered by O'Donoghue (2001), the CAPP_DYN model presents the following features:

• It is a closed model: it simulates the life-cycle evolution of a population's main demographic and economic features. There are new members of the population each year, due to births and net migratory inflows, whilst others exit due to death.

• It is a dynamic ageing model: individual characteristics are periodically updated due to dynamic ageing processes based on discrete stochastic transitions between one state and another.

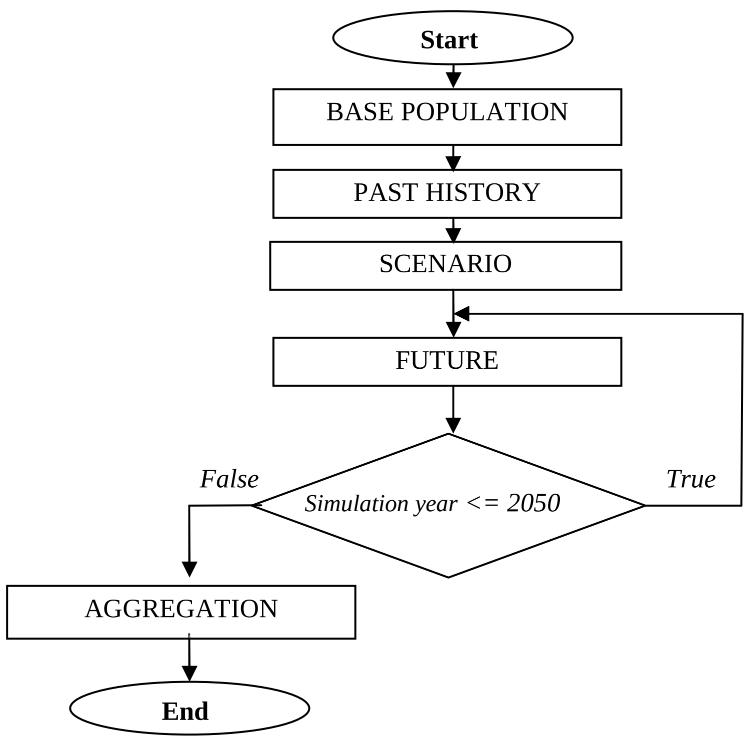
• It is a discrete time model: transition and updating processes are carried out at the end of each year.

• The ageing process is probabilistic: given a particular event subdivided into a number of mutually exclusive states at each point in time, transitions between states are achieved using probabilistic methods, and in particular, using a *Monte Carlo* technique.

• The units of analysis are both individuals and households.

More specifically, the model is structured in four blocks as shown in figure 1 below.

Figure 1 CAPP_DYN structure



• Base population: this block contains the procedures needed to generate the base year population. Socio-economic information for the basic units are drawn from the 2002 Survey of Households' Income and Wealth (SHIW). A set of statistical methods is then employed in order to improve sample representativeness.

• Past history: this retrospectively reconstructs the working path and earnings for basic units with a history of contributions going back to the base year.

- Scenario: this defines the model's exogenous parameters. In particular, it depicts the dynamic path of macro-demographic (mortality, fertility and migration) and macroeconomic (GDP and earnings growth) variables. Policy parameters and certain behavioural rules – in particular pension-related decisions - are also established within this section of the model.
- Future: this constitutes the main section of the model. This block contains a set of modules which simulate the socio-economic evolution of the micro-units according to observed individual characteristics. More specifically, the model recursively applies the modules and sub-modules reported in table 1 below. Each module in the "Future" block produces a yearly cross-section of outputs for the period from 2005 to 2050.
- Aggregation: this is the last step of the simulation process. The set of cross-sections of annual outputs is aggregated in order to produce a panel containing socio-economic information for the population for the period 2005-2050.

The following is a detailed description of the contents of each block.

Table1

"Future" block modules

	EVENTS	POTENTIAL CANDIDATES
	DEMOGRAPHIC MODEL	
1	Ageing	All individuals
2	Mortality	All individuals
3	Fertility	Married Women aged 16-49
4	Migration	Additional individuals aged 16-65
5	Leaving the family household	Children aged 18-34
6	Marriage	Single, divorced or widowed persons aged 16-60
7	Divorce	Married persons aged below 50
	HEALTH MODULE	
8	Disability	All individuals
	EDUCATION, LABOUR MARKET MODEL	
9	Compulsary schooling	Individuals aged below 16
10	Choice of post-compulsory educational level	Individuals aged 16 who have completed their compulsory education
11	Tertiary education	Individuals enrolled in tertiary education
12	Entry into the labour market	Individuals leaving or abandoning school
13	Transitions between labour and non-labour statuses	All individuals except pensioners and students
14	Transitions between contractual types	All individuals active in the labour market
15	Wages and Salaries	All individuals active in the labour market
	SOCIAL SECURITY MODULE	
16	Retirement	All non-pensioners accruing retirement pension requirements
17	Survivors' pension entitlements	Survivors (spouse and children) satisfying legal requirements
18	Social Pension entitlements	Individuals aged 65 and over entitled to care benefits
19	Pension benefits	All pensioners (old-age and retirement) in the three systems (defined as a benefit, defined as a contribution and mixed)
20	Supplements to minimum pensions (integrazioni al minimo) and social welfare supplements (maggiorazioni sociali)	Pensioners satisfying age and economic status requirements

3. Blocks description

• The base year population

The SHIW_02 is the most frequently used Italian data base for micro-econometric and distributional analyses. The survey unit is the household, i.e. a "group of individuals linked by ties of blood, marriage or affection, sharing the same dwelling and pooling all or part of their incomes" (Brandolini, 1999); however, as information is gathered at the individual level (only interest, dividends and financial assets are recorded at the family level), analyses of personal income are also permitted. SHIW income is net of tax and social security contributions.

It represents the Italian population, and sampling is organized in two stages: firstly, municipalities are non-randomly selected according to 51 strata; subsequently, households are randomly selected from within the strata. Hence, statistical inference must allow for sampling design: to this end, a bootstrapping method is employed.

The 2002 wave contains information on 21,148 individuals within 8,011 household units.

As in other surveys, differing response rates among groups, under-reporting and misreporting (especially with regard to capital income) are likely to bias any estimation based on this source. In particular, under-reporting seem significantly widespread among the self-employed (to a level of almost 20% in 1987, according to Cannari and D'Alessio's (1992) estimates), and is inversely related to household income and wealth, thus causing an underestimation of mean income and inequality[2]. Furthermore, a comparison with National Accounts data (through a grossing-up procedure) shows a slight overestimation of wages, together with a substantial underestimation of self-employed income and net interest on financial assets (the said figures are underestimated by 50% and 65-70% respectively): the end result is that total income is underestimated by about 30% (32% after interest and dividends have been included - Brandolini, 1999).

Finally, top and bottom coding problems have to be accounted for when the top or the bottom of the income distribution are analyzed.

Therefore, in building the base year population, we have tried to reduce, as far as possible, the biases resulting from the use of a not fully representative data set. To this end, we applied a post-stratification procedure to the original sample weights, using information provided be the most recent ISTAT census of households and the population as a whole. This procedure, which was developed by Gomulka and is currently employed by EUROMOD (Atkinson *et al.*, 1988), allows

[2] Response rate seems declining sharply from 26% of poorest to 14% of richest (Cannari, D'Alessio, 1992).

for the increasing representativeness of the sample in terms of the set of socio-economic characteristic that we are controlling[3].

As far as regards the size of the initial population, there is a trade off between, on the one hand, improving the heterogeneity of the simulation by using a larger sample, thus reducing estimation variance (Orcutt *et al.*, 1986), and on the other hand, the technological constraints involved in processing a set of sample members which, by the end of the simulation period, can number several million.

The substantial experience of several important research groups in the micro simulation area, reveals that in the present setting, the model simulates the evolution of a base year population of 107,000 household units and 270,000 individual observations.

- **The 'historical' block**

In order to obtain a complete contributory history for each micro unit present in the base year sample, the historical module retrospectively constructs the past working history of each active individual present in the base year[4].

The life-cycle profile of past earnings is built using econometric estimations implemented within the "income" module. Individual earnings are then discounted by an annual variable rate amounting to the growth of real earnings in the period 1952-2001[5].

- **The 'scenario' block**

This block enables us to set the values of the exogenous parameters. Table 2 displays the list of exogenous variables, together with the official data sources from which the values used in the simulations are drawn.

It should be pointed out that demographic dynamics and macroeconomic variables are not independent. Therefore, at this stage, the model uses the central demographic forecasts provided by ISTAT, which are the same employed by the *Ragioneria Generale dello Stato* (RGS) to forecast future GDP growth and earnings, which in turn represent the benchmark macroeconomic scenario.

Furthermore, retirement decision rules are set within these sub-blocks ; these rules account both for The inter-temporal choice optimizing framework, and for those elements linking retirement decisions to the achievement of a certain replacement rate level (i.e. the last gross earnings to first pension benefit ratio).

[3] A detailed report of the procedure can be found in Morciano (2007).

[4] The re-construction of active individuals in 2002 employs information regarding contributory seniority, professional attainments and sectors (actual and previous) taken from the SHIW_02 survey.

[5] Values are taken from Golinelli 2002.

Table 2 Data sources and reference scenarios for the exogenous variables

EXOGENOUS VARIABLES	*SOURCE*	*REFERENCE SCENARIOS*
	Demographic Variables	
Age, gender and geographical area specific mortality rates	ISTAT 01/01/2007	High, median, low
Age, gender and geographical area specific fertility rates	ISTAT 01/01/2007	High, median, low
Net Migration	ISTAT 01/01/2007	High, median, low
	Macroeconomic Variables	
Real GDP growth	Ragioneria Generale dello Stato 2007	Country base and programmed
Productivity growth	Ragioneria Generale dello Stato 2007	Country base and programmed

• The 'future' block

This block contains the entire set of dynamic ageing procedures representing the core of the model. They can be grouped into four main modules:

1. Demography
2. Health
3. Education, labour market and related incomes
4. Social security.

Each module is in turn composed of sub-modules. The sequence of modules and sub-modules is shown in figure 3, which also illustrates the order of simulated events.

Two crucial aspects of the model are worth mentioning here, namely that:

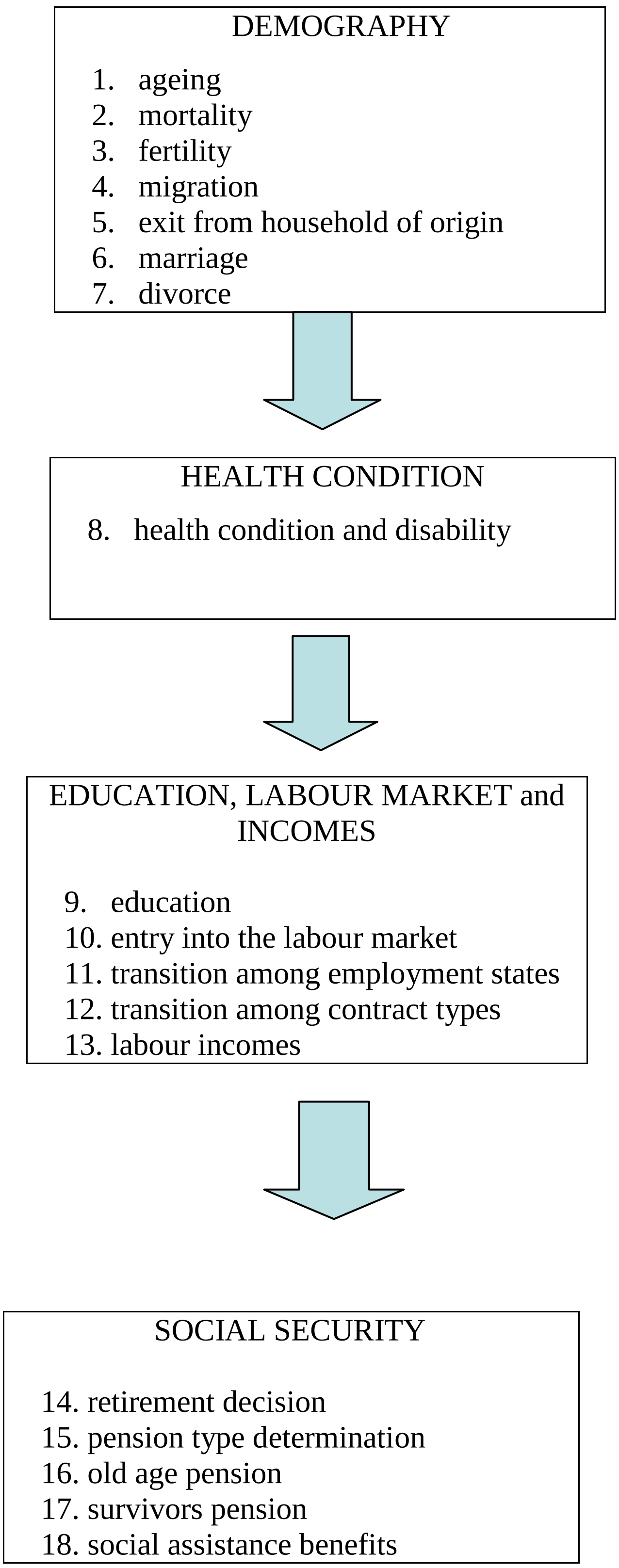
i) the model is *sequential*

ii) the model is *recursive*

The first feature rules out from the analysis any interaction between behaviours modelled within each single module. The second implies that once all the modules have been run, the model recommences the analysis of the same modules, in the same order, for the following year. These are two hypotheses that are frequently employed in dynamic population micro-simulation models. The introduction of a reaction function in such a model would be better in long-term General Equilibrium Models (Auerbach and Kotlikoff, 1987), employed mainly in the study of aggregate supply, inter-temporal consumption saving and capital accumulation choices.

The general rule for the dynamic ageing of socio-economic variables - which are not exogenously defined in the "scenario" block – is *probabilistically* based. In practice, the model estimates the probability of transition among states by means of models estimated using different statistical sources. The predicted probability is then matched to a random number drawn from a uniform distribution with support [0;1] (i.e. *Monte Carlo* technique). The set of events simulated using this technique are reported in figure 2.

Figure 2 Events simulated by CAPP_DYN



In theoretical terms, the general dynamic-ageing rule governing the socio-economic characteristics of those units in the population is based on the discrete, finite Markovian processes (chain) theory.

Given an event X, the probability of a transition from state x_i at time t to state x_j at time t+1 does *not* depend on past history, but is solely determined by the current characteristics at time t. Therefore, the transition probabilities $P_{ij} = P(X_{t+1} = x_j | X_t = x_i)$ can be represented by a strictly positive matrix, called a transition or stochastic matrix, where the *m* rows (*n* columns) identify the space of events in year t (t+1).

$$P_{m*n} = \begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1j} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2j} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{i1} & p_{i2} & \cdots & p_{ij} & \cdots & p_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mj} & \cdots & p_{mn} \end{vmatrix}$$

The i-th row of the transition matrix P : | p_{i1} p_{i2} ... p_{ij} ... p_{in} |, called the *probability vector,* represents the probability of all possible transitions of state x_i into any other state within the range of states, during period *t+1*.

Matrix *P* has the following properties:

- • it is a square matrix, the number of states being the same in year t and in year t+1;
- • $0 \leq \leq 1 \; {}_{ij}p \; \forall i, j$;
- $\sum_{j=1}^{n} p_{ij} = 1$ *i=1,2…,m*;
- the main diagonal elements represent the probability of inertia.

Transitions among states are simulated annually by means of a Monte Carlo experiment: every year the simulator generates a random number (u_{ks}) for the k-th observation and the s-th event drawn from a uniform distribution with support [0,1]. The transition occurs if $p_{ks}-u_{ks}<0$.

4. The core of the model: the ‘future’ block

This section examines the set of modules composing the “Future” block. Table 3 shows all the events simulated annually by the model, the method employed to estimate the transition probabilities, the set of covariates and the data source.

Table 3 Estimation methods, covariates and data sources for the simulation of the model events

Event	***Estimation***	***covariates***	***Source***
Demography			
Mortality	Transition matrix	Age, gender, birth year	ISTAT forecast, 2005
Fertility	Transition matrix	Age, gender, birth year, area	ISTAT forecast, 2005
Migration	Transition matrix	Age, gender, birth year, area	ISTAT forecast, 2005
Exit from household of origin	Transition matrix	Age class, gender	“Famiglie e Soggetti Sociali” ISTAT, 2005
Marriage	Transition matrix	Age class, gender, area, education, marital status	“Famiglie e Soggetti Sociali” ISTAT, 2003
Divorce	Transition matrix	Wife’s age class, area	“Famiglie, Soggetti Sociali e Condizioni dell’Infanzia” ISTAT, 2003
Health			
Disability	Transition matrix	Age, gender, area	Indagine sulle Condizioni di Salute, ISTAT 2003
Economy			
Education	Ordered Probit	Parents’ education, gender, area	ISFol PLUS 2003
Entry in the labour market	Transition matrix	Education, age, gender, area	Rilevazione Trim. forze di Lavoro ISTAT
Transitions between labour and non labour statuses	Multinomial Logit	Education, polynomial of age , area, birth cohort, sector marital status	Rilevazione Trim. forze di Lavoro ISTAT
Transitions between contractual types	Logit	Education, age, gender, area	ISFOL Plus 2003
Work income	OLS	Age, contributory seniority, gender, area, citizenship, professional qualification, work time (part time/full time), contract, sector, education	ISFOL Plus 2003

4.1 The demographic module

The set of demographic events can be divided into two groups: external events, which modify the population’s structure in terms of age, gender and geographical area; and internal events, which affect the household structure only. Ageing, mortality, fertility and immigration are included in the former group, while exit from the family unit, marriage and divorce are part of the latter.

The general functioning of the demographic module is depicted in figures 4 and 5. First, external events are simulated. Each yearly simulation ages the population by one year. Then, simulation goes on to establish the number of observations that exit the model due to the death of the persons in question. During the following step, the model simulates new births. Finally, the population stock also varies every year due to net migration.

Once the population size and composition have been defined for each period, the model starts the simulation of processes modifying the structure and the composition of household units (internal events). Children between 18 and 34 can leave their household unit of origin. Singles, living or not with their parents, can get married. The marriage event determines the creation of a new household unit. Widowed or divorced/separated individuals can get married, following the same rules applied to singles. Finally, the model simulates divorce for a share of married people, this event determining the split of the original household unit.

Figure 3 Demographic module - external events

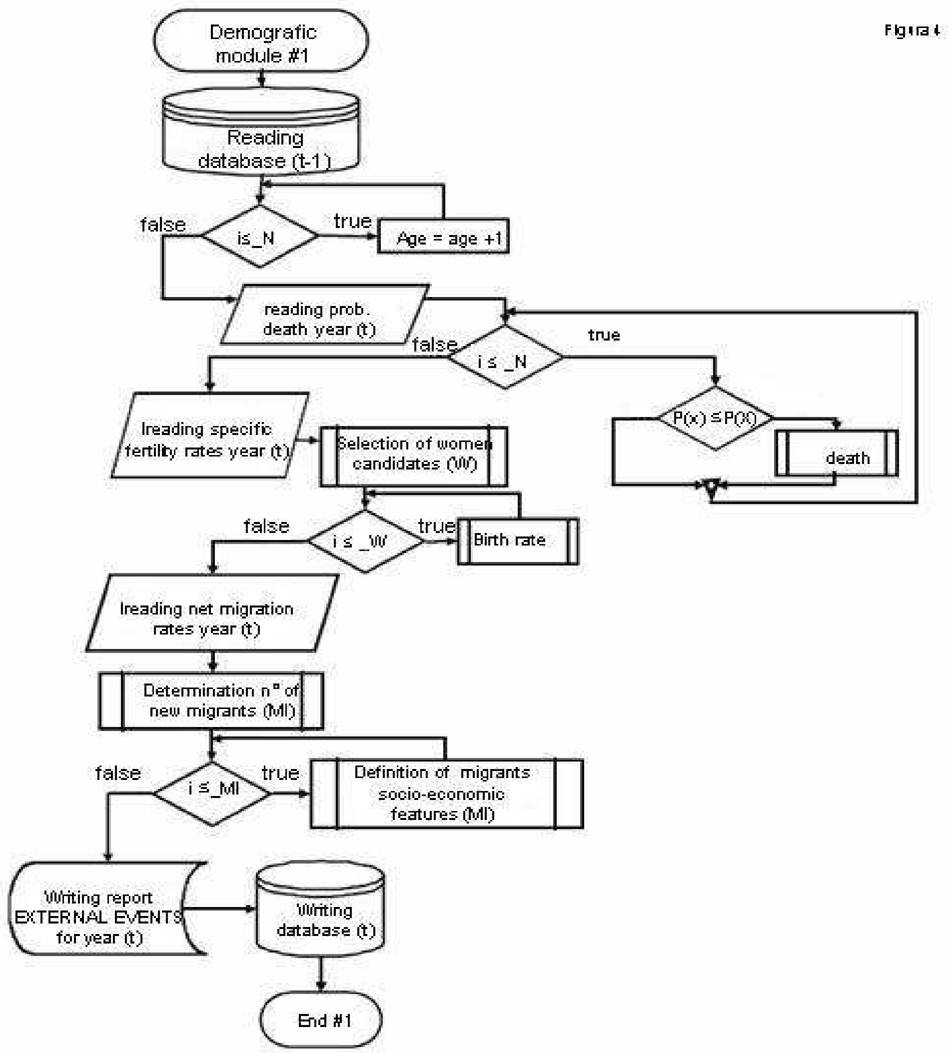
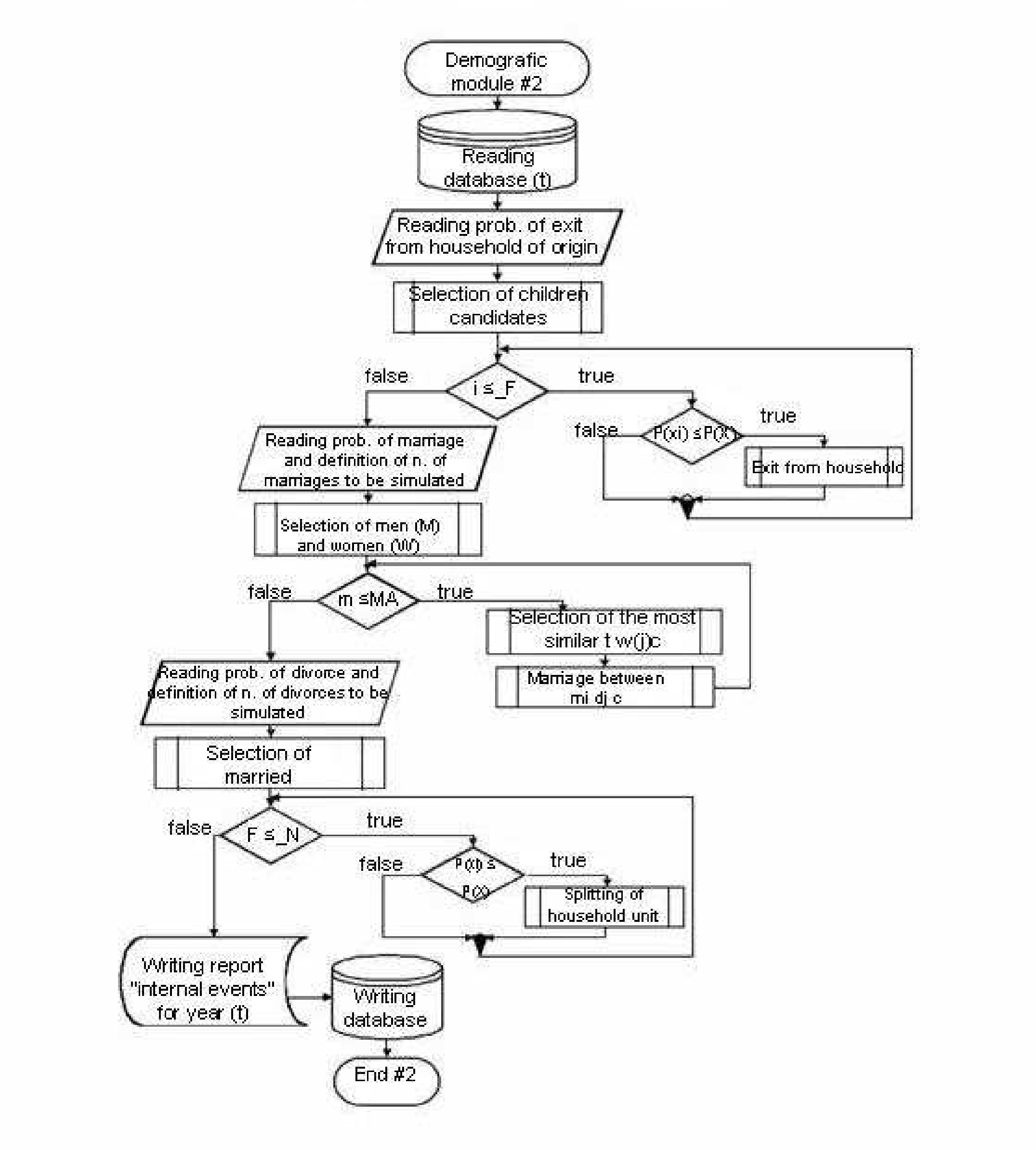
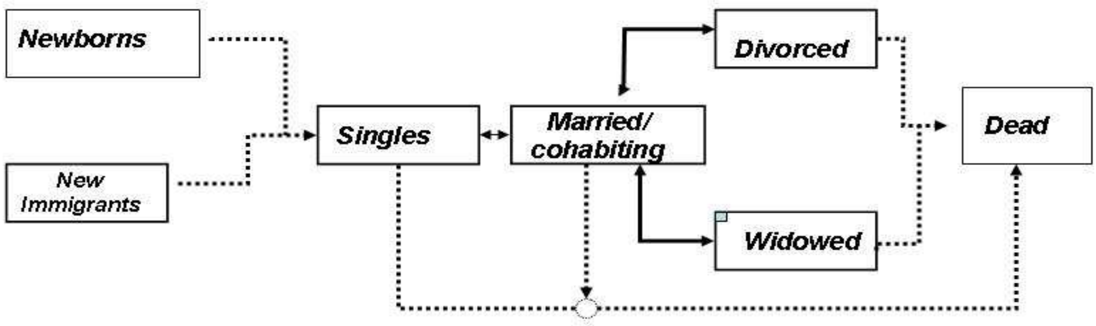


Figure 4 Demographic module - internal events



The model identifies four marital statuses (single, married/cohabiting, divorced, widower), allowing possible transitions across statuses according to the scheme showed in Figure 5.

Figure 5 Marital statuses transitions

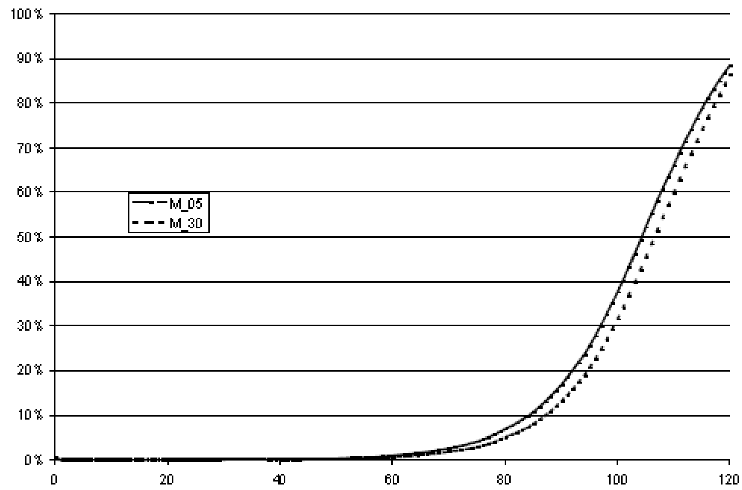


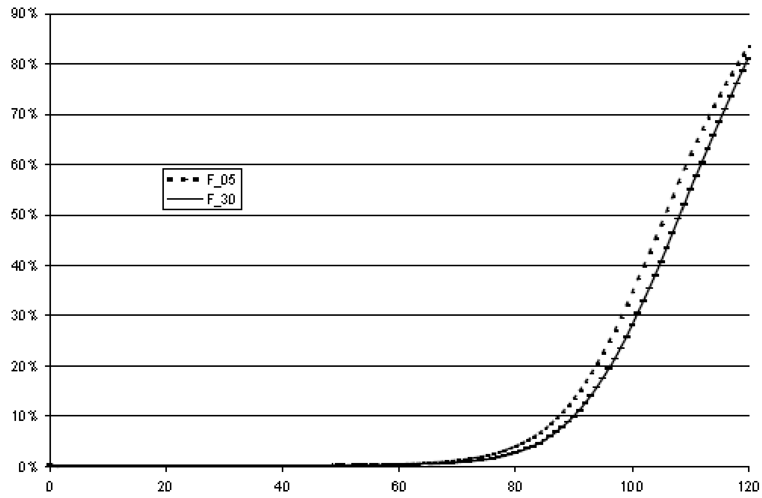
In the following, a detailed analysis of the main demographic sub-modules is presented.

4.2 The mortality module

The survival probabilities for the simulation are drawn from ISTAT's official projections (1/2005). It is worth remembering that ISTAT adopts an age-cohort approach when estimating death probabilities, in order to allow for the recently accepted phenomena (widely used in all developed countries) of the decreasing probability of death across all ages, and the substantial increase in old-age survival probabilities, particularly for women.

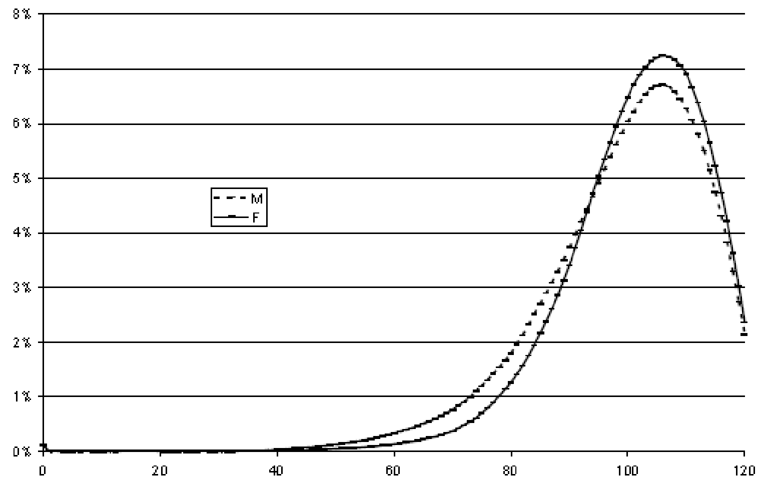
Figure 6 a and b Death probability by age and gender





Source: ISTAT, main projections at 1.1.2005. Central scenario.
Death probability on the right axis (national average).

Figure 7 Death probability variation by age and gender for year 2005 and 2030



Source: ISTAT, main projections at 1.1.2005. Central scenario.
Death probability on the right axis (national average).

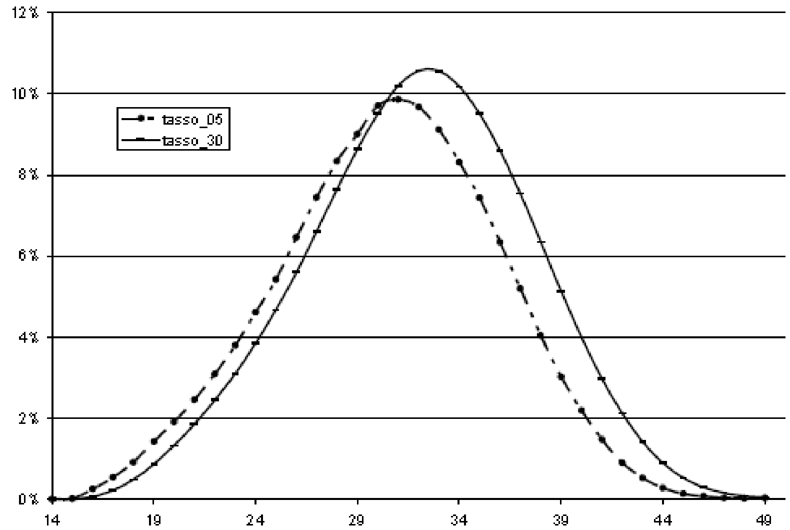
The mortality module works in the following way: given the year of simulation, age and gender, a random number drawn from a uniform distribution [0,1] is assigned to each observation. If the random value is smaller than the age-cohort-specific ISTAT death probability rate, then the model simulates death and consequently modifies the cohabitant's marital status.

On the other hand, if the random number is greater than the said rate, then the model ages the observation by one year.

4.3 The fertility module

The annual flow of newborn children is a function of the stock of women of child-bearing age (16-49) and of the ISTAT-specific fertility rates. ISTAT adopts an age-cohort approach in this case too. Figure 8 compares specific fertility rates for 2006 and 2030 according to the ISTAT median Scenario, which reveals a slight increase in total fertility rates due to the higher specific rates for women over the age of 31. On the other hand, the fertility rate of women younger than 31 is expected to decrease due to couples getting married later than before.

Figure 8 Specific fertility rates by mother age : 2005 and 2030 compared



Source: ISTAT forecasts, central scenario

Once the flow of newborn children, classified according to the age category of the mothers, has been determined, the model selects those women who are likely to have a child. Letting $f_a(c)$ be the probability distribution function for a married woman aged *a* with a number of previous children equal to *c*, the probability of that woman, of child-bearing age, having another child in year t+1 will be:

$P(c_{t+1} = c_t + 1 | a_{t+1}, c_t) = (1 - F_{a(t+1)}(c_t))$

where $F_{a(t+1)}(c_t)$ is the cumulative distribution function of $f_{a(t+1)}(c_t)$.

The aforementioned procedure enables us to establish the flow of newborn children according to the mother's age, taking account of the number of children previously living in the household unit. Once the newborn child is provided with household *id*, the model determines her/his socio-demographic characteristics and updates the household unit's size and composition. Gender is randomly assigned, given that the probability of the child being either male or female is the same.

4.4 The immigration module

The model simulates the net flows of immigrants each year according to official forecasts provided by ISTAT[6]. Hence, the forecast net immigration flow for the next few years lies within the range of 145,000 to 150,000 individuals per year.

The entry age of immigrants is classified according to the age-distribution of legally registered immigrants, according to the figures supplied by ISTAT. The model, however, excludes other members of the immigrant's family joining that person, and hence it is assumed that they are own on their own (single) as they enter the country. The flow of new immigrants is added to the stock of individuals (both immigrants and natives) who have previously settled, net of simulated deaths and births.

All the model's modules are applied to the entire simulated population, it being assumed that immigrants' behaviour is the same as that of the native population[7]. The imputation of socio-economic characteristics is carried out using the Monte Carlo method.

4.5 The exit from household unit

This sub-module allows the selection of children likely to leave the family household.

The increasing delay in the timing of young people's exit from the household is a well-established fact in Italy. Indeed, according to ISTAT, in 2003 60.2% of young people aged 18-34 lived with at least one parent (table 5): individual expectations and the increasingly difficult economic plight of the new generation were the main reasons for this phenomenon.

Recent ISTAT estimates show an increase in the percentage of employed young people living with their parents, together with a decrease in the percentage of those looking for their first job, while 32.3% of those young people who live with their parents are studying (ISTAT, 2004).

---

[6] In official demographic forecasts, international migration is usually considered less important than fertility or mortality. In fact, forecasts regarding migrations tend to be aleatory, since the mobility of populations is affected by social, economic, psychological, political factors which are very difficult to predict (Blangiardo, 1997).

[7] This hypothesis may be too narrow for the simulation of certain events, such as fertility, while other empirical evidence would seem to suggest a less marked difference in behaviour. For example, although the careers of immigrants are more mobile than those of native Italians, Anastasia Gambuzza Rasera (2005) saw evidence in the Giove 2004 archive of similar patterns of behaviour in the labour market
between immigrants and Italian workers after their initial entry into the labour market. Immigrants' income levels, on the other hand, were lower than those of Italians doing similar jobs.

Table 4 Singles, aged 18-34 living with at least one parents

	1993			*1998*			*2003*		
Class	*Males*	*Females*	*Total*	*Males*	*Females*	*Total*	*Males*	*Females*	*Total*
18-19	98,4	95,4	96,9	99,0	97,9	98,4	97,6	97,1	97,4
20-24	90,9	78,9	85,0	92,8	83,7	88,2	92,3	83,7	87,9
25-29	60,5	36,8	49,0	70,6	46,0	58,7	70,5	51,7	61,0
30-34	24,9	12,2	18,5	30,6	16,0	23,2	37,4	21,4	29,5
Total	64,0	48,9	56,5	66,2	51,1	58,7	66,8	53,6	60,2

Source: ISTAT (2004) "Indagine Multiscopo sulle famiglie:
Aspetti della Vita quotidiana; Famiglia, Soggetti Sociali 2003".
Mean value for 1993-1994, 1998 and 2003 for 100 young in the same age class.

Given the lack of forecasts concerning future trends for this phenomenon, the model uses expost probabilities drawn from table 5, in order to establish a steady-state exit rule: the likely yearly flow of young people leaving home was selected on the basis of a Monte Carlo process employing transition probabilities, that are conditional upon gender and age class, equal to 1- the probabilities shown in columns 8-9 of table 4.

4.6 Marriage module

The model allows for single people marrying each year, and the simulation of this event involves three steps: firstly, the flow of yearly marriages is defined as 4.3‰ of the total population[8]. Once the number of marriages is known, potential candidates (aged 16-60) are selected by means of a Monte Carlo process relying on marriage probabilities , depending on gender and age, provided by ISTAT's multi-purpose survey "Famiglie e soggetti sociali" (ISTAT, 2004)[9]. Candidates are then included in two separate databases, male and female, in order to enable new household units to be created.

Research points to the presence in Italy of positive assortative mating in marriage (Becker, 1991), whereby people choose their marriage partners in a nonrandom way, opting for those persons with similar educational backgrounds (Rossetti, Tanda, 2000) and professional status (Del Boca et al.,

[8] http://www.istat.it/salastampa/comunicati/non_calendario/20060424_00/indicatori_demografici.pdf.
The steady state hypothesis does not appear in this framework particularly strict: in fact, the marriage rate has not substantially modified in the last year.

[9] ISTAT does not publish marriage probability by age and gender but reports the number of individuals getting married each year only. Starting from this information, cohort and periods effects apart, we obtain yearly marriage rates dividing the number of individual getting married by age and gender for the total number of marriages each year.

2000). Furthermore, Borlini and Zajczyk (2001) discovered a high probability of individuals from the same geographical area, and with similar educational and professional backgrounds, getting married.

Women tend to get married at a younger age than men, and the likelihood of their marrying men of a similar age is assigned to all those women chosen, in order to allow for this gap[10]. Therefore, a Monte Carlo technique dependant upon a wife's age generates a variable containing the age class of a potential husband.

A marriage is then simulated, allowing for the matching of similar spouses according to a vector of observable features including dummies on educational background, marital status (single, divorced and widowed), geographical area and age group, in keeping with the propensity score method adopted by Rosembaum and Rabin (1983), Holland (1986), and Rubin and Thomas (2000). Each new household unit (including children from previous relationships) is provided with a HID (Household IDentification number), which remains unvaried for the entire simulation period.

4.7 The divorce module

Married couples are allowed to divorce with the resulting splitting of the household into two different units, each headed by one of the two divorced individuals. As with the marriage module, the divorce simulation is carried out in three stages. Firstly, the yearly flow of divorces is defined as 3‰ of the total number of married couple (ISTAT, 2003)[11]; secondly, those couples that are likely to divorce are selected: as ISTAT finds a different incidence of divorce events both at geographical level and according to age, this selection process consists of a Monte Carlo process based upon ISTAT probabilities conditional on the geographical area in question and the wife's age group. Within this group, a number of couples amounting to the yearly number of divorces to be simulated, is randomly selected; the splitting up of the household into two different units, and the updating of marital status and household composition variables, are then carried out[12].

---

[10] Probabilities are computed from ISTAT data, where the distribution of women's age at marriage is considered to be a function of the spouse's age group, on the total number of marriages each year. The mean age gap between men and women getting married is about three years.

[11] The steady-state hypothesis used in the divorce simulation exercise would appear stricter than the hypothesis used in the case of marriages, as statistics on this topic suggest a growing propensity towards divorce in recent years.

[12] Any children shall be deemed as belonging to the mother's household unit. According to ISTAT, in 85% of all cases, the mother is given custody of the children (those under the age of 18).

4.8 The disability module

The simulation of the disability condition is based on external information taken from the ISTAT Survey on public health and the use of the national health service; this survey, which is carried out every five years, covers a sample of more than 100,000 individuals of all ages. The most recent edition of the survey, which is the one used for the purposes of this paper, was conducted in 2005. The survey collects information about individuals' ability to perform certain basic daily tasks such as washing, eating and dressing, without the need for the help of others. There are 19 questions of this type, and they may be grouped into four categories, each of which may point to a different form of disability, namely: being unable to get out of the house, or having serious difficulties with movements, everyday activities, or in communicating with others[13]. For each of the four categories, therefore, we end up with a dummy variable which is given the value 1 if the individual is unable to perform that set of activities. This classification has been used to distinguish three levels of disability, each of which depends on how many of these dummy variables take the value of 1: the lowest disability condition (level 1) is that where the person is disabled in terms of only one of the four groups of variables; medium (level 2) disability corresponds to two dummy variables equal to 1; finally, a person is deemed severely disabled (level 3) if three or four areas of disability take a value of 1. Table 5 provides some basic descriptive statistics regarding the survey.

Table 5 Descriptive statistics of the Survey on health conditions and use of health services

	Whole sample	Disabled, level 1	Disabled, level 2	Disabled, level 3
Age	41.8	68.1	75.5	78.8
Woman	51.4%	63.0%	68.4%	70.0%
Compulsory education	59.7%	88.6%	89.5%	93.5%
High-school diploma	26.4%	8.8%	7.7%	4.4%
University degree	8.2%	2.6%	2.7%	2.1%
North	45.2%	41.5%	38.8%	39.2%
Centre	19.2%	18.8%	20.1%	21.8%
South	35.6%	39.6%	41.1%	38.9%
Widow	7.9%	36.9%	45.6%	53.7%
Remaining life expectancy (in years)	40.9	19.2	13.5	11.3
Number of observations	128040	2797	1869	1324

[13] For example, a person is defined as unable to perform basic everyday activities if he/she indicates a serious difficulty as a reply to at least one of the questions that fall within this category.

Average age increases with the seriousness of the disability condition, as does the proportion of women. The level of education is negatively correlated with the level of disability, as is the state of widowhood.

In order to assign to each individual in the simulation database a disability status, we propose and compare two alternative approaches:

a) *Pure ageing*: the ISTAT Health Survey is used to compute the proportion of disabled people within classes defined by gender and age (Costello and Przywara, 2007). These relative frequencies by gender and age are employed, by means of Monte Carlo methods, to select which sample members are attributed disability status. Three levels of disability have been identified. Note that under this scenario no cohort effect is taken explicitly into account, nor is it assumed that any future gains in life expectancy will be spent in a state of bad health. As a consequence, these projections are rather mechanical and as such risk producing distorted estimations of the number of disabled.

b) *Compression of disability*: the probability of being disabled is not constant within groups of the same age and gender, but depends on a vector of socio-demographic determinants. If these variables change, the probability of suffering from a disability should change accordingly. In order to take account of this endogeneity, we have performed an ordered probit estimation on the 2005 Health survey, where the dependent variable may be classified at four different levels: no disability (95.7% of total sample), low disability (2.1%), average disability (1.3%), severe disability (0.9%). The explanatory variables must be restricted to those socio-demographic characteristics that are common to both the Health Survey and the microsimulation model database, namely: age, gender, educational level, geographic area, widowhood. In addition to these explanatory variables, we have also included the residual life expectancy (in years) of each person, such data (depending on age and gender) being taken from the latest ISTAT estimates. The introduction of residual life expectancy is important, since if overall life expectancy rises, one would not expect the probability of becoming disabled to remain constant for any given age. Indeed, it is now widely recognised that this probability increases rapidly during the final years of one's life. In the presence of an ageing population, the omission of residual life expectancy from the regression would, at the simulation stage, result in an overestimation of the probability of becoming disabled, and therefore also of future LTC costs (Norton and Stearns, 2004). This second hypothesis may be considered to be a variant of the diverse theories asserting that the number of years spent in poor health should decrease as life expectancy increases (Manton et al. 2006). It is, nevertheless, more accurate and consistent with the data used to build the model, than the mere application of a simple ad hoc rule whereby the probability of being disabled increases with life expectancy.

In order to check the results of the application of this rule, we also create a comparison scenario with a very simple rule whereby the probability of becoming disabled changes each year in proportion to the increase in life expectancy. Costello and Przywara (2007) refer to this rule as the "constant health scenario".

The effect of observable socio-demographic characteristics on the probability of becoming disabled is modelled in terms of an ordered probit model, which takes the following form. Define an ordinal variable y {i: 1 . . .,N} indicating the observed level of disability among the sample members and y_i^* is the associated latent variable. The model has the following general structure:

$$y_i^* = X_i \beta$$
$$y_i = j \quad \text{if} \quad c_{j-1} < y_i^* \le c_j$$

where X_i denotes the vector of observable explanatory variables; β is a vector of coefficients, and ε is a random variable distributed as a normal. Given the nature of the data available, we ignore the possibility of unobservable personal characteristics which might influence both the level of disability and some of the explanatory variables. There are four different disability levels, denoted by j: 0 no disability, 1 low level of disability, 2 intermediate level of disability, 3 serious level of disability. The cut-off parameters *c* are estimated as part of the model. (A constant term is not identified in the model). Table 6 shows the results of the ordered probit estimate on the 2005 Health Survey. The explanatory variables relating to age are introduced using a spline function, and their coefficients how a marked increase in the probability of becoming disabled over the age of 70.

Disability status is strongly dependent on the level of education (the omitted variable is the graduate level), and on being resident in the southern part of Italy (the omitted geographic area). Residual life expectancy has a significant effect: if, in the future, life expectancy increases, this will lead to a reduction in the probability of becoming disabled for each year of a person's life.

Table 6 Ordered probit estimates of the probability of being disabled

	Coef.	Robust Std. Err.
<=30 years	-0.0551	0.0100
31-50 years	-0.0400	0.0097
51-60 years	-0.0255	0.0104
61-70 years	-0.0012	0.0089
71-80 years	0.0463	0.0077
>=81 years	0.0604	0.0054
Female (D)	0.3137	0.0424
Compulsory education (D)	0.4726	0.0409
High-school diploma (D)	0.1472	0.0476
Northern Italy (D)	-0.2386	0.0188
Central Italy (D)	-0.1518	0.0234
Widow (D)	0.0997	0.0234
Residual life expectancy (in years)	-0.0541	0.0099
Cut-points:		
C_1	-1.5867	0.7821
C_2	-1.1616	0.7821
C_3	-0.6392	0.7819

Number of obs = 128040; LR $chi^2(13)$ = 15988.04; Prob > chi^2 = 0.0000; Log likelihood = -21515.021; Pseudo R^2 = 0.2709. (D) indicates dummy variables.

Since CAPP_DYN projects all the model predictors, we are able to use the estimated coefficients and the cut-off parameters of this regression to predict the probability each year of an individual with characteristics X being in a condition of disability j as:[14]

$$pr(y_i^* = 0) = \int_{c_0}^{c_1} y_i^* dy = Norm[(c_1 - (X_i\beta + \varepsilon_i)]$$

$$pr(y_i^* = 1) = \int_{c_1}^{c_2} y_i^* dy = Norm[(c_2 - (X_i\beta + \varepsilon_i)] - pr(y_i^* = 0)$$

$$pr(y_i^* = 2) = \int_{c_2}^{c_3} y_i^* dy = Norm[(c_3 - (X_i\beta + \varepsilon_i)] - pr(y_i^* = 1)$$

$$pr(y_i^* = 3) = \int_{c_3}^{1} y_i^* dy = Norm[(X_i\beta + \varepsilon_i) - c_3]$$

In order to identify the disability level for the sample members for each year of the simulation, we use a Monte Carlo process. We assign no disability to those sample members who receive a random number z, drawn from a uniform distribution between 0 and 1 below the

[14] 14 We assume that the gradient of the disability rates (by levels), with respect to the socio-economic characteristics observed in the cross-section (year 2005), will remain constant in the future.

conditional probability of having no disabilities $pr(y_i^* = 1)$; we assign low level of disability if the random number are between $pr(y_i^* = 1)$ and $[pr(y_i^* = 1) + pr(y_i^* = 2)]$; we assign intermediate level if $[pr(y_i^* = 0) + pr(y_i^* = 1)] \leq z \leq [pr(y_i^* = 0) + pr(y_i^* = 1) + pr(y_i^* = 2)]$; finally, if z is between $[pr(y_i^* = 0) + pr(y_i^* = 1) + pr(y_i^* = 2)]$ and 1 the individual is assigned the most serious level of disability.

We assume also that if a person is deemed to be disabled in year t, he/she cannot then return to being classified as non-disabled at any point in his/her future; however, if that person is deemed to be less than seriously disabled, then he/she may be attributed a worse degree of disability in any subsequent year, up until death. We then randomly select (for each of the two alternative imputation approaches described above), from among those persons classified as being severely disabled, and who have been disabled for more than three years, a sub-sample of individuals to be admitted to nursing homes; this number corresponds to official estimates of the number of people living in such homes in Italy (ISTAT, 2007b).

Table 7 shows the association between socio-demographic characteristics of the population and the level of disability predicted by the probit model on the basis of the method used in Ermish and Francesconi (2001). The predicted probabilities are computed at the sample values, using estimated parameters (and cut-off points) from the model presented in table 2. The results can be read as follow. The "baseline probabilities" row displays the predicted average probabilities for each of the four levels of disability, when all the characteristics are set at their sample values for each person. The overall baseline predicted probability of being in one of the four states can be shown to be equal to 1, while the sum of the predicted probabilities of having one of the three levels of disability is equal to the rate of disability presented in the text (sum of the baseline probability of having a low level of disability=2.1%, an intermediate level=1.3%, a serious level=0.9%). The remaining rows in Table 3 show predicted probabilities related to particular values of the explanatory variables. In the case of age, for example, all the characteristics other than age are set at their sample values for each person, and predicted probability values for each person are averaged out over the entire sample.

Women have a disability probability that is always higher than that of men. Individuals with a low level of education display a 2.3% probability of having a low level of disability, compared with 1.4% for those with a high school diploma, while those the disability probability of those with a higher level of education is lower still.

People living in the South of Italy have a higher probability of disability than those living in the Centre and North of Italy.

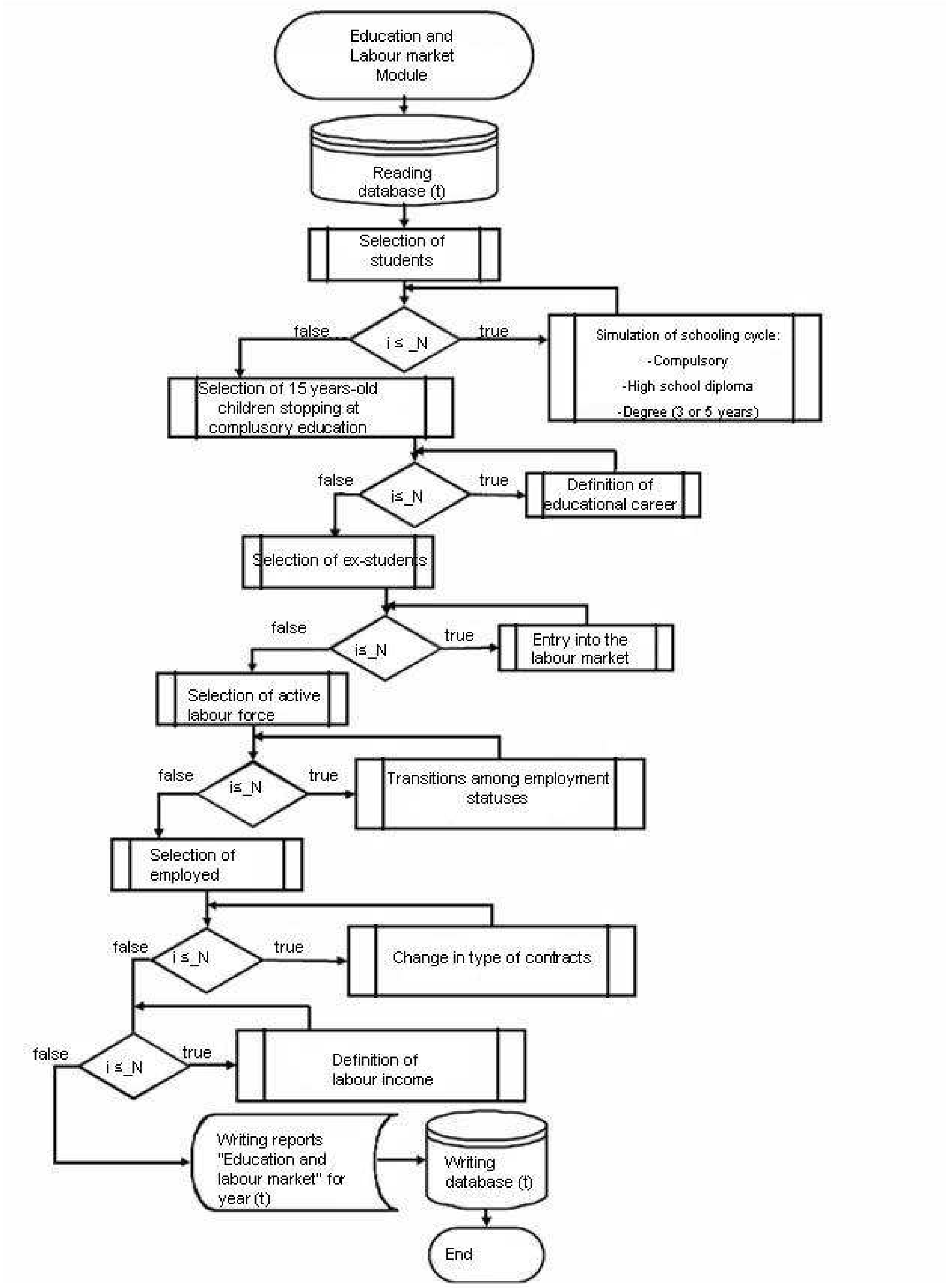
Table 7 Predicted probabilities of being disabled by level of gravity

	Not disabled	Low level	Intermediate level	Worst level
Baseline probabilities	0.957	0.021	0.013	0.009
55 years	0.970	0.016	0.009	0.004
65 years	0.976	0.013	0.007	0.003
75 years	0.966	0.018	0.011	0.006
85 years	0.925	0.034	0.024	0.017
Female	0.947	0.026	0.016	0.011
Male	0.967	0.017	0.010	0.006
Compulsory education	0.953	0.023	0.014	0.010
High school diploma	0.972	0.014	0.009	0.005
Degree	0.978	0.012	0.007	0.004
North	0.963	0.018	0.011	0.008
Center	0.958	0.020	0.013	0.009
South	0.948	0.025	0.016	0.012
Widow	0.952	0.023	0.015	0.010
Living with spouse	0.959	0.020	0.013	0.008

4.9 Education and labour market

After the demographic module, there is another module pertaining to education, to entry into and movement within the labour market, and to the calculation of earnings. This module is structured as shown in figure 10. Firstly, all individuals aged 16 are deemed to have completed their compulsory education. A higher educational level than this delays entry into the labour market until achievement of the imputed educational level (high school certificate, three year degree or five year degree). The end of the educational process is followed by entry into the labour market. Inputs/outputs into/from the labour force, together with changes in employment, are then simulated. The active population is divided into two sub-groups: on the one hand public and private employees, and on the other the self employed (part time/full time). A proportion of the population is employed on atypical, temporary contracts. Finally, the model simulates labour income and updates individuals social security and pension contributions records.

Figure 9 “Education and labour market” module structure



The steps illustrated in figure 9 are organized into the following sub-modules:

- education,
- entry into the labour market,
- employment transitions,

• transition between different forms of employment contract,

• labour income,

which will be analyzed below.

4.10 The education module

This module accounts for three educational levels:

1) compulsory education

2) upper secondary school

3) higher education (three- or five-year degree).

All individuals aged 16 are deemed to have completed their compulsory education; individuals can then decide whether to continue their education or look for a job.

Educational attainments are simulated by imputing coefficients obtained by an ordered probit estimation, the results of which are reported in table 6. The sample includes those people aged over 16, who have completed their education or are enrolled at a university[15], and are included in the 2004 ISFol PLUS survey[16]. The sample consists of 34,324 individuals.

The empirical model is structured as follows: we define y_i as the observed achieved educational Level, and y_{i*} as the corresponding latent variable. The alternatives have an ordinal form which implies the following general structure:

$$y_i^* = X_i\beta + \varepsilon_i$$
$$y_i = j \quad se \quad c_{j-1} < y_i^* \leq c_j$$

where X_i is the vector of individual (gender, geographical area and cohort dummies) and household (parents' presence and educational level[17]) characteristics; the parameters c_j, representing the threshold values, are estimated together with the column vector of β coefficients.

Regression results are reported in table 6: the first column shows estimated coefficients, while columns 2-4 display marginal effects for every single value of the dependent variable. As the results

[15] Following Checchi, Flabbi (2005) students enrolled at university are supposed to end their education by obtaining a degree.

[16] A problem for the empirical analysis in the determinants of "educational careers in Italy" comes from the lack of suitable statistical data for dynamic estimation purposes. Estimations have been conducted on different sample surveys. Several cross-section data fail to allow for the extrapolation of cohort and period effects, which as such have been less frequently examined in the empirical studies that have been conducted in Italy so far. ISTAT's polled cross-section of the work forces allows for the analysis of cohort effects in educational-rate dynamics (Leonbruni, Richiardi 2006), but it does not allow for educational choices to be conditioned by the nature of individuals' family backgrounds. A pooling of SHIW surveys allows for the joint analysis of the aforesaid two effects. However, in the case of youngsters living with their parents, implying potential estimation distortions due to the "selection effect" (Heckman, 1979), the recent ISFOL PLUS survey of a sample of more than 40,300 individuals aged between 15 and 64 (ISFOL, 2006), provides detailed information (from telephone interviews) about respondents' educational attainments and their families' socio-economic conditions. The information furnished by this survey is therefore more suitable, compared with other information, for the purposes of the estimation of educational choice determinants allowing for recent social and cultural, as well as legislative, changes in Italy.

[17] Presence of parents refers to the year interviewed was 15.

of the regression show, educational attainments appear to be strongly dependent on parents' educational level and the geographical area in which the family lives. Women are more likely to achieve a higher level of education: the likelihood of enrolling at university is *coeteris paribus* 1.5% higher for women than for men, and the probability of going no further than compulsory schooling is 2.9% lower for women than for men. Cohort dummies suggest a positive trend in schooling for younger generations. *Coeteris paribus* an individual born after 1979 has a greater probability (+4.4%) of getting a university degree than an individual born in the period 1971-75. On the other hand, the same individual is 7.9% less likely to go no further than compulsory secondary education.

In the dynamic simulation, the coefficients in table 6 will be employed to impute educational choices. In practice, the probability of the dependent variable assuming a value of 1, 2 or 3 are calculated for each observation at the end of compulsory schooling[18]. In other words:

Table 8 Ordered probit of educational level

Education(y)	coefficient	y=Pr(j==1) 0.4214	y=Pr(j==2) 0.4589	y=Pr(j==3) 0.1196
compulsory_mother***	-.8062	.2803	-.0655	-.2147
	(.0702)	(.0202)	(.0042)	(.0231)
High School_mother***	-.3054	.1209	-.0680	-.0529
	(.0689)	(.0273)	(.0171)	(.0102)
Compulsory_father***	-1.328	.4176	-.0312	-.3864
	(.0554)	(.0121)	(.0090)	(.0196)
High School father***	-.5229	.2062	-.1235	-.0826
	(.0550)	(.0212)	(.0146)	(.0068)
No_mother***	-.3204	.1270	-.0734	-.0536
	(.0581)	(.0230)	(.0151)	(.0079)
No_father***	-.1977	.0782	-.0425	-.0356
	(.0437)	(.0174)	(.0103)	(.0071)
women***	.0755	-.0295	.0144	.0151
	(.0211)	(.0082)	(.004)	(.0041)
Centre***	.1987	-.0765	.0339	.0425
	(.0287)	(.0108)	(.0043)	(.0066)
South***	.0961	-.0374	.0179	.0195
	(.0232)	(.0090)	(.0042)	(.004)
Co_min_1950***	-.4847	.1914	-.1123	-.0790
	(.0404)	(.0155)	(.0104)	(.0055)
Co_1951_1960***	-.2435	.0961	-.0515	-.0445
	(.0428)	(.0169)	(.0098)	(.0071)
Co_1961_1965**	-.1265	.0498	-.0259	-.0239
	(.0483)	(.0191)	(.0105)	(.0086)

[18] The same procedure is applied to the students over 15 of the base year, in order to define the human capital level. The imputed value is constant all over the simulation.

Co_1966_1970**	-.1279	.0503	-.0260	-.0242
	(.0457)	(.0181)	(.0099)	(.0082)
Co_1976_1978***	.1895	-.0726	.0311	.0414
	(.0419)	(.0157)	(.0058)	(.0099)
Co_1979_plus***	.2047	-.0786	.0341	.0444
	(.0387)	(.0145)	(.0056)	(.0090)
_cut1	-21.282			
	(.0777)			
_cut2	-7530			
	(.0767)			
N	34323			
R:squared	0,935417			

Thereafter, each observed individual aged at least 15 is assigned a random number (z) drawn from a uniform distribution with support [0,1] which will be compared to the probabilities estimated using the previous formula. If the number is lower than $pr(y_{i^*}=1)$, then the i_{th} individual will have attained the basic educational level; if z of the i_{th} is between $pr(y_{i^*}=1)$ and $[pr(y_{i^*}=1)+ pr(y_{i^*}=2)]$ the corresponding educational attainment will be a high school diploma. If z is higher than $[pr(y_{i^*}=1)+pr(y_{i^*}=2)]$ higher education is assigned[19].

4.11 Entry into, and transition within, the labour market

Missing information regarding employment status, occupational attainments and type of activity of students from the previous year are assigned using conditional probabilities drawn from transition matrices built on the sub-group of individuals moving from the educational sphere to the labour market (*Rilevazione trimestrale sulle forze di lavoro ISTAT 2001-2002* – hereinafter *RTFL*).

Occupational attainments and sector are assumed to be time-invariant over the whole simulation period for each individual, whereas employment status and contractual arrangements are allowed to change over the course of time.

Changes in jobs are simulated in the sub-modules “transitions”, while transitions from one type of contract to another (typical, atypical, permanent,temporary) are determined in the sub-module “employment contract”.

Concerning the number of transition in the labour market, CAPP_DYN allows for four employment statuses and, and in keeping with other dynamic MSMs, it assumes that employment decisions depend solely on individual characteristics, and are thus independent of demand-side factors.

[19] Since detailed information on university careers is not available, we assume 30% of enrolled university students finish their studies with a three-year degree, while the remaining 70% attain a five-year degree.

Transitional probabilities are estimated on the basis of the 1993-2003 RTFL. Individuals aged 16 to 64, excluding pensioners and students, can be classified as:

- full time workers (those working at least 31 hours)
- part-time workers (those working less than 31 hours)
- unemployed
- outside the labour market (unemployed individuals not looking for a job)

In table 7, ex-post transition probabilities are estimated on the selected sample which is composed as follows: 68.8% workers, 9.24% unemployed individuals, and 22.70% individuals outside the labour market. Each cell shows the proportion of individuals who, starting from status *i* , have attained status *j* by the following year.

Results suggest a strong degree of immobility between statuses, as shown by the main diagonal figures.

Full-time workers are characterised by greater stability, while part-time workers appear more mobile: 28.4% of the latter move toward a full-time position the following year, 5.55% become unemployed, while 4.64% exit the labour market altogether.

More than 23% of the unemployed find a job after one year (19.41% of them find a full-time job), while 7.8% exit the labour market. 3.69% of inactive individuals enter the ranks of the employed, while 2.70% begin to look for a job[20].

Table 9 Transition probabilities among employment statuses from year t to year t+1

Initial status	Employed_FT	Employed_PT	Unemployed	Non_active
Employed_FT	95,16	1,89	1,95	0,99
Employed_PT	28,41	61,4	5,55	4,64
Unemployed	19,41	3,92	69,18	7,48
Non_active	2,58	1,11	2,7	93,61
Mean	63,44	4,62	9,24	22,7

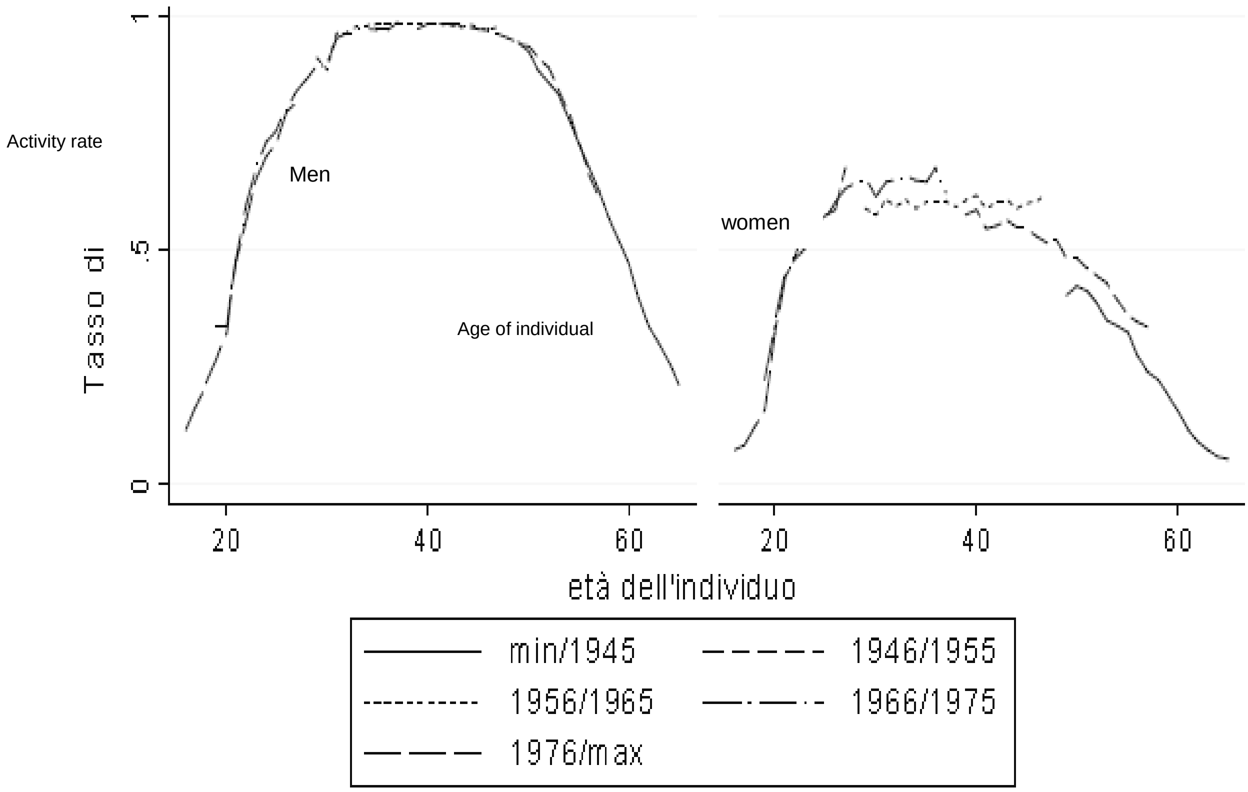
The figures shown in Table 9 can be interpreted as mean trend indicators for the whole sample during the ten years period in question. However, one may reasonably assume that transition probabilities depend on individual characteristics (gender, age, education, geographical area).

[20] Probability refers to occupational mobility from one employment status to another after 1 year. Of course, during the course of the year an individual may have experienced more than one transition. In such cases, account is taken of the transition between the initial and the final status.

Furthermore, it should be pointed out that the period in question is characterized by a strong, prolonged increase in employment due to the greater degree of female participation in the labour market, and to increased presence of part-time employment of various kinds, depending on birth cohorts.

The graphs in Figure 10 show some results according to cohort, gender and geographical area. Results are in the main in keeping with expectations and with the empirical evidence set forth in other studies[21]: while participation rates are approximately invariable among male cohorts, among women there is a general increase in participation in the case of those persons aged 30 to 50, with rates among younger women nearly always suggesting that a greater share of them are making educational choices which delay entry into the labour market. Moreover, the younger cohorts contains a significant share of part-time workers.

Figure 10 Prticipation rates to labour market 1993-2003 by age, gender and cohort.



Transition probabilities are estimated on a pseudo-panel RTFL 1993-2003 by using a multinomial logit. The dependent variable is the employment status in the final year[22]. Estimations have been

[21] See Trivellato et al. (2005), Leonbruni, Richiardi (2005), ISTAT (2004).
[22] ISTAT data are not free from *caveats*. Trivellato et al. (2005) suggest not using too detailed classifications. Notwithstanding the acknowledged limits, the transition matrices analysis of the RTFL allows for the use of a broad data set which is representative of the Italian population.

carried out for each initial status and gender (four initial statuses *per* gender, therefore 8 models have been estimated).

Given the initial status of the h_{th} individual, the conditional probability of transition or immobility in the following year (J=j) can be represented as[23]:

$$P(J = j \mid I, X_h) = \frac{\exp(X_h \beta_j)}{1 + \sum_{j=1}^{4} \exp(X_h \beta_j)}, j = 1, \ldots, 4$$

where *j* is one of the four feasible statuses, Xh is the covariates vector (education, second order polynomial in age, geographical area, marital status, activity and eight cohort dummies) and $_j\beta$ is the vector of coefficients which vary according to each state.

Table 10 shows 24 coefficients for both women and men, while column 12 presents the different types of transition[24]. A positive coefficient, *coeteris paribus*, increases the transition probability from the initial to the final status. For instance, the negative coefficient of the "public" variable for full-time women moving to other employment statuses suggests that they have a smaller chance of moving from a full time job to other statuses in the public sector; the negative coefficient for women in the Centre and in the South regarding their moving from unemployment or non-activity to employment, suggests a greater likelihood of finding a job in the North than in the Centre or the South.

As Chies et al. (1998) have pointed out, this table can be read by row and by column: by row it provides information as to whether, and to what degree, a variable affects transition frequency, hile by column it displays the characteristics of those individuals who are more/less likely to move from one status to another.

The move away from a situation of unemployment or non-activity is more frequent in the North than in the Centre or South for both men and women. Education positively affects the probability of getting and keeping a job. The older a person is, the lower the likelihood of a transition from employment to unemployment, for both women and men, although the same variable has a negative effect in the case of unemployed or inactive individuals willing to enter into the labour market. As one would expect, the careers of public employees tend to be more stable than those of private-sector workers. The share of young people losing their jobs falls as the age cohorts rise. Young

[23] The multinomial logit model is valid under some condition. The most important, known as the independence of irrelevant alternatives (IIA), establishes that errors ε_{ij} are independent of j, i.e. the odds-ratios are assumed to be constant between two alternatives, even if the number of alternatives increases.

[24] Coefficients are computed with reference to the inertia condition i.e. when the status is unchanged from period t to period t+1 (i = j).

women with full-time jobs are increasingly moving towards part-time employment, unlike their older counterparts, and this trend is stronger among married women, probably due to the presence of children[25].

The transition module uses the coefficients in table 8 to compute the individual odds-ratios of employment transitions. Finally, a Monte Carlo process allows the simulation of mobility for each active individual[26].

Table 10 Multinomial logit for transitions in the labour market

	Employed_FT-> employed_PT	Employed_FT-> Unemployed	Employed_FT-> Inactive	Employed_PT-> Employed_FT	Employed_PT-> Unemployed	Employed_PT-> Inactive
Compulsory	0.273***	0.746***	1.109***	-0.482***	0,15	0.657**
High_school	-0.03	0.19	0.253*	-0.418***	-0.025	0,158
Age	0.069**	-0.04	-0.125***	-0.082**	-0.206***	-0.171***
Sq_age	-0.001**	0.00	0.001***	0.001	0.002	0.002**
centre	-0.125*	0.197**	0.13	0.066	0.264*	-0.151
South	-0.345***	0.965***	0.898***	0.441***	1.329***	0.810***
public	-0.200***	-0.556***	-0.652***	-0.016	0,103	0.024
Married	0.540***	-0.432***	1.035***	-0.170**	-0.540***	0.611***
C_48_52	-0.19	-0.346*	-0.321**	-0.363**	0,502	-0.496**
C_53_57	-0.16	-0.692***	-0.328*	-0.472***	0,162	-0.803***
C_58_62	0.08	-0.671**	-0.18	-0.733***	0,138	-1.140***
C_63_67	0.370*	-0.731**	-0.01	-0.898***	-0.227	-0.980***
C_68_72	0.397*	-0.949***	-0.18	-1.020***	-0.482	-0.879**
C_73_77	0.499*	-0.968**	-0.31	-1.133***	-0.923	-1.793***
c_78_max	0.526*	-1.038**	-0.15	-0.825**	-1.311*	-2.005***
_cons	-5.022***	-1.273*	-2.288***	2.368***	2.600*	1.292

	unemployed-> employed_FT	Unemployed-> Employed_PT	Inactive-> Inactive	Inactive-> Employed_FT	Inactive-> Employed_PT	Inactive -> Unemployed
Compulsory	-0.711***	-0.642***	0.669***	-1.074***	-0.884***	-1.006***
High_school	-0.386***	-0.533***	0.316**	-0.603***	-0.671***	-0.594***
Age	0.007	0.118**	0.009	0.102***	0.255***	0.059*
Sq_age	0	-0.001	0	-0.002***	-0.003***	-0.002***
centre	-0.692***	-0.750***	-0.295***	-0.067	-0.394***	-0.115
South	-1.378***	-1.534***	-0.228***	-0.355***	-1.206***	0.297***
Married	-0.003	0.299***	1.280***	-0.794***	-0.610***	-1.171***
C_48_52	-0.268	0.005	-0.362*	-0.239*	0.087	0.077
C_53_57	-0.38	0,233	-0.254	-0.371**	0.028	0.011
C_58_62	-0.533	0,491	-0.188	-0.513***	0,316	-0.063
C_63_67	-0.574	0.806*	-0.232	-0.502**	0.618*	0.029
C_68_72	-0.644	0,739	-0.383	-0.533**	0.789**	0.019
C_73_77	-0.543	1.127*	-0.364	-0.424	0.934**	0,152
c_78_max	-0.538	1.222**	-0.319	-0.641*	0,307	-0.007
_cons	0,383	-4.647***	-2.556***	-2.035**	-7.230***	-1.479**

[25] The presence of children in the household is not controlled by the model.

[26] Professional qualifications (blue collar, white collar, self-employed, manager) remain unvaried during one’s lifetime. This is a widely-employed simplification.

4.12 The income module

The "income" module simulates the yearly labour income of the active population. The estimation of life-cycle labour income is based on the 2004 ISFol PLUS cross-section survey. This survey specifically focused on the active population, and includes useful information for the purposes of dynamic simulation, such as nationality and type of employment contract.

The econometric model specification is the following[27]:

$$\ln y_i = \alpha + \beta X_i + \varepsilon_i$$

$$\varepsilon_i \sim N(0,\sigma^2_\varepsilon)$$

where $\ln y_i$ is the log of individual labour income gross of personal taxation, and the X vector includes the set of observables which are usually employed in a human capital model *a là* Mincer.

The income level is separately determined for employees and self-employed workers. The group of employees is in turn sub-divided according to educational level and gender[28]. Table 9 shows the estimation results based on certain ISFol sample subsets.

The signs of coefficients are in line with expectations. In particular, income is "bell" shaped with respect to age, and increases with educational level and seniority, while it appears, *coeteris paribus*, lower for women, people from the South of Italy, immigrants, blue-collar workers and public employees.

On average, blue-collar workers' wagesare, *coeteris paribus*, significantly lower than those of white-collar workers, while managers with secondary or higher education receive in turn 6% to 10.5% higher salaries than white collar workers.

Public employees earn less than employees in the private sector. The negative value of the gender dummy suggests lower income for self-employed women, while the income of self-employed individuals with a university education is 45.4% higher than that of those who left school after they had finished their compulsory education. Finally, atypical workers earn 45.8% less than the self-employed.

---

[27] Errors are assumed to be normally distributed with a zero mean and a σ2 variance.

[28] The limited availability of observations for self-employed graduates leads us not to disaggregate data by gender. For the same reason we have decided not to break down the sub-sample of the self-employed by gender and education.

Once the coefficients are estimated, the level of gross income is computed for each individual, taking into account observable characteristics only (age, work experience, education, etc).
However, individual income differs from mean income for two reasons: firstly, due to the presence of an individual component (constant over the entire period) which can be interpreted as a proxy of professional ability and effort; secondly, due to the presence of a yearly component which can be thought of as the increase in productivity of all workers in each simulation period.

Table 11 OLS estimation coefficients of log gross labour income

	Employees			Self-employed
	Men	Women	Graduated	
Age	0.0351***	0.0233***	0.0436***	0.0790***
	-	-	-	-
Sq_age	0.0003***	0.0002***	0.0004***	0.0008***
	-	-	-	-
Women	0.2052***	0.5418***	0.2052***	0.5418***
Noth	0.0423**	0.0460***	0.0303	0.0662
	-	-		-
South	0.0600***	0.0733***	-0.0643**	0.1807***
immigrants	-0.0577	-0.0662	-0.1942**	-0.0969
				-
Atypical				0.4584***
	-	-	-	
Partime	0.4079***	0.3460***	0.3556***	0
Years_contrib	0.0029***	0.0069***	0.0032**	0.0013
High_school	0.1176***	0.1360***	0.3593***	
Degree				0.4536***
	-	-	-	
fixed_term	0.0702***	0.0614***	0.1644***	
	-	-	-	
Blue_collar	0.0780***	0.0990***	0.1832***	
Mager	0.0602***	0.0620***	0.1057***	
	-		-	
Public	0.0708***	-0.0104	0.0852***	

_cost	9.0353***	9.0260***	9.0727***	7.9228***
N	4772	4819	3144	3127
Sq_R	0.3291	0.395	0.38306	0.2045

The distinction between these components enables us to compute the level of gross labour income for each individual according to the following formula:

$$\hat{y}_{i,t} = e^{\frac{1}{2}\hat{\sigma}^2 + \log \hat{y}_{it}} e^{(u_i)} (1 + \tau_t)$$

The first term on the right-hand side is an unbiased estimation of the mean gross wage for individuals with similar observable characteristics, where $\hat{\sigma}^2$ is an unbiased estimator of the error variance, and log $\hat{y}_{i,t}$ is the estimated log $y_{i,t}$ (Wooldridge, 2003).

The term e^{u_i} represents interpersonal variability among workers with similar observable characteristics who nevertheless display differing levels of ability, talent, etc. This component is typically non-observable, and in practice, deviations from the mean are recovered as residuals from the first stage regression. Of course, this procedure is feasible for those individuals employed in the base year only. For those individuals who enter the active population subsequently (future births, students, unemployed), this term is randomly generated - when they enter into the labour market – from a normal distribution with zero mean and variance equal to the estimated variance from the first-stage regression (Root MSE).

Thus all workers are provided, in this way, with an individual fixed effect, the distribution of which is constant in time, implying differences in wages resulting from unobservable characteristics that remain constant throughout the simulation period.

Finally, the $(1+\tau_t)$ factor allows the wage level to be linked to the medium/long-term growth in productivity, which is calibrated through the "scenario" block. Once again, the demographic evolution and the increase in human capital stocks over the following decades raise the mean income level, since age and education have a positive effect on labour income[29]. However, in this model, endogenous growth is lower than that forecast by the RGS, since the model does not allow for the expected increase in productivity.

In order to avoid over/under-estimations of earnings growth rates for subsequent decades, the following procedure is adopted: every year, a *pro-quota* growth factor *e* - equal to the difference

[29] Other factors could have a negative effect; for instance, an increase in female participation in the labour market, an increase in immigration, and the divulgation of part-time contracts.

between the exogenous earning growth fixed in the “scenario” and the growth estimated by the model - is added to the endogenous growth due to socio-demographic evolution.
The term ${}_t\tau$ is given by:

$$\tau_t = \prod_{i=1}^{t} (1+m_i) - \left(\frac{E(\bar{y}_t)}{E(\bar{y}_{t-1})} \right)$$

where *m* is exogenously determined in the “scenario”, while $E(\bar{y}_t) / E(\bar{y}_{t+1})$ describes the endogenous growth rate generated by the model.

5. The social security module

Individual retirement choices and the computation of old age, seniority and survivors pension benefits, and of social welfare allowances, social welfare increases (*maggiorazioni sociali*) and supplements *(integrazioni al minimo*), are all simulated in this module.
The individual pension transfer depends on the following variables:

- the life-cycle profile of labour incomes;
- the seniority of social security contributions upon retirement;
- the contribution rate during one’s working life;
- exogenous, macroeconomic growth during the period of pension contributions;
- the pension scheme;
- retirement age.

The first four variables depend on the results provided by the previous “demographic” and “education and labour market” modules. In particular, the life-cycle profile of labour incomes depends on the evolution of characteristics controlled for in the regressions reported in table 2.2.9.
Seniority at the moment of retirement depends on the total number of years during which an individual has received a positive labour income according to the model. Due to the chance of transitions between one employment status and another during the course of an individual’s working life, a proportion of simulated individuals may display certain periods during which no contributions were paid in.
The model simulates the following pension benefits:

1. Old age and seniority pensions,
2. Survivors’ and indirect pensions,
3. INPS disability pensions,
4. Civil disability pensions,

5. Social allowances,

6. Supplements to the minimum (*integrazioni al minimo)*

7. Social assistance supplements (*maggiorazioni sociali*).

The amount of the benefit is held constant in real terms for the whole retirement period, in accordance with the general pension indexation system introduced by the 1992 reform[30]. Minimum pensions, contribution caps and minimum and maximum thresholds employed to calculate the benefit according to different pension schemes, all increase over the course of time on the basis of the real GDP growth forecasts selected in the "scenario"[31].

5.1 The decision to retire

New retirees are selected for each year of the simulation following a two-stage procedure. At the first stage, all individuals satisfying the necessary conditions (in terms of age and contributory seniority) for old age and/or seniority pensions are duly identified. This identification process is implemented according to the law in force, in particular according to laws 243/2004 and 247/2007[32].

The first opportunity to retire arises when an individual satisfies the age requirements for early retirement (rather than the statutory age - which is currently set at 60 years for women and 65 for men[33]. In this case, the model checks whether an individual's exit from the labour market is inter-temporally [financially] advantageous. In practice, it compares two options: either the individual keeps working a further year, or immediately exits the labour market (that is, he/she retires early). If net social security wealth is greater under the second option, then the retirement choice is effectively simulated if the replacement rate exceeds a certain threshold, set at 60%.

Therefore, the model provides for two relevant factors affecting retirement choice: the first is an evaluation of inter-temporal convenience, and the second concerns the adequacy of the pension benefit provided by the social security system[34].

---

[30] The model does not allow for inflation.

[31] The path of real GDP growth is exogenous to the model. The implemented procedure differs from the pension scheme provisions in force, which allow for adjustments to pensions and minimum social benefits through legislative action. However, we believe that the option of keeping the minimum pensions constant at base year values is unrealistic, and that requiring discrete adjustments in certain years is quite arbitrary.

[32] Early retirement requirements were modified by with respect to those established by Law 243/2004: the age requirement for seniority retirement was increased to 58 in 2008, while the seniority requirement itself (i.e. the minimum number of years of welfare contributions made by a worker) remained at 35. From 2009, a system of quotas is going to be introduced: a worker may then take early retirement if the sum of his/her age and the years of contributions is higher than 95 (until the end of 2010), 96 (until the end of 2012) and 97 (from 2013 onwards), with the agerequirement however increasing from 59 in 2009 to 61 in 2013. Requirements for the self-employed are one year higher.

[33] This option will also remain with the DC system for those workers with 40 years contributions, regardless of their age.

[34] The choice of the particular threshold value adopted is clearly sensitive to the determination of the actual mean age of retirement. The choice of 60% seems to be a "reasonable" option which in practice restricts the seniority

The second way of retiring is to satisfy the statutory age requirements, which are subject to the minimum requirements for social security contribution seniority, which in turn amount to 20 years for those individuals who come under the defined benefit (DB) or mixed system, and 5 years for those under the defined contribution (DC) system. In the latter case, the exit rule and the provision of the (old age) pension benefit depends on the gender of the worker. It should be pointed out once again that the regulation in force establishes a different retirement age for men and women (65 for men, 60 for women). Once they reach the age of 65, and their minimum contributory requirements have been fulfilled, then men are supposed to retire, regardless of the financial convenience or adequacy of the pension benefit they receive. In the case of women, a different rule has been adopted in order to avoid a substantial flow of pensioners being replaced at a very low rate, especially during the period in which the DC system is in place. We believe that retirement at 60 to be unrealistic for those individuals who, especially during the latter part of the simulation, are on average better educated –and thus coming into the labour market later on – and whose pensions will be established according to the DC rule, which tends to penalize early retirement.

In order to avoid the formation of a stock of young pensioners receiving very low pensions, we calibrate the model so that the actual retirement age progressively increases every ten years until it reaches 65 in the year 2050.

5.2 The computation of old age and seniority pensions

The computation formulas used for the estimation of the amount of the first pension transfer, for each of the three pension schemes (the defined benefit, the mixed and the defined contribution varieties) are shown below. In general, computing the pension benefit is not an easy task, due to the need to take account of both individual and household income when providing for social transfers such as supplements to the minimum (*integrazioni al minimo*) and social assistance supplements (*maggiorazioni sociali).* Where possible, the calculation of such social benefits takes account of the socio-economic characteristics of individuals, together with the temporal evolution of such characteristics.

- The defined benefit regime

Individuals with at least 18 years of contribution history by the end 1995 fall under the DB regime. The rule used to calculate pension benefits may be expressed by the formula:

$$P_{DB} = r^*(N_1W_1 + N_2W_2)$$

---

retirements in the first part of the simulation period only.

where r is the pension rate of return, N_1 and N_2 are the years of contribution before and after 1992 respectively, while W_1 and W_2 represent those earnings that may be used to estimate pensions in the case of contributions paid before and after 1992 respectively.

The terms in the DB formula are not the same for all workers, as they depend on the pension scheme and on the pensionable wage level in question. In particular, W_1 is equal to the last salary received in the case of public-sector employees, and to the five- and ten-year average for private-sector employees and the self-employed respectively. W_2 is the ten-year mean wage for public and private sector employees, and the 15–year mean wage for the self-employed. The rate of return r is equal to 2% for the pensionable earnings bracket 0 to 36,980 Euros in 2002, and it decreases as earnings levels rise, down to 1.1% for the pensionable earnings bracket over 49,156 Euros.

Pensioners who have paid at least 20 annuities of contributions, but have not reached the minimum pensionable amount, are granted a supplement bringing them up to the minimum level.

- The mixed regime

This regime applies to those workers with a social security contribution history of less than 18 years in 1995. In such cases, the old age/seniority pension benefit is calculated as the sum of two components; the first component, P_A , is computed according to the DB formula on those contributions paid in before 1995, while the second component, P_B, is computed according to a DC rule on post-1995 contributions. The formula is as follows:

$P_{mixed}=P_A+P_B$

where the general rule for determining P_A is similar to the formula used in the DB regime.

Nevertheless, in the “mixed” regime, the pensionable wage for those contributions paid between 1992 and 1995, is determined differently, being computed as the mean wage over the years after 1992 indexed to a 1% yearly rate according to a simple compounding rule. The P_B term of the mixed pension is computed according to a DC rule.

Pensioners who have at least 20 years of contributions, but who have not reached the minimum pensionable amount, are topped up to the minimum level.

- The defined contribution regime

This regime covers those individuals who came onto the labour market after 1995, whose pensions shall be defined according to the following formula:

$P_{DC}=k*MC$

where k is the transformation coefficient that varies with retirement age so as to guarantee a quasi-actuarial equity between the present value of paid contributions and the present value of expected pension benefits. In order to allow for the expected evolution of mortality, which is represented in the model by ISTAT’s official forecasts, we have used the coefficients computed by RGS in its forecast model which adjusts pensions every ten years according to varying demographic conditions.

MC is the "*montante contributivo*" i.e. the total contributions accrued during an individual's entire working life, capitalized at the rate of growth of nominal GDP, and defined by the sum of the contribution paid compounded at the GDP growth rate. The annual contribution is computed as a proportion of an individual's gross wage, in the case of employees, and of gross income in the case of the self-employed. The contribution rate is set at 33% for employees and 20% for the self-employed. A contributory cap is set at 82,404 Euros. In order to receive the minimum pension, a person must have paid at least five annuities of contributions. Finally, the entity of the pension cannot exceed the amount of the social allowance increased by 20%. Furthermore, individuals will not receive a pension if requirements are satisfied prior to the statutory retirement age of 60 for women, and of 65 for men. For the pensions provided for under the DC scheme, no supplement up to the minimum is allowed, while a supplement up to the level of social allowance is provided if income requirements have been fulfilled.

- Survivors and indirect pensions

The death of a pensioner or of an insured individual entitles the survivors to a survivor's or an indirect pension, respectively.

The model allows the payments of indirect pensions if the dead person had paid in at least 5 years' contributions. When the total amount due to the survivor has been determined, shares are assigned to every single component of the household, their economic conditions being taken into account when doing so[35].

In particular, the model distributes the total amount among the survivors by assigning:

- 60% to the spouse
- 20% to each child if the spouse is still alive
- 40% to each child if the spouse is not alive.

The sum of the share cannot exceed the 100% of the pension the retired person would have been entitled to.

If only one child is entitled to this survivor's pension, the share is set at 70% of the original entitlement. The entity of this type of pension depends on the economic conditions of the beneficiary. The allowance is reduced by 20%, 40%, and 50% should the income earned by the beneficiary exceed amount of the minimum pension benefit by a factor of 5, 4, or 3 times

[35]Current regulations can be found by consulting the INPS website in the section *"la pensione ai superstiti"*.

respectively. This latter rule does not apply should underage students or disabled children be jointly entitled to the said pension.

5.3 Civil and INPS disability pensions

The social security module selects the beneficiaries of disability allowances, disability and civil inability pensions.

The disability allowance is calculated according to the standard system of old-age pension computation, which is based on the pensionable earnings and on the contributory seniority of the insured individual.

If the allowance is lower than the minimum benefit, it can be supplemented to bring it up to this minimum amount.

The disability pension consists of two components: one part is calculated according to the pensionable earnings and to the contributive seniority, as is the case with the inability allowance, while the remaining part, known as the "*maggiorazione*" (additional amount), is calculated according to the difference between the inability allowance and the pension he/she would receive had he/she accrued a level of seniority increased by a period amounting to the difference between the year the inability allowance started to operate and the pensionable age (set at 60 years for men and 55 years for women with a dispensation of at least 80% for disabled persons). No contributions beyond 40 years will be taken into consideration.

In order to calculate disability pensions using the DC system, the sum of contributions accrued is added to a share of contributions for the gap period between the initial pension year and the person's 60th birthday. Once again, seniority may not exceed the 40 years mark.

The entity of civil disability pensions are in line with the values provided by current law, and are updated according to growth in GDP from the second year of the simulation onwards.

The module providing entitled individuals with the civil disability pension benefit, checks whether the yearly income requirements have been satisfied or otherwise.

5.4 Social allowances

Over the age of 65, those individuals with an individual/household income that falls below the statutory limit, are entitled to a social allowance. In fact, Italian regulations establish that the said social allowance shall be paid out when the economic conditions of the beneficiary or his/her spouse fall below a certain level. In order to allow – at least partially - for these conditions, each year the model adds up the labour and pension incomes of the spouses in question. The benefits are then paid out when the total income is lower than the statutory thresholds, which are annually revised in accordance with growth in real GDP.

In general, the entity of this allowance is designed to bring overall income up to the legal minimum. For example, in 2007 the monthly allowance amounted to 389.36 Euros. This means a yearly allowance of 5,061.68 Euros (389.36 x 13 monthly installments), which also corresponds to the income limit for an unmarried applicant; the yearly limit is thus 5,061.68 x 2 = 10,123.36 when the applicant is married. In general, the yearly allowance is equal to the gap between the current annual social allowance and the applicant's actual total income–summed to the spouse's income if this exists-[36]. Unlike the social pension[37], the social allowance - or a share of it - is paid to the applicant even if he/she has a personal income over and above the individual legal threshold, as long as the overall income of the spouses is lower than the legal threshold for the couple.

5.5 Supplements to the minimum (integrazioni al minimo) and social assistance supplements (maggiorazioni)

As has already been mentioned, when the pension benefit entitlement falls below the minimum legal amount, the model provides the pensioner with a supplement or a social assistance increase, allowing for personal and household income as well as for the pensioner's age[38].

REFERENCES

Ando A. and Nicoletti Altimari S., (2004), A Micro Simulation Model of Demographic Development and Households' Economic Behavior in Italy, Temi di discussione Banca d'Italia n. 533.

Atkinson A B, Gomulka J., Sutherland H. (1988), Grossing-up FES data for Tax-Benefit Models, in *Tax Benefit Models*, (ed.) Atkinson A. B., Sutherland H., STICERD Occasional Paper no. 10, LSE.

---

[36] So for instance, if the applicant is married and the couple's yearly income is 9,000 Euros, the allowance is reduced to 1,123.36 Euros (the difference between 10,123.36, current annual income threshold for a married applicant, and the actual income of 9,000 Euros).

[37] From the 1st January 1996, the social allowance replaced the social pension, which however continued to be paid to those entitled individuals who applied for it by the 31st of December 1995.

[38] 38 Minimum benefits are calculated according to the provisions reported on the INPS website in the section "il trattamento minimo".

Auerbach A. and L: Kotlikoff, (1987), Dynamic Fiscal Policy, Chicago University Press.

Borlini B., Zajczyk F., (2001), "Matrimonio e lavoro: l'influenza del coniuge", in Bianco M. L. (ed), L'Italia delle diseguaglianze, Carocci, Roma

Blangiardo, G.C., (1997), Elementi di demografia, Il Mulino, Bologna.

Brandolini A., (1999), The distribution of personal income in post-war Italy: source description, data quality, and the time pattern of income inequality, Temi di discussione, n. 350, Banca d'Italia, Roma.

Cannari L., D'Alessio G., (1992), Mancate interviste e distorsione degli stimatori, Temi di Discussione, No. 172, Servizio Studi, Banca d'Italia, Roma.

Checchi D., (2003), Scelte di scolarizzazione ed effetti sul mercato del lavoro, in Lucifora, C. (a cura di), *Mercato, occupazione e salari: la ricerca sul lavoro in Italia*, Mondatori, Milano.

Checchi D., and Flabbi L., (2005), Intergenerational mobility and schooling decisions in Italy and Germany, mimeo.

Chies L., Lucchetti R., Staffolani S., (1998) Occupazione, disoccupazione, inattività: determinanti della mobilità tra stati in Italia, Rivista Italiana degli Economisti, n.3.

Costello D, Przywara B (2007) Projecting future health care expenditure: drivers, trends, policy recommendations, forthcoming in *European Economy*, European Commission, Directorate General for Economic and Financial Affairs.

Del Boca D., Locatelli M., Pasqua S., (2000), Employment Decision of Married Women: Evidence and Explanations, Labour, vol.14.

Ermisch J, Francesconi M (2001) 'Family Matters: Impacts of Family Background on Educational Attainments', *Economica*, 68 (270), 137-156.

Heckman J.J., (1979), Sample selection bias as a specification error, Econometrica, 47:153-161.

ISFOL, (2006), Impianto dell'indagine PLUS, Roma.

ISTAT, (2000), Le strutture familiari, Collana Informazioni n. 17, Roma.

ISTAT, (2002), La mobilità nel mercato del lavoro: principali risultati aprile 1998-aprile 2002, Collana Approfondimenti, Roma.

ISTAT, (2004), Forze di lavoro Istat. Anni 1993-2003, Annuario n 13, Roma.

ISTAT, (2004), Matrimoni, separazioni e divorzi. Anno 2000, Annuario n 13, Roma.

ISTAT (2007), Condizioni di salute, fattori di rischio e ricorso ai servizi sanitari (anno 2005), May (http://www.istat.it/salastampa/comunicati/non_calendario/20070302_00/) [accessed 15 July 2007].

Leombruni R., Richialdi M., (2006), LABORsim: an agent-based microsimulation of labour supply. An application to Italy, mimeo.

Klevmarken, N. A (2005)Dynamic Microsimulation for Policy Analysis – Problems and Solutions, paper presented at The Economics Conference of Australia 2005, Melbourne Sept 26-28

Mandrone E., Radicchia D., (a cura di.) (2006), PLUS: participation, labour, employment, survey. Indagine campionaria nazionale sulle caratteristiche e le aspettative degli individui sul lavoro, Rubbettino Editore, Soveria Mannelli (CZ).

Manton K G, Gu X, Lamb V L (2006) 'Changes in the prevalence of chronic disability from 1982 to 2004/05 as measured by long term changes in function ad health in the US elderly population', *Proceedings of the National Academy of Sciences*, 103, 18374-79.

Mazzaferro C., Morciano M., (2005), Un modello di microsimulazione a popolazione dinamica per la stima degli effetti distributivi della riforma pensionistica, paper presentato alla Conferenza SIEP 2005, Università di Pavia.

Ministero del Lavoro e delle Politiche Sociali, (2005), Aspetti distributivi del sistema pensionistico nella prospettiva del passaggio al sistema contributivo, Roma disponibile sul sito http://www.welfare.gov.it e sul sito www.capp.unimo.it.

Ministero della Solidarietà Sociale, (2008), Studio prospettico della povertà in Italia, Rapporto finale, Roma.

Morciano, M. (2007), Un modello di microsimulazione a popolazione dinamica per l'analisi del sistema di protezione sociale italiano, Unpublished PhD Thesis, University of Bologna.

Norton E C, Stearns S C (2004) 'Time to include time to death? The future of health care expenditure projections', *Health economics*, 13(4), 315-327.

O'Donoghue C., (2001), Dynamic Microsimulation: A Methodological Survey, Brazilian Electronic Journal of Economics, Universidade Federal de Pernambuco, vol. 4(2).

Orcutt G.H., Mertz J., Quinke H (Ed.), (1986), Microanalytic Simulation Models to Support Social and Financial Policy, North-Holland, Amsterdam.

Rosembaum P.R., Rubin D.B., (1983), The Central role of the Propensity Score in Observational Studies Causal Effects, Biometrika, n. 70, 1.

Rossetti S., Tanda P., (2000), Human capital, wages and family interactions, Labour, vol.14.

Trivellato U., Battistin E., Rettore E,. (2005), Choosing among alternative classification criteria to measure the labour force state, IFS Working Papers W05/18, Institute for Fiscal Studies.

Vagliasindi P., (2004), Effetti redistributivi dell'intervento pubblico: esperimenti di micro simulazione per l'Italia, Giappichelli, Torino.

Zaidi A., Rake K., (2002), Dynamic microsimulation models: a review and some lessons for SAGE, Sage discussion paper n. 2, London.