T-DYMM 3.0
FORECAST MODEL REPORT

L'Istituto nazionale per l'analisi delle politiche pubbliche (INAPP) è un ente pubblico di ricerca che si occupa di analisi, monitoraggio e valutazione delle politiche del lavoro, delle politiche dell'istruzione e della formazione, delle politiche sociali e, in generale, di tutte le politiche economiche che hanno effetti sul mercato del lavoro. Nato il 1° dicembre 2016 a seguito della trasformazione dell'Isfol e vigilato dal Ministero del Lavoro e delle politiche sociali, l'Ente ha un ruolo strategico – stabilito dal decreto legislativo 14 settembre 2015, n. 150 – nel nuovo sistema di governance delle politiche sociali e del lavoro del Paese. L'Inapp fa parte del Sistema statistico nazionale (SISTAN) e collabora con le istituzioni europee. Da gennaio 2018 è Organismo intermedio del PON Sistemi di Politiche Attive per l'Occupazione (SPAO) del Fondo sociale europeo, delegato dall'Autorità di Gestione all'attuazione di specifiche azioni ed è Agenzia nazionale del programma comunitario Erasmus+ per l'ambito istruzione e formazione professionale. È l'Ente nazionale all'interno del consorzio europeo ERIC-ESS che conduce l'indagine European Social Survey.

Presidente: *Sebastiano Fadda*
Direttore generale: *Santo Darko Grillo*

Riferimenti
Corso d'Italia, 33
00198 Roma
Tel. +39 06854471
www.inapp.org

MOSPI – MODERNIZING SOCIAL PROTECTION SYSTEMS IN ITALY
VS/2018/0414

T-DYMM 3.0 FORECAST MODEL REPORT

This report was produced for the MOSPI project, which received funding under the call for proposals VP/2018/003 of the Directorate-General for Employment, Social Affairs and Inclusion of the European Commission. The opinions expressed in this report reflect only the authors' view. The European Commission is not responsible for any use that can be made of the information contained therein.

Acknowledgements
This research project is indebted to the National Social Security Institution (INPS), the Italian National Institute of Statistics (ISTAT) and the Finance Department of the Ministry of Economy and Finance (MEF-DF) for sharing their databases within the framework of the Italian National Statistical System (SISTAN) and for providing precious insights on metadata for the scopes of our research. In this regard, we are particularly grateful to Leda Accosta (INPS), Francesca Lariccia (ISTAT), Paolo Acciari (MEF-DF) and their respective units.
We are thankful for the constant support of SOGEI S.p.A. in the elaboration of the present report and in the development and maintenance of the Treasury Dynamic Microsimulation Model (T-DYMM) through the years. We are particularly grateful to Francesca Di Brisco and her unit, 'Economic Modelling and Statistical Analysis for Policy Making'.
We are much obliged to Professor Marco Alfò, who provided precious advice and guidance on methodological and statistical aspects of the analysis.

The present report is the result of a joint work by MEF and FGB. Authors of each chapter are listed below.

Authors
Chapters 1 and 7: *Stefano Boscolo*, *Riccardo Conti*, *Alessandra De Castro*, *Elena Fabrizi*, *Elena Giachin Ricca*, *Ottavio Ricchi*, *Massimo Tangucci*, *Serena Teobaldo*, *Federico Tullio*
Chapter 2: *Feliciano Iudicone*
Chapter 3: *Elena Fabrizi*
Chapters 4 and 5: *Michele Raitano*
Chapter 6: *Michele Bavaro*, *Paolo Liberati*, *Simone Tedeschi*

Report completed on March 2020

Table of contents

Introduction

The aim of the Modernising Social Protection Systems in Italy (MOSPI) project is to "support the modernisation of the social protection system in Italy allowing to respond to the challenges of digitalisation, the changing world of work, the aging of population". This research subject requires the development of analytical tools for the analysis of the effectiveness of public policies in addressing long-term consequences of labour market outcomes. Keeping the primary focus on the adequacy of the pension system, there will also be scope for considering more general issues related to social security programs. The current work stream grounds on two previous projects, respectively titled 'Innovative Datasets and Models for Improving Welfare Policies' (INDIW) and 'Improving Effectiveness in Social Security' (IESS), which were implemented in the periods 2009-2012 and 2014-2016. The goal is to improve on what was already achieved by enhancing:

1. The unique and innovative dataset called 'Administrative SILC' (AD-SILC hereafter), which is developed by matching longitudinal information from several administrative archives, held by the National Social Security Institution (INPS), with IT-SILC survey data (i.e. the Italian component of the EU-SILC) collected by the National Institute of Statistics (ISTAT);
2. The dynamic microsimulation model called T-DYMM (Treasury Dynamic Microsimulation Model), designed to simulate the evolution in terms of demographic and labour market outcomes of a representative sample of the Italian population.

The present document follows on the first delivery of the MOSPI project, titled 'Background report on future of work scenario: The dynamics of non-standard work in Italy'. At the outset, that report reviewed the main legal characteristics of Non-Standard contracts and showed the recent trends in their utilisation by Italian companies, relying on administrative data provided by INPS. The available information was also used to provide insight on Non-Standard Workers (NSWs), identified as the potentially most vulnerable individuals engaged in the labour market. The Report dealt with the dispersion of wages earned by workers exposed to different contract frames and analysed labour market transitions, such as frequency of hires and contractual transformations, singling out the patterns of NSWs. Finally, it provided an original contribution by looking at the impact of digitalization on the labour market by delving into the issue of platform workers; such an analysis was enabled by the availability of the innovative PLUS 2018 survey dataset, managed by the National Institute for Public Policy Analysis (INAPP), which provides information on the matter.

Concerning the specific goals of the MOSPI project, in our proposal we pledged to look into different issues related to the changing labour market conditions by identifying and evaluating, by means of different scenarios, possible outcomes in the medium to long term. Our analytical and micro-simulation modelling efforts will allow focusing on risks associated to individuals characterised by discontinuous participation in the labour market and suffering from several specific deficiencies or disadvantages within the social protection system.

There is a wide range of risks faced by less protected/more vulnerable workers as a result of the interaction between unfavourable labour market outcomes, characterised by discontinuous careers and low upward mobility, and the provisions of labour law and social insurance legislation. These individuals could face little to no access to social protection and/or enjoy very poor end-of-career pension entitlements, especially in the context of a (Notional) Defined Contribution computation formula. The latter problem, namely inadequate tertiary age income, would be aggravated by individual myopia or limited financial capabilities 'distorting' saving behaviour and investment decisions over different financial instruments, for instance by forgoing voluntary contributions. Indeed, against this backdrop, one objective of our project is to assess the relevance of the risks of inadequacy of public pension benefits, whose long-term performance can be simulated in T-DYMM by projecting working careers and simulating the resulting pension benefits. Several enhancement to the previous version of the model will allow the delivery of a more accurate analysis of non-standard and atypical workers' career and remuneration prospects, monitoring if – under the current rules – they could face difficulties in accessing social security and identifying contribution gaps leading to discontinuity in social protection. The overall adequacy of the current welfare system will be evaluated by assessing the pros and cons of the current design of the Italian second and third (private) pension pillars. We shall focus on the impact of different options for the integration of public and supplementary pension schemes, while taking into account the different propensity of workers in contributing to pension funds.

The final aim is to include in the final report simulations that could provide insight for the improvement of the current policy setting by identifying measures that might mitigate the impending risks and, more in general, help addressing social needs in the present as well as in a forward-looking perspective. The existing instalment/release of T-DYMM (henceforth, T-DYMM 2.0) is already suitable for investigating some of the aforementioned issues. The 2016 Final Report of the IESS project[1] provided an assessment of the adequacy of the Italian welfare system, projecting poverty and inequality indicators into the long-term and focusing on the effects of the many reforms operated in the pension system over the past 30 years and, more recently, in the legislation relative to unemployment benefits. Our effort aims at enhancing the model features in several dimensions. The present report describes what has been achieved in the initial stages of the project.

This report pursues three main aims. First, it provides a detailed explanation of the characteristics and advantages of the new release of the AD-SILC dataset, where a large sample of Italian workers is followed longitudinally since their entry in the labour market up to 2018. Before making use of the upgraded dataset within the dynamic microsimulation in future stages of the project, in this report we extensively explore the available data to provide quantitative assessments over the characteristics and trends of individuals' labour market outcomes up to 2018. Finally, this work presents the main advances in the development of the T-DYMM model. More in detail, the report is structured as follows.

Chapter 1 looks into the new dataset, 'AD-SILC 3.0'. The core of AD-SILC is built by merging data from the Italian version of the EU-SILC survey with longitudinal information provided by administrative archives managed by INPS. This allows to reconstruct individuals' work and life patterns by controlling for a much higher number of variables than those included in both the original INPS and SILC datasets. This chapter provides details on aspects of data processing and describes the additional information content included in the newest version of AD-SILC, due to both the higher number of variables obtained by SILC and the new data sources involved. These include the SHIW survey data from the Bank of Italy, Tax Declarations and Cadastral data from the Department of Finance and survey data from INAPP.

The following chapters represent the core of the present report, as they provide descriptive and econometric analyses on AD-SILC 3.0. Before presenting these results, with the aim of shedding some light on normative aspects relating to non-standard and atypical work, Chapter 2 provides a brief review of the various types of contractual arrangements existing in Italy and covered in AD-SILC, in light of the recent developments of the Italian labour market legislation. Afterwards, Chapters 3-5 present detailed analyses on the labour market run by using AD-SILC. First, in Chapter 3, applying survival analysis techniques, we analyse the duration of the various types of contractual arrangements in the 2004-2018 period and investigate the role played by various individual characteristics. Chapter 4 focuses on the trends of earnings distribution in private employment in 2004-2018, with a further specific focus on in-

1 See: http://www.fondazionebrodolini.it/sites/default/files/iess-final-report.pdfhttp://www.fondazionebrodolini.it/sites/default/files/iess-final-report.pdf.

work poverty risks It also shows original results about the trend of various indicators of that distribution and, applying decomposition techniques, about the role played by individuals' and firms' characteristics as a major driver of the trend of earnings inequality observed in 2004-2018. Chapter 5 devotes its attention to longitudinal individual working careers (i.e. from the entry in the labour market up to 2018), with the explicit aim of assessing the position of individuals who may have experienced penalising patterns in the context of a Notional Defined Contribution (NDC) pension scheme. To this aim, two explicit focuses are carried out: i) labour market outcomes over a 20-year period of the first cohorts of workers entirely enrolled in the NDC scheme (i.e. those who entered the labour market in 1996-1998), in order to provide insights about the possible spread of risks of inadequate pension benefits due to unfavourable career patterns; ii) on the same outcomes achieved by individuals who started working in 1996-2013 observed over a shorter – 5-year – period, in order to observe whether individual labour market patterns in the early phase of the career have deteriorated in recent years.
In order to complete the outlook with a view on financial wealth, Chapter 6 presents evidence on data from the Bank of Italy's SHIW survey, one of the main additional sources of the new AD-SILC dataset. Descriptive analyses are presented together with some preliminary evidence on the determinants of financial decision, with specific focuses on the role of financial literacy and of specific working conditions (non-standard employment and atypical contracts) on investment choices.
Lastly, Chapter 7 is devoted to our microsimulation model. It outlines the modelling strategy of T-DYMM in the context of the relevant literature on the subject and describes the enhancements to implement within the MOSPI project. The new model, henceforth 'T-DYMM 3.0', shall include a number of structural innovations: a Migration Module will be developed to account for immigration and emigration processes; a Wealth Module aimed at modelling the whole spectrum of movable and immovable assets at the household's disposal will be developed; a sub-module on working pensioners will be included to simulate more realistically the transition between work and retirement. The Labour Market, Pension and Taxation Modules will all be expanded and updated in order to account for the recent innovations in the relevant legislation. The information on Non-Standard Workers available in INPS archives, Tax Declarations records and INAPP's PLUS survey will allow an accurate identikit and simulation of this target group of fragile workers. Finally, the simulated sample will enjoy a higher representativeness of the reference population, as *ad hoc* re-weighting techniques are to be applied.

1. The AD-SILC dataset[1]

Introduction

The first step in our quest to identify, assess and tackle the issues relating to the changing features of the labour market is the construction of an appropriate database. In doing so, we have made use of the experiences of the INDIW and the IESS projects to develop an updated and expanded version of the AD-SILC dataset.
The next section presents the key features of the old versions of AD-SILC, which is constructed matching longitudinal information from administrative archives, held by the INPS, with survey data collected by the National Institute of Statistics (ISTAT). The second section describes the contents of the data sources of AD-SILC 3.0, namely, IT-SILC and the INPS archives, providing a description of the main innovations in terms of variables included. Section 3 turns the attention to the new data sources (survey data from the Bank of Italy, administrative data from the Department of Finance) to be used in the current version of the dataset. Section 4 concludes by outlining the privacy and methodological issues faced in merging the data and the consequences in terms of representativeness.

1.1 The previous versions of AD-SILC

Longitudinal data is crucial in the analysis of labour market outcomes. It allows analysing the dynamics of earnings and the individual transitions among different occupational statuses, highlighting the relation with socio-economic characteristics that are traditionally identified as drivers of labour dynamics. We postpone to Chapters 3-4 a discussion on such drivers. In this framework, *linked* administrative datasets (i.e. datasets combining survey and administrative data) provide a tool, constantly growing

[1] Due to delays in the delivery from the National Institute of Statistics (ISTAT), data for the analyses was only made available on February 6, 2020. At the time of drafting of the present report (early March 2020), a handful of SILC variables still have to be delivered.

in popularity, for researchers conducting analyses in labour and public economics (see, e.g., Gottschalk and Huynh 2010, and Meyer and Mittag 2019). Indeed, linked administrative datasets exploit the main advantages of administrative records and survey data, overcoming some of the shortcomings of the two sources. Administrative archives are consistently and accurately collected, covering a large number of observations (sometimes, collecting information on the whole reference universe) and long-time periods. However, the number of variables is often limited, because administrative data is collected for administrative purposes, rather than research purposes (e.g. educational attainments and marriage statuses may not be recorded in administrative archives). Thus, the identification of relevant drivers of labour outcomes is not always feasible. On the other hand, self-reported survey data provides detailed information on individual characteristics, but it is typically associated with higher measurement errors (Kim and Tamborini 2012). Moreover, panel surveys are usually limited either in the cross-sectional or time dimension. For instance, the EU-SILC dataset is based on a rotation scheme where individuals are followed over four years, at most.

In the present context, longitudinal information is provided by two different sources:

- Administrative archives from the INPS, providing information on the working history of individuals;
- Survey data from IT-SILC, the Italian database of the European Union Survey on Income and Living Conditions (EU-SILC), collected by ISTAT and providing information on the socio-demographic and economic status of interviewed individuals.

To build the first version of AD-SILC on the occasion of the INDIW project, detailed micro-data from the 2005 cross-sectional wave of IT-SILC was merged with information collected in several administrative archives held by INPS. The administrative records provide very granular information on several characteristics of private and public employees, self-employed workers, recipients of unemployment benefits and other social benefits and retirees. Later, the IESS project[2] expanded the AD-SILC database, updating the micro-data collected in the administrative archives to the end of 2013, and adding new variables from the administrative datasets (providing information on the so-called '*parasubordinati*'[3] and further details on contractual arrangements and type of unemployment benefit received). Moreover, the sample size increased approximately from 50 to 200 thousand individuals. While the first version of AD-SILC made use of the 2005 wave of IT-SILC only, in the updated version all individuals interviewed in the waves spanning the period 2004-2012 were included.

2 See: http://www.fondazionebrodolini.it/en/projects/iess-improving-effectiveness-social-security.

3 These workers represent a class of *atypical* workers. In 1996, INPS created a fund called *Gestione Separata*, for self-employed workers belonging to specific categories of professionals, who did not have a special social security scheme and fund, and fixed-term contract workers, other than employees (so called 'co.co.pro'and 'co.co.co'). In what follows, we shall refer to these workers as "para-subordinates".

The key features of AD-SILC concern several aspects:

- Individual working career histories (collected in the administrative archives) and socio-economic variables (e.g. educational attainment, parental education and occupation, marital status, family composition, citizenship) are merged together;
- Labour force status is registered at the individual level, using the information collected by INPS in the Registers of Active People and Retirees;
- The dataset includes variables at the firm level, providing reliable information on employers' characteristic (e.g. industry classification and firm's size, for holding companies and subsidiaries);
- By means of the longitudinal component of IT-SILC, it is possible to track changes occurred in individual social statuses (e.g. change in civil status and educational attainment, childbirths, etc.).

In this framework, AD-SILC has been employed to analyse the Italian labour market evolution over the last decades, at the individual level[4]. A special focus has been given to the dynamics of the earnings distribution, the individual transitions among occupational statuses, and the adequacy of accrued contributions, paying specific attention to the cohorts of workers enrolled in the Notional Defined Contribution (NDC) pension scheme[5]. Furthermore, AD-SILC has been crucial for the development of the Treasury Dynamic Microsimulation Model (T-DYMM), whose main goal was to assess the adequacy of the Italian pension system.

1.2 AD-SILC 3.0

For the scopes of the MOSPI project, the dataset for both the analyses and the microsimulations on T-DYMM has been updated and expanded. First, a new merge of IT-SILC (Section 1.2.1) and administrative data from INPS (Section 1.2.2) has been carried out, this time encompassing all waves in the period 2004-2017 and a wider set of variables, including information on respondents' health and migration status. A number of additional data sources (Section 1.3) has also been considered in order to enrich the information contained in AD-SILC 3.0.

4 See, amongst others: Barbieri *et al.* (2020), De Villanova *et al.* (2019), Lallo and Raitano (2018), Naticchioni *et al.* (2016), Raitano and Fantozzi (2015), Raitano and Vona (2018), Raitano and Vona (2017).

5 Legislative Decree no. 335/1995 set for the gradual replacement of the traditional Earnings Related (ER) scheme with a Notional Defined Contribution (NDC) scheme. Only workers enrolled beyond 1995 have their pension benefit entirely computed according to NDC rules, while for senior workers pro-quota computation rules are in place.

1.2.1 Structure and variables of IT-SILC

IT-SILC is the Italian database of the European Union Survey on Income and Living Conditions (EU-SILC). The EU-SILC project is coordinated by Eurostat and has been developed as an instrument to collect comparable cross-sectional and longitudinal microdata on income, poverty, social exclusion and living conditions. It provides:

- Cross-sectional data pertaining to a given time or a certain time period with variables on income, poverty, social exclusion and other living conditions;
- Longitudinal data pertaining to individual-level changes over time, observed annually over a four-year period.

An important feature of IT-SILC is the large number of survey variables. Two types of variables are provided: variables measured at household level and variables measured at individual level. Social exclusion and housing condition information is collected mainly at household level, while information on labour, education and health is obtained for individual persons aged 16 and older[6] (Eurostat 2020).

The sample design of IT-SILC is based on a two-stage procedure (Ceccarelli *et al.* 2008). First, municipalities are partitioned into auto-representative and non-auto-representative municipalities in each region. Then, households are randomly drawn. In the first group, where municipalities are characterised by larger population sizes, households are drawn from the register office records according to a systematic sampling scheme. For the latter group, composed of smaller municipalities, households are randomly drawn from a subsample of selected municipalities, according to a multistage sampling scheme.

In line with Eurostat directives, the scheme of IT-SILC envisages two components: a cross-sectional and a longitudinal component. In particular, a rotated sampling design is implemented (Duncan *et al.* 1989): a new panel of households and individuals is introduced each year to replace one fourth of the existing sample. Table 1.1 illustrates the four-year rotation scheme adopted in IT-SILC. For instance, in 2007, panel D was interviewed for the fourth time (D4), panel E for the third time (E3), panel F for the second time (F2), and panel G is introduced for the first time (G1).

The first round of the survey, taken in 2004, was composed of 32,000 households, around 8,000 households belonging to each longitudinal sample. In the second wave, relative to 2005, the sample comprises ¾ of the households interviewed in 2004 (corresponding to B4, C3 and D2 in Table 1.1) and 8,000 newly selected households, belonging to the new longitudinal panel (E1 in Table 1.1). The same rotation pattern was followed until the last available round of the survey in 2017. Overall, 255,465 people were interviewed between 2004 and 2017, corresponding to 703,198 individual records.

---

6 To provide comparability with other Italian surveys, the Italian version of EU-SILC (IT-SILC) also includes 15-year-old individuals.

Table 1.1 The rotational sampling scheme of IT-SILC

Year	Panel samples																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
2004	A4	B3	C2	D1													
2005		B4	C3	D2	E1												
2006			C4	D3	E2	F1											
2007				D4	E3	F2	G1										
2008					E4	F3	G2	H1									
2009						F4	G3	H2	I1								
2010							G4	H3	I2	J1							
2011								H4	I3	J2	K1						
2012									I4	J3	K2	L1					
2013										J4	K3	L2	M1				
2014											K4	L3	M2	N1			
2015												L4	M3	N2	O1		
2016													M4	N3	O2	P1	
2017														N4	O3	P2	Q1

At each wave *t*, the data refers to two different time periods: some variables refer to the same year (time *t*), others (e.g. income) to the previous year (*t*-1).
Graf *et al.* (2011) provide a summary list for the main target variables of EU-SILC (i.e. variables that are collected every year)[7], which can be organised according to the following criteria:

- Kind of data:

A. If the variable is a household variable, it is either:
   1. Basic data (basic household data including degree of urbanisation);
   2. Income (total household income and gross income components at household level);
   3. Social exclusion (non-monetary household deprivation indicators, including problems in making ends meet, extent of debt and enforced lack of basic amenities);

[7] In addition to the harmonised variables, ISTAT adds a number of additional inquiries to the IT-SILC survey. Until 2015, the dataset consisted of four components, all including harmonized variables (at the European level) as well as variables specific to the Italian survey. Since 2016, Italy-specific variables have been provided in two additional files; thus, six different files are provided in total today.

    4. Labour information (childcare);
    5. Housing (dwelling type, tenure status and housing conditions, amenities in dwelling and housing costs).
B. If the variable is an individual variable, it is either:
    1. Basic data (basic personal data and demographic data);
    2. Education (educational attainment, years of schooling, current enrolment);
    3. Labour information (basic labour information on current activity status and on current main job, including information on last main job for the unemployed, basic information on activity status during income reference period, total number of hours worked on current second/third... job, detailed labour information, activity history and calendar of activities);
    4. Health (health status and chronic illness or condition, access to healthcare);
    5. Income (net and gross personal income, total and components at personal level).

- Type of variable: cross-sectional or longitudinal.
- Reference period: constant, current, income reference period, last twelve months, since last year, working life and childcare reference period.
- Unit: household, household member, former household member, selected respondent and household member aged 16 (15 for IT-SILC) and over.

Besides the ground information, for every year after 2004 the survey has been devoted to a particular topic and set to collect secondary target variables according to said topic. Table 1.2 reports the modules on which secondary target variables were collected in the period 2005-2017.

The main functions of AD-SILC include allowing the preliminary analyses that render the development of T-DYMM possible as well as operating as starting dataset for the microsimulations.

IT-SILC provides a wide range of variables about individual backgrounds. The level of detail provided in month-to-month transitions and in the main activity carried out by workers, for instance, can be fruitfully confronted with the information contained in the administrative data to obtain a clear identikit of atypical workers, who may incur in a number of (involuntary) changes in working status within a short time span. Information on social insurance contributions is also crucial. IT-SILC provides valuable information on supplementary pension plans, as the questionnaire asks respondents to indicate whether they have paid voluntary contributions to banks, insurance and financial institutions for private pension plans, and if so, in what amount.

An important extension to the previous version of the dataset is the inclusion of variables pertaining to health condition and migration status of respondents (see Annexes 1.1 and 1.2).

Table 1.2 SILC modules for secondary variables, 2005-2017

Year	Module
2005	Intergenerational transmission of poverty
2006	Social participation
2007	Housing conditions
2008	Over-indebtedness and financial exclusion
2009	Material deprivation
2010	Intra-household sharing of resources
2011	Intergenerational transmission of disadvantages
2012	Housing conditions
2013	Well-being
2014	Material deprivation
2015	Social and cultural participation and material deprivation
2016	Access to services
2017	Health and children's health

Source: Eurostat

1.2.2 Structure and variables of INPS administrative archives

To build AD-SILC, we make use of several administrative datasets, provided by INPS and updated to 2019 in the occasion of the MOSPI project. All individuals surveyed at least once in IT-SILC in the 2004-2017 waves are linked to all available information stemming from the following archives:

- EC_INPS (*Register of Active Workers*): it collects information on workers paying contributions to INPS funds, i.e. private employees, farmers, dealers, artisans, atypical workers and professionals without a private fund managed by their association. The units of observation in EC_INPS are the individuals' single working records in a given year. In practice, individuals may present more than one record per year (that is the case if, for a given year, they experience more than one employment status or receive more than one type of substitute contribution, e.g. stemming from unemployment benefits, maternity or sickness allowances, etc.). Each year, the following variables are recorded: age; year of birth; date of death; sex; province and nation of birth; starting and ending date of every working record or period of substitute contribution; number of weekly contributions of each working record or number of weeks in which benefits/allowances are paid; tax code of the firm (for private employees); province where the individual works; gross wage (i.e. including workers' contributions) or amount of welfare benefit; specific INPS funds where the individual pays his contributions (that allows to identify private employees, atypical workers, farmers, artisans and dealers); type of contributions (that allows to identify working periods and periods of unemployment, sickness allowances and maternity allowances); occupational classification referring to employees, that allows to distinguish managers, white-collars, blue-collars, apprentices and, since 1998, amongst part-time/full-time workers and fixed-term/open-ended employment arrangements.

- AZ_INPS: it collects information on firms' characteristic. This dataset is merged to EC_INPS by means of the tax codes of the firms. The following variables are recorded: industry classification of the firm (ATECO three-digits); business structure of the firm (distinguishing among single firms and companies with a parent/subsidiary relation); number of workers of the holding and number of workers of the subsidiary.
- PENSIONI (*Register of Retirees*): it collects information on people receiving pension benefits. The following variables are recorded: sex; age; date of death; date of retirement; monthly gross amount of the pension benefit; region of residence; professional status before retirement (e.g. employee or self-employed); seniority at retirement (weeks of contributions accrued); type of pension benefit (which allows to distinguish among old-age/early-retirement, survivor, invalidity and social pensions).

Table 1.3 lists the number of records per year provided by the Register of Active Workers and the Register of Retirees[8], for all individuals included in AD-SILC. The total number of observations is equal to 8,475,883 and the vast majority of them (7,153,807 records, approximately 85%) stem from the Register of Active Workers contributory records. Around 45% of the observations belong to the 2004-2017 interval, covered by SILC survey information. Many individuals (32.4%) report more than one record per year: they hold multiple contributory relationships with INPS, either because they have multiple jobs, or because they are both employed and receive some type of benefit within the same year.

Therefore, the total number of individuals recorded in all years is smaller, amounting to 5,468,192 (Table 1.4).

[8] A small amount of records (less than 0.01 percent) was dropped from the sample, due to misreporting of the reference period.

Table 1.3 Number of records in AD-SILC by year

Year	Units	%	Year	Units	%	Year	Units	%
1901	60	0.00	**1941**	300	0.00	**1981**	103,588	1.22
1902	59	0.00	**1942**	388	0.00	**1982**	106,807	1.26
1903	59	0.00	**1943**	554	0.01	**1983**	106,100	1.25
1904	59	0.00	**1944**	591	0.01	**1984**	106,876	1.26
1905	59	0.00	**1945**	663	0.01	**1985**	111,339	1.31
1906	59	0.00	**1946**	773	0.01	**1986**	110,950	1.31
1907	59	0.00	**1947**	964	0.01	**1987**	112,982	1.33
1908	59	0.00	**1948**	1,214	0.01	**1988**	117,038	1.38
1909	59	0.00	**1949**	1,568	0.02	**1989**	120,333	1.42
1910	62	0.00	**1950**	2,171	0.03	**1990**	143,049	1.69
1911	65	0.00	**1951**	2,853	0.03	**1991**	126,037	1.49
1912	59	0.00	**1952**	4,452	0.05	**1992**	127,202	1.50
1913	59	0.00	**1953**	5,078	0.06	**1993**	124,031	1.46
1914	59	0.00	**1954**	6,600	0.08	**1994**	123,820	1.46
1915	59	0.00	**1955**	8,650	0.10	**1995**	123,029	1.45
1916	59	0.00	**1956**	10,588	0.12	**1996**	128,120	1.51
1917	59	0.00	**1957**	19,077	0.23	**1997**	128,661	1.52
1918	59	0.00	**1958**	21,900	0.26	**1998**	130,267	1.54
1919	59	0.00	**1959**	25,230	0.30	**1999**	133,243	1.57
1920	59	0.00	**1960**	28,128	0.33	**2000**	138,968	1.64
1921	59	0.00	**1961**	31,537	0.37	**2001**	143,715	1.70
1922	59	0.00	**1962**	34,082	0.40	**2002**	147,938	1.75
1923	59	0.00	**1963**	37,256	0.44	**2003**	148,905	1.76
1924	59	0.00	**1964**	38,398	0.45	**2004**	232,325	2.74
1925	59	0.00	**1965**	41,903	0.49	**2005**	258,038	3.04
1926	59	0.00	**1966**	44,627	0.53	**2006**	264,512	3.12
1927	60	0.00	**1967**	47,853	0.56	**2007**	276,607	3.26
1928	60	0.00	**1968**	51,003	0.60	**2008**	277,843	3.28
1929	59	0.00	**1969**	53,786	0.63	**2009**	285,651	3.37
1930	60	0.00	**1970**	56,222	0.66	**2010**	286,075	3.38
1931	838	0.01	**1971**	58,827	0.69	**2011**	284,645	3.36
1932	64	0.00	**1972**	65,972	0.78	**2012**	289,183	3.41
1933	67	0.00	**1973**	71,248	0.84	**2013**	281,696	3.32
1934	75	0.00	**1974**	85,440	1.01	**2014**	281,748	3.32
1935	75	0.00	**1975**	82,490	0.97	**2015**	286,549	3.38
1936	86	0.00	**1976**	87,923	1.04	**2016**	285,544	3.37
1937	100	0.00	**1977**	88,241	1.04	**2017**	285,518	3.37
1938	124	0.00	**1978**	92,034	1.09	**2018**	182,343	2.15
1939	174	0.00	**1979**	96,040	1.13	**2019**	144,364	1.70
1940	217	0.00	**1980**	99,990	1.18	**2020**	27	0.00
						Total	8,475,883	100.00

Source: own elaborations from AD-SILC data

Table 1.4 Number of recorded individuals in AD-SILC by year

Year	Number	%	Year	Number	%	Year	Number	%
1901	23	0.00	**1941**	172	0.00	**1981**	73,072	1.34
1902	24	0.00	**1942**	221	0.00	**1982**	73,630	1.35
1903	24	0.00	**1943**	297	0.01	**1983**	74,178	1.36
1904	24	0.00	**1944**	357	0.01	**1984**	74,245	1.36
1905	24	0.00	**1945**	405	0.01	**1985**	75,715	1.38
1906	24	0.00	**1946**	499	0.01	**1986**	77,092	1.41
1907	24	0.00	**1947**	650	0.01	**1987**	79,427	1.45
1908	24	0.00	**1948**	804	0.01	**1988**	81,703	1.49
1909	24	0.00	**1949**	1121	0.02	**1989**	83,122	1.52
1910	27	0.00	**1950**	1539	0.03	**1990**	84,401	1.54
1911	30	0.00	**1951**	2056	0.04	**1991**	85,691	1.57
1912	24	0.00	**1952**	2766	0.05	**1992**	86,200	1.58
1913	24	0.00	**1953**	3606	0.07	**1993**	84,369	1.54
1914	24	0.00	**1954**	4660	0.09	**1994**	84,137	1.54
1915	24	0.00	**1955**	5867	0.11	**1995**	84,661	1.55
1916	24	0.00	**1956**	7,237	0.13	**1996**	87,590	1.60
1917	24	0.00	**1957**	14,051	0.26	**1997**	88,726	1.62
1918	24	0.00	**1958**	15,795	0.29	**1998**	90,171	1.65
1919	24	0.00	**1959**	18,117	0.33	**1999**	92,241	1.69
1920	24	0.00	**1960**	20,006	0.37	**2000**	94,683	1.73
1921	24	0.00	**1961**	22,266	0.41	**2001**	96,860	1.77
1922	24	0.00	**1962**	24,442	0.45	**2002**	99,649	1.82
1923	24	0.00	**1963**	26,155	0.48	**2003**	100,956	1.85
1924	24	0.00	**1964**	27,132	0.50	**2004**	153,324	2.80
1925	24	0.00	**1965**	29,678	0.54	**2005**	156,160	2.86
1926	24	0.00	**1966**	31,783	0.58	**2006**	159,647	2.92
1927	25	0.00	**1967**	33,820	0.62	**2007**	163,899	3.00
1928	25	0.00	**1968**	35,810	0.65	**2008**	166,447	3.04
1929	24	0.00	**1969**	37,955	0.69	**2009**	167,149	3.06
1930	25	0.00	**1970**	40,140	0.73	**2010**	168,234	3.08
1931	51	0.00	**1971**	43,038	0.79	**2011**	168,841	3.09
1932	27	0.00	**1972**	47,340	0.87	**2012**	168,912	3.09
1933	30	0.00	**1973**	51,106	0.93	**2013**	166,729	3.05
1934	34	0.00	**1974**	57,572	1.05	**2014**	166,350	3.04
1935	33	0.00	**1975**	59,526	1.09	**2015**	167,195	3.06
1936	41	0.00	**1976**	62,040	1.13	**2016**	168,155	3.07
1937	50	0.00	**1977**	63,470	1.16	**2017**	167,776	3.07
1938	61	0.00	**1978**	66,758	1.22	**2018**	105,276	1.93
1939	89	0.00	**1979**	69,007	1.26	**2019**	99,162	1.81
1940	124	0.00	**1980**	71,859	1.31	**2020**	23	0.00
						Total	5,468,192	100.00

Source: own elaborations from AD-SILC data

1.3 Additional data sources

IT-SILC survey data and INPS administrative archives have been employed ever since the construction of the first AD-SILC dataset. The first objective of the MOSPI project was to update the IT-SILC and INPS data, and by doing so we have expanded the number of variables at our disposal and enlarged the sample. In the present paragraph, we shall explore the additional data sources considered for a further expansion of AD-SILC 3.0.

1.3.1 Administrative data on Tax Declarations and Cadastral Records (Department of Finance)

The Department of Finance of the Italian Ministry of Economy and Finance holds all national data on annual Tax Returns[9] and all cadastral records[10]. As a result of recent interactions, the Finance Department has agreed to share data relative to the individuals surveyed in SILC waves 2010, 2012, 2014, 2016 and 2018, with the information pertaining to the previous tax year.
Despite the fact that the amount of information shared will not allow us to build a complete panel structure (not all the SILC waves are covered), the informational gain will be very high.
From the Tax Returns data, we shall obtain: i) Gross income subject to Personal Income Tax (PIT), including labour income, pension benefits, cadastral income, capital income and gains; ii) PIT tax credits and deductions; iii) Gross income subject to proportional taxation (e.g. rental income subject to *cedolare secca*, productivity bonuses, self-employment income subject to substitute tax regimes such as *regime fiscal di vantaggio* and *regime forfetario*).
From the Cadastral Records, we shall obtain: i) Municipality where the real estate unit is located; ii) Cadastral category; iii) Surface in squared meters (only for residential dwelling units); iv) Cadastral value; v) Share of possession. In addition to this data from the Cadastre, the Department of Finance will also provide a match with the market value of the residential dwelling units according to the Observatory of the Real Estate Market (*Osservatorio del Mercato Immobiliare,* OMI)[11], managed by the Italian Revenue Agency (*Agenzia delle Entrate*).
The potentialities of the inclusion of the listed information in the development of the model and in the analyses of labour market dynamics are under evaluation. The level of detail of the data on income sources from the Tax Returns can be useful in better identifying a number of categories of Non-Standard Workers (casual workers,

---

9 For official statistics and methodological information, see: https://bit.ly/3cxwURV.
10 For official statistics and methodological information, see: https://bit.ly/2MsK48j.
11 See: https://bit.ly/36XPQYU.

self-employed under substitute tax regimes), while the data on real estate shall provide a significant contribution to the development of the Wealth Module (see further). Furthermore, these data sources will be of great value for a more precise identification of households and/or individuals receiving means-tested benefits, where both income and wealth information are used for their computation.

1.3.2 The Survey on Household Income and Wealth (SHIW)

The Survey on Household Income and Wealth (SHIW) is a household survey conducted every two years from 1977 to 2018 (with a three-year gap in the period 1995-1998) by the Bank of Italy. The main objective of the survey is to study the economic behaviour of Italian households, defined as groups of individuals related by blood, marriage or adoption and sharing the same dwelling. The sample size comprises about 8,000 households per year, drawn from population registers. The survey contains a sizable panel component, which allows for the estimation of target variables' processes and transitions. The head of the household is the person responsible for the household finance, he is the main earner in the family and is labelled with an order number equal to one (NORD=1). The longitudinal component allows following potentially over 50% of the households in two spells of twice-repeated observations.

Data collection is entrusted to a specialized company using professional interviewers and CAPI (Computer Assisted Personal Interview) methodology. The survey collects the following information:

- Characteristics of the household and its members (number of income earners, gender, age, education, job status, industry sector, and characteristics of the dwelling);
- Income (wage and salaries, income from self-employment, pensions and other financial transfers, income from financial assets and real estate);
- Consumption and saving (food consumption, other nondurables, expenses for housing, health, insurance, spending on durable goods, and household saving);
- Wealth in terms of real estate, financial assets, liabilities;
- Special modules such as capital gains, inheritance, risk aversion, unpaid work, economic mobility, social capital, tax evasion, financial literacy.

For the scope of the MOSPI project, we are going to use 2004-2016 SHIW waves (in line with the time span of the reference AD-SILC 3.0 dataset) in order to integrate AD-SILC with household net wealth[12] and consumption/savings information, whenever they are not available from an administrative source.

The underreporting of financial assets constitutes one of the issues of this data source, especially with respect to financial variables. Several works have addressed this issue, in

[12] In the form of real estate (houses), financial assets, mortgages, supplementary pension schemes, plus information on financial literacy.

particular D'Aurizio *et al.* (2006). The authors adjust wealth data by matching the 2002 SHIW wave with anonymous data from a sample survey of customers of a private bank (Unicredit Bank) on the assets actually owned. Our aim is to apply the same correction to the data, combining it with another form of correction proposed by Brandolini *et al.* (2009) that helps to reduce the distortions in terms of ownership of financial assets. This combined procedure follows what Boscolo (2019) has recently proposed. The link between the SHIW dataset and AD-SILC will be carried out through a merging technique at the household level. Our aim is to adopt the Propensity Score Matching technique following Tedeschi *et al.* (2013), who merged SHIW with the Italian *Indagine sui Consumi delle Famiglie* (Household Budget Survey, HBS). One crucial variable to link the data is the household labour income that in the last available wave of SHIW referring to 2016, this implying a perfect overlapping with the tax year of the income information contained in IT-SILC. In this procedure, we will conceive AD-SILC 3.0 as the recipient sample and SHIW as the 'donor' of some missing information.
The imported survey information on real estate will be compared and adjusted with administrative data such as tax and cadastral records provided by the Finance Department as well as listing of real estate values from OMI.

1.3.3 Survey data from INAPP

The National Institute for Public Policy Analysis (*Istituto Nazionale per l'Analisi delle Politiche Pubbliche*, INAPP) carries out a number of surveys directed to analyses on the labour market, on both sides of demand and supply. Two of them seem promising for our aim: the Survey on Labour Participation and Unemployment (PLUS), a sample survey on the Italian labour market supply and (to a lesser extent) the Longitudinal Survey on Enterprises and Labour (RIL), which collects information from enterprises. As noted in 'Background report on future of work scenario: The dynamics of non-standard work in Italy' (MOSPI first intermediate report by INAPP) notes, the primary objective of the PLUS survey is to provide statistically reliable estimates of phenomena that are either rare or marginally explored by other surveys concerning the Italian labour market. The survey was collected in 2018 on a sample of about 45,000 individuals. It dedicates a special focus on platform workers (divided into labour platforms, capital platforms, and online sale of consumer goods) and reports detailed information on education, age, occupational status, relevance of platform-related earnings, type of contract (for labour platforms only) and ability to deal with unexpected expenses (medical treatment and dental care).
Even though the survey is only available for the year 2018, this data may contribute to extend the information on non-standard workers already at our disposal through data from INPS and from Tax Returns, thereby allowing for a more accurate depiction of work trajectories and therefore welfare-specific issues. A correct stylisation of the

niche phenomenon of platform workers may also allow us to implement scenarios and elaborate targeted policy proposals to test in T-DYMM.
Similarly, information on automation and investments on digitalisation contained in both the PLUS and RIL surveys can be helpful in defining future of work scenarios to assess the effect of the digitalisation process on specific sectors and categories of workers.
Because SILC and INAPP's surveys are conducted on different samples (and since administrative data on the same topic is not available), the information cannot be directly linked to AD-SILC. Nonetheless, the vast information AD-SILC already holds could allow the use of statistical matching techniques to associate its observations to variables reported in external surveys for cases that display similar features to the ones in AD-SILC. Techniques similar to the ones mentioned above for SHIW may also be employed here.

1.3.4 Possible further expansions of the dataset

As noted above, the overall data employed within the MOSPI project has a multitude of sources. All of the gathered information can be essentially divided into two categories: i) data that are directly linkable to individual observations via the tax code key (IT-SILC, INPS and Finance Department data); ii) survey data that do not allow for a direct link (SHIW and INAPP's surveys).
Since the Treasury Department has access to the tax codes pertaining to IT-SILC surveyed individuals, upon agreements with the pertinent Institutions, the first category may be further expanded in the future. The Revenue Agency has recently developed an archive that gathers information on account balances and financial operations[13]. The database is under construction and its use is restricted and regulated by law. Nonetheless, we have opened a channel for potential dialogue of AD-SILC with the Revenue Agency data on financial operations in the future. Tax codes may also be linked to administrative data on health-related information (e.g. health-related expenditure, Hospital Discharge Cards), thereby allowing for analyses on the socio-economic determinants of health status and for the development of a Health Module within T-DYMM.

[13] Decree Law 201/2011 (so-called *Salva Italia*) introduced the obligation for financial operators to communicate to the Tax Registry – called the Archive of relations with financial operators – information on the balances and movements of active relationships. The communication – made through the SID infrastructure – is accompanied by that relating to the Registry of financial relations, regulated by the provisions of 19 January 2007 and 29 February 2008. Since January 2016, communications relating to monthly information (openings and terminations of relationships) and annual information (balances, movements and other accounting data) take place on the basis of the layout and technical specifications established in the same month through the SID infrastructure on the basis of the layout and technical specifications established in the same month.

1.4 Data treatment and preliminary analyses

In the present Paragraph, we shall devote our attention to the procedures carried out to collect the data that composes the AD-SILC dataset (1.4.1), to merge it (1.4.2) and to address concerns related to drop-outs (1.4.3).

1.4.1 Data collection procedure

The AD-SILC database comprises a great deal of information subject to protection under the legislation on the protection of citizens' personal data. In May 2018, the European General Data Protection Regulation (EU) 2016/679 (GDPR) was applied in all European countries. In compliance with the Regulation, within the MOSPI project we have followed requirements and instructions for the development of services that require the processing of personal data, limiting the risks for the rights and freedoms of all interested parties.

The data transfer saw a triangulation of efforts from ISTAT, the Department of the Treasury (DT) and INPS. Following the procedure outlined within the National Statistical System (SISTAN), ISTAT shared with DT all IT-SILC data pertaining to the waves 2004-2017 and the tax codes relative to all surveyed individuals. DT subsequently shared with INPS the tax codes to allow for the merge of administrative data relating to workers, pensioners and firms. Upon delivery from INPS of the administrative information, DT operated the 1:1 merge between IT-SILC and INPS data. DT then stored the tax codes in a dedicated server area with heavily restricted access, while in the database used for the analyses anonym identification codes substituted tax codes. More recently, DT has approached the Department of Finance (DF) of the Ministry of Economy and Finance concerning the possibility of getting access to the data on individual income tax returns and the real estate registry, which the DF holds. Individuals from IT-SILC waves would be linked to that information with a 1:1 merge. Because DT and DF (Departments of the Italian Ministry of Economy and Finance) hold autonomous positions within the SISTAN, the procedure to apply is identical to the one followed for ISTAT and INPS.

As all institutions involved in the merging process (ISTAT, INPS, DT and DF) are part of SISTAN, works have been launched for the institution of a specific project within the SISTAN platform that would entail the periodic update of the dataset. Such a framework would ensure the reliability of the data provided, as SISTAN projects are subject to high quality standards, as well as set the ground for a routine use of AD-SILC and T-DYMM not only for periodical research, but also as consolidated policy tools.

For what concerns the additional data taken into consideration, data related to the Survey on Household Income and Wealth (SHIW) is public and available on the website of the Bank of Italy, while the PLUS, QDL and RIL surveys are available upon request to INAPP's statistical service.

1.4.2 Merging the data: checks for representativeness

As discussed above, the AD-SILC dataset is made up of two main data sources. On the one hand, we have the IT-SILC cross-sectional waves for the 2004-2017 period. On the other, we have individuals' working histories and pension accrued from INPS archives. Table 1.5 reports the number of IT-SILC's observations with tax code and observed working history. It is worth noting that we do not have the respective tax codes for all surveyed individuals[14]. The percentage of observations with tax code ranges in the interval [92.9%, 98.8%], and increases over time in relative terms.

Table 1.5 From IT-SILC to AD-SILC: merged observations

YEAR	IT-SILC (sample size)	%	IT-SILC with tax code	%	IT-SILC ↔ INPS (via tax code)	%
2004	61,542	100.0	57,160	92.9	52,248	84.9
2005	56,105	100.0	53,259	94.9	48,346	86.2
2006	54,512	100.0	51,494	94.5	46,269	84.9
2007	52,772	100.0	50,515	95.7	44,970	85.2
2008	52,433	100.0	50,110	95.6	44,123	84.2
2009	51,196	100.0	48,946	95.6	42,600	83.2
2010	47,551	100.0	45,737	96.2	39,515	83.1
2011	47,841	100.0	45,450	95.0	38,981	81.5
2012	47,365	100.0	45,100	95.2	38,385	81.0
2013	44,622	100.0	43,264	97.0	36,358	81.5
2014	47,136	100.0	46,508	98.7	38,698	82.1
2015	42,987	100.0	41,838	97.3	34,423	80.1
2016	48,316	100.0	47,704	98.7	39,062	80.8
2017	48,819	100.0	48,257	98.8	39,471	80.9
TOTAL	703,197	100.0	675,342	96.0	583,449	83.0

Source: own elaborations from AD-SILC data

[14] Tax codes are reported as missing when they are not validated by SOGEI (Società Generale Informatica, the State-owned IT company in charge of assigning tax codes).

Each IT-SILC wave was merged with data from INPS archives using the tax code as key. The percentage of IT-SILC surveyed individuals with at least one contributory record diminishes over time, ranging within the interval [80.1%, 86.2%]. This can be explained by the fact that INPS information exceeds the year of the survey. Therefore, young individuals surveyed in earlier IT-SILC waves have had years to grow and enter the labour market by the last year for which INPS archives deliver information (2019), while young cohorts surveyed in 2017 are less likely to have accrued any contribution in the following two years.
Focusing on the Register of Active Workers (see Paragraph 1.2.2, 'Structure and variables of INPS administrative archives') a few adjustments were required due to data inconsistency. As explained above, INPS data keeps track of the entire working careers of individuals. Every record refers to the date of beginning and end of the individual contract, which is recorded on a yearly basis. The Register of Active Workers consists of 7153807 records referring to 204779 individuals.
First, implausible data was excluded from the dataset, such as observations with: i) incorrect years (i.e. beginning in 1900s); ii) ending date of the relative contract that refers to years later than 2020; iii) imputation errors in the starting/ending date of the contract (i.e. the 31st of November); iv) ending date of the contract recorded before the starting one; v) death date prior to new contracts or the IT-SILC interviews; vi) birth date preceding the starting one of the contract; and vii) working contracts referring to years where individuals were younger than 16 years old. Furthermore, we excluded from the dataset all the contributions accrued abroad.
As a result of the data cleaning process, 82,805 observations (1.2% of the original dataset) were dropped from the Register of Active Workers, for a final dataset of 7071002 records.

1.4.3 Informative Drop-out

In this paragraph, we perform a sensitivity analysis in order to check whether the AD-SILC dataset presents issues related to non-ignorable or Informative Drop-out (ID) based on the observed individual features. The analysis has been concentrated in two steps.
As previously specified, the IT-SILC and INPS datasets were linked by the individual tax code. Focusing on the universe of respondents to IT-SILC in the 2004-2017 period (only P-files)[15], we do not dispose of the tax code of 18,915 observations out

[15] Each IT-SILC cross-sectional wave from 2004 to 2015 is made up of four different sub-files (D-, R-, H- and P-files), where data harmonized at the European level is gathered together with integrative Italian variables. From 2016 onwards, these latter variables have been separated from the remaining ones, for a total number of seven sub-files (D-, R-, H- and P-files; QI-, QF- and RF-files). P-files refer to personal data of individuals with more than 14 years of age at the income reference year.

of 601,082 (3.15%). As a result, these respondents were excluded from a number of analyses on the AD-SILC dataset.
The aim of the first analysis it to establish whether this exclusion affects the representativeness of our sample, as well as to verify to what extent the excluded observations differ from the included ones with respect to the observed covariates.
A second analysis checks for coherence between IT-SILC respondents and INPS records. Out of 601,082 IT-SILC observations, we were able to distinguish 222,130 respondents[16]. Then, individuals were matched to the respective working careers collected in INPS archives (both the Register of Active Workers and the Register of Retirees). Out of 222,130 individuals, 216,357 workers were successfully matched (the unmatching is due to inactive individuals). Among the 177,121 individuals who declared to have worked at least once according to IT-SILC data, 2,883 were found to have no records in INPS archives[17]. Within this subsample, as was done in the case of missing tax codes, we tested for differences in the observed covariates.
In the two analyses, we have fitted a logistic regression model for the drop-out process to evaluate the probability of informative drop-out, depending on observed covariates. Parameters in the model are estimated by maximum likelihood and inferences drawn through conventional likelihood procedures[18].
To select the best model and perform model validation, we randomly partitioned the original sub-sample[19]. The original dataset was divided into two subsets: training and testing. The training set was used for model selection and to fit the final prediction models; the test set was used to evaluate the performance (in terms of prediction accuracy) of the final model. The prediction accuracy of models was assessed by using the Brier score, which quantifies the distance between predictions and the corresponding estimated probability.
The Brier score is a metric that combines discrimination (where accurate predictions discriminate between observations with and without the outcome) and calibration (the agreement between observed outcomes and predictions)[20]. Therefore, the model with the smaller value of the score was chosen.
In both the analyses, we included the following independent variables: age, gender, income, self-reported activity status, general self-reported health status, civil status and educational attainment. Based on the results of the Brier score, the selected (best) model was performed after model validation.

16 On average, individuals are followed for approximately three IT-SILC waves in our sample.
17 Three (main) reasons can explain why missing matches occur: i) undeclared work; ii) misunderstanding of the question; iii) incorrect linkage between the two data sources.
18 See: Diggle and Kenward (1994).
19 See: Zhang *et al.* (2018).
20 See: Brier (1950).

In the first analysis (Table 1.6), the dependent variable indicates whether we dispose of the respondent's tax code in P-files (value 0) or not (value 1). As a result of the analysis, the magnitude of the effect we have observed looking at the covariates is not negligible. However, our analysis deals with a large sample. In these cases, there is a tendency to reject the null hypothesis[21] even in the case of small differences[22]. This means that standard errors estimated in large datasets can be biased. In order to estimate the significant effect of each variable, instead of observing the p-value and determining the result in the usual way (e.g. significant if $p < 0.05$), we followed a Bootstrap approach (Efron 1979). The SILC dataset has been treated as a population and our estimates have been based on the distribution of regression coefficients[23] related to 500 samples with size of 10,000, drawn from SILC population (Table 1.7). We have also considered a different number of samples combined with different sample sizes, obtaining quite stable results.

As can be observed in Table 1.7, only a few variables are now significant. In the case of years, this can be interpreted as the improvement in the accuracy of the estimates over time. Being a student is also significant. This is in line with expectations and is probably because respondents (in SILC) have a better level of education if compared with the rest of the population.

The second analysis aims at investigating coherence between IT-SILC and INPS data. The results are reported in Table 1.8, where the final model was selected by looking at the Brier score. The magnitude of the effects is relatively small and most of the variables were dropped out through cross-validation techniques. We found a moderate (negative) effect of the gender category *male*, probably due to the higher presence of undeclared work among domestic (female) workers. However, the magnitude of observed significant effects is still small enough to ensure that the use of the AD-SILC dataset will not lead to biased results.

[21] Referring here to the t-test for βj to decide whether Xj has a significant effect on Y in the population, with all other X variables held constant.

[22] See: Gneiting and Raftery (2007).

[23] A logistic regression analysis, GLM.

Table 1.6 Logistic regression Y=1 if fiscal code is missing

	HR	HR low	HR up	P-value
(Intercept)	0.131	0.118	0.145	0.000
Year=2005	0.656	0.619	0.694	0.000
Year=2006	0.727	0.687	0.769	0.000
Year=2007	0.535	0.502	0.569	0.000
Year=2008	0.526	0.493	0.559	0.000
Year=2009	0.511	0.479	0.544	0.000
Year=2010	0.461	0.431	0.493	0.000
Year=2011	0.603	0.566	0.642	0.000
Year=2012	0.569	0.534	0.607	0.000
Year=2013	0.365	0.338	0.394	0.000
Year=2014	0.131	0.117	0.146	0.000
Year=2015	0.289	0.266	0.315	0.000
Year=2016	0.087	0.076	0.099	0.000
Year=2017	0.110	0.097	0.124	0.000
Gender: man	0.891	0.862	0.920	0.000
Age	0.988	0.987	0.990	0.000
Income	1.000	1.000	1.000	0.000
Self economic status (ref. Worker)				
Disable	1.042	0.908	1.190	0.551
Inactive	1.082	1.027	1.139	0.003
Retired	0.863	0.812	0.918	0.000
Student	0.576	0.537	0.618	0.000
Unemployed	1.092	1.027	1.160	0.005
General health (ref. Bad)				
Fair	0.965	0.904	1.031	0.288
Good	1.068	1.000	1.141	0.051
Very bad	1.126	0.996	1.271	0.056
Very good	1.060	0.983	1.144	0.132
Civil status (ref. Single)				
Separated/divorced	1.374	1.280	1.472	0.000
Widowed	1.198	1.113	1.288	0.000
Married	0.871	0.836	0.907	0.000
Educational attainment (ref. Primary school)				
High school	1.036	1.001	1.073	0.047
University	1.342	1.278	1.409	0.000

Source: own elaborations from AD-SILC data

Table 1.7 Bootstrap approach. Results based on the distribution of regression coefficients (logistic regression analysis, GLM) based on 500 samples with size of 10,000, drawn from the SILC population

	Analysis of Maximum Likelihood Estimate	Odds Ratio Estimates		
		HR	HR low	HR up
(Intercept)	-1.490	-2.820	-0.420	0.230
Year=2005	-0.440	-0.900	0.010	0.650
Year=2006	-0.340	-0.780	0.100	0.710
Year=2007	-0.680	-1.140	-0.190	0.510
Year=2008	-0.640	-1.250	-0.190	0.530
Year=2009	-0.690	-1.230	-0.210	0.500
Year=2010	-0.790	-1.330	-0.280	0.450
Year=2011	-0.530	-1.020	-0.040	0.590
Year=2012	-0.570	-1.130	-0.090	0.570
Year=2013	-1.030	-1.710	-0.500	0.360
Year=2014	-2.060	-3.060	-1.350	0.130
Year=2015	-1.260	-2.050	-0.650	0.280
Year=2016	-2.540	-3.980	-1.650	0.080
Year=2017	-2.290	-3.760	-1.480	0.100
Gender: female	0.130	-0.130	0.380	1.140
Age	-0.050	-0.090	0.010	0.950
Income	0.000	0.000	0.000	1.000
Self economic status (ref. Worker)				
Disable	0.020	-1.550	0.970	1.020
Inactive	0.050	-0.400	0.440	1.050
Retired	-0.240	-0.730	0.190	0.790
Student	-0.720	-1.370	-0.170	0.490
Unemployed	0.050	-0.460	0.500	1.050
General health (ref. Bad)				
Fair	0.030	-0.430	0.530	1.030
Good	0.130	-0.310	0.680	1.130
Very bad	0.080	-1.160	0.950	1.080
Very good	0.100	-0.460	0.690	1.100
Civil status (ref. Single)				
Separated/divorced	0.390	-0.280	0.940	1.470
Widowed	0.160	-0.540	0.750	1.170
Married	-0.060	-0.410	0.340	0.940
Educational attainment (ref. Primary school)				
High school	0.030	-0.230	0.290	1.030
University	0.320	-0.100	0.680	1.370

Source: own elaborations from AD-SILC data

Table 1.8 **Logistic regression analysis (GLM) Y=1 if individual in P file has ever worked but there is no trace of fiscal contribution**

	HR	HR low	HR up	P-value
(Intercept)	0.120	0.088	0.164	0.000
Year=2005	0.821	0.688	0.980	0.029
Year=2006	0.800	0.654	0.976	0.028
Year=2007	0.826	0.673	1.013	0.067
Year=2008	1.035	0.848	1.261	0.736
Year=2009	0.956	0.778	1.170	0.662
Year=2010	0.971	0.792	1.189	0.779
Year=2011	0.976	0.794	1.198	0.819
Year=2012	0.972	0.790	1.193	0.789
Year=2013	0.762	0.607	0.951	0.017
Year=2014	1.020	0.840	1.237	0.844
Year=2015	1.235	1.037	1.474	0.019
Year=2016	1.940	1.643	2.297	0.000
Year=2017	1.543	1.279	1.860	0.000
Gender man	0.512	0.473	0.553	0.000
Age	0.973	0.961	0.985	0.000
Age^2	1.000	1.000	1.000	0.572
Income	0.999	0.999	0.999	0.000

Source: own elaborations from AD-SILC data

Annex 1.1 Migration-related variables in the AD-SILC 3.0 dataset: variables at our disposal and non-weighted proportion of individuals who provided a response

WAVE	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
PB210	(.855)	(.853)	(.853)	(.853)	(.853)	(.851)	(.847)	(.846)	(.850)	(.852)	(.853)	(.850)	(.858)	(.867)
PB220A/ citesa	(.854)	(.854)	(.853)	(.100)	(.100)	(.100)	(.100)	(.846)	(.850)	(.852)	(.852)	(.851)	(.859)	(.867)
PB220B/ citesb	(.001)	(.001)	(.001)	(.004)	(.007)	(.009)	(.008)	(.014)	(.015)	(.016)	✓	(.019)	(.017)	(.015)
cittadx								(.100)	(.100)	(.100)	✓	(.100)	(.100)	(.100)
ncitt						(.766)	(.749)	(.956)	(.954)	(.952)	(.948)	(.950)	(.942)	(.935)
secitt								(.100)	(.100)	(.100)	✓	(.100)	(.100)	(.100)
RB031/aita						(.059)	(.044)	(.063)	(.065)	(.066)	(.069)	(.070)	(.073)	(.088)
mita						(.059)	(.044)	(.063)	(.065)	(.066)	✓	(.070)	(.073)	(.088)
aallo						(.003)	(.001)	(.002)	(.002)	(.002)	✓	(.002)	(.002)	(.001)
mallo						(.003)	(.001)	(.002)	(.002)	(.002)	✓	(.002)	(.002)	(.001)
Sample size	**61,542**	**56,105**	**54,512**	**52,773**	**52,433**	**51,196**	**47,551**	**47,841**	**47,365**	**44,622**	**47,136**	**42,987**	**48,316**	**48,819**

Note: *pb210*: country of birth; *pb220a/citesa*: first citizenship; *pb220b/citesb*: second citizenship; *cittadx*: possession of the Italian citizenship at the time of the interview; *ncitt*: possession of the Italian citizenship from birth; *secitt*: whether the individual has a second citizenship; *rb031/aita*: year of first immigration to Italy; *mita*: as in *rb031/aita*, but it takes the month instead of the year; *aallo* the year after which the individual has lived in Italy without leaving for a period of one year or more; *mallo*: as in *aallo*, but it takes the month instead of the year.
✓ Data is not available yet where proportions are not reported.
Source: own elaborations from AD-SILC data

Annex 1.2 Health-related variables in the AD-SILC 3.0 dataset: variables at our disposal and non-weighted proportion of individuals who provided a response

WAVE	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
PH010	(.855)	(.854)	(.853)	(.847)	(.848)	(.843)	(.852)	(.823)	(.826)	(.820)	(.829)	(.829)	(.838)	(.839)
PH020	(.855)	(.854)	(.853)	(.841)	(.837)	(.836)	(.845)	(.818)	(.820)	(.815)	(.824)	(.824)	(.832)	(.836)
PH030/limit_d	(.855)	(.854)	(.853)	(.837)	(.837)	(.834)	(.846)	(.814)	(.819)	(.815)	(.824)	(.825)	(.826)	(.836)
PH040/sp	(.855)	(.854)	(.853)	(.855)	(.855)	(.852)	(.859)	(.846)	(.851)	(.852)	(.855)	(.851)	(.581)	(.356)
PH050/spmot	(.069)	(.057)	(.056)	(.054)	(.060)	(.059)	(.059)	(.056)	(.051)	(.062)	(.061)	(.062)	(.043)	(.015)
PH060/den	(.855)	(.854)	(.853)	(.855)	(.855)	(.852)	(.859)	(.846)	(.851)	(.852)	(.855)	(.851)	(.531)	(.236)
PH070/denmot	(.093)	(.087)	(.085)	(.080)	(.093)	(.083)	(.086)	(.092)	(.083)	(.094)	(.094)	(.091)	(.069)	(.019)
PH100														✓
PH110														(.868)
PH120														(.645)
PH130														(.868)
PH140														(.868)
PH150														(.868)
motsep								(.014)	(.013)	(.014)	(.015)	(.015)	(.016)	(.017)
malat	(1.00)	(1.00)	(1.00)	(.999)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Sample size	**61,542**	**56,105**	**54,512**	**52,773**	**52,433**	**51,196**	**47,551**	**47,841**	**47,365**	**44,622**	**47,136**	**42,987**	**48,316**	**48,819**

PH010: General health; PH020: Suffering from any chronic (long-standing) illness or condition; PH030: Limitation in activities because of health problems; PH040: Unmet need for medical examination or treatment; PH050: Main reason for unmet need for medical examination or treatment; PH060: Unmet need for dental examination or treatment; PH070: Main reason for unmet need for dental examination or treatment; PH100: Number of consultations of a medical or surgical specialist; PH110: Body mass index (BMI); PH120: Type of physical activity when working; PH130: Time spent on physical activities (excluding working) in a typical week; PH140: Frequency of eating fruit; PH150: Frequency of eating vegetables or salad.
✓ Data is not available yet where proportions are not reported.
Source: own elaborations from AD-SILC data

References

Barbieri T., Bloise F., Raitano M. (2020), Intergenerational Earnings Inequality. New Evidence from Italy, *Review of Income and Wealth*, 66, n.2, pp.418-443

Boscolo S. (2019), Quantifying the redistributive effect of the Erosion of the Italian personal income tax base: a microsimulation exercise, *Economia Pubblica*, n.2, pp.39-80

Brandolini M., Giarda E., Moriconi M., Loi M. (2009), Possibili effetti dell'under-reporting sull'analisi della ricchezza finanziaria basata sull'indagine dei bilanci delle famiglie di Banca d'Italia, in Prometeia (ed.), *Rapporto di Previsione*, Bologna, Prometeia, pp.123-131

Brier G.W. (1950), Verification of Forecasts Expressed in Terms of Probability, *Monthly Weather Review*, 78, n.1, pp.1-3

Ceccarelli C., Di Marco M., Rinaldelli C. (eds.) (2008), *L'indagine europea sui redditi e le condizioni di vita delle famiglie (EU-SILC)*, Metodi e Norme n.37, Roma, Istat

D'Aurizio G., Faiella I., Iezzi S., Neri A. (2006), *L'under-reporting della ricchezza finanziaria nell'indagine sui bilanci delle famiglie*, Temi di discussione n.610, Roma, Banca d'Italia

De Villanova C., Raitano M., Struffolino E. (2019), Longitudinal employment trajectories and health in middle life: Insights from linked administrative and survey data, *Demographic Research*, 40, n.47, pp.1375-1412

Diggle P., Kenward M.G. (1994), Informative Drop-Out in Longitudinal Data Analysis, *Journal of the Royal Statistical Society*, *Applied Statistics Series C*, 43, n.1, pp.49-73

Duncan G., Kalton G., Kasprzyk D., Singh M.P. (eds.) (1989), *Panel Surveys*, New York, Wiley & Sons

Eurostat (2020), EU statistics on income and living conditions (EU-SILC) methodology <https://bit.ly/2Ba8zFj>

Gneiting T., Raftery A. (2007), Strictly Proper Scoring Rules, Prediction, and Estimation, *Journal of the American Statistical Association*, 102, n.477 pp.359-378

Gottschalk P., Huynh M. (2010), Are Earnings Inequality and Mobility Overstated? The Impact of Nonclassical Measurement Error, *The Review of Economics and Statistics*, 92, n.2, pp.302-315

Graf M., Wenger A., Nedyalkova D. (2011), *Quality of EU-SILC data*, Ameli, Deliverable 5.1, Trier, University of Trier

Kim C., Tamborini C.R. (2012), Response Error in Earnings. An Analysis of the Survey of Income and Program Participation Matched With Administrative Data, *Sociological Methods & Research*, 43, n.1, pp.39-72

Lallo C., Raitano M. (2018), Life expectancy inequalities in the elderly by socioeconomic status: evidence from Italy, *Population Health Metrics*, 16, n.7 <https://bit.ly/31tsE3V>

Meyer B.D., Mittag N. (2019), Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net, *American Economic Journal: Applied Economics*, 11, n.2, pp.176-204

Naticchioni P., Raitano M., Vittori C. (2016), La Meglio Gioventù: Earnings Gaps across Generations and Skills in Italy, *Economia Politica: Journal of Analytical and Institutional Economics*, 33, n.2, pp.233-264

Raitano M., Fantozzi R. (2015), Political cycle and reported labour incomes in Italy: a quasi-experimental evidence on tax evasion, *European Journal of Political Economy*, 39, pp.269-280

Raitano M., Vona F. (2017), Competition, Firm Size and Returns to Skills: evidence from currency shocks and market liberalizations, *The World Economy*, 40, n.12, pp.2676-2703

Raitano M., Vona F. (2018), From the cradle to the grave: the effect of family background on the career path of Italian men, *Oxford Bulletin of Economics and Statistics*, 80, n.6, pp.1062-1088

Tedeschi S., Mazzaferro C., Morciano M., Pisano E. (2013), Modelling private wealth accumulation and spend-down in the Italian microsimulation model CAPP_DYN: A life-cycle approach, *International Journal of Microsimulation*, 6, n.2, pp.76-122

Zhang Z., Cortese G., Combescure C., Marshall R., Lee M., Lim H.J., Haller B. (2018), Overview of model validation for survival regression model with competing risks using melanoma study data, *Annals of Translational Medicine*, 6, n.16, pp.325

2. Contractual arrangements covered in AD-SILC

Introduction

The present chapter provides an overview of the key legal provisions affecting contractual arrangements in the private sector recorded in the AD-SILC dataset, including the main reforms implemented over the covered time span. The rationale is to provide inputs supporting the following analyses of transitions between working statuses as well as income trends for the different workers' categories.

In this respect, the arrangements considered are:

- Open-ended contracts;
- Part-time contracts;
- Fixed-term contracts;
- Temporary agency work[1];
- Para-subordinate contracts;
- Self-employed work;
- Standard self-employed.

The contribution does not address non-standard contractual arrangements not covered by AD-SILC, such as casual workers[2], voucher-based workers, occasional self-employed workers and other minor contracts abrogated over time.

2.1 Open-ended contracts

Open-ended contracts are the 'standard' employment contract, characterised by the indefinite duration of the relationship duration. Like other employment contracts, they

1 Temporary agency workers are not recorded as such in AD-SILC but are classified as open-ended or fixed-term workers depending on the duration of the contract.

2 Casual workers are actually split in AD-SILC between open-ended or fixed-term workers in reason of the duration of their contract. Due to their overall limited size and share in the two broad categories considered, the present contribution does not provide a focus on the related legal provisions and reforms.

balance a set of rights for the employee while granting the employer a 'directive power' and a 'disciplinary power' over the work provision (Article 2094 of the Italian Civil Code). In reason of the extension of non-standard work arrangements and, especially, of the growth of fixed-term contracts, the law (Law no. 247/2007) reaffirmed employment contracts are permanent by default, buttressing inspectorate's orders and case law converting bogus atypical arrangements into open-ended contracts.

Termination of the contract by the employer is allowed in presence of misconduct by the employee or on economic grounds. Changes meant to reduce protections against unfair dismissal and, this way, making the contract more 'appealing' for companies were implemented in 2012 (Law no. 92/2012), and in 2015 (Legislative Decree no. 23/2015)[3]. The first act weakened the system built by Article 18 of Law no. 300/1970, ruling the reinstatement of workers unfairly dismissed plus an indemnity for companies staffed with more than 15 employees (indemnities prevailing instead in smaller companies as per Law no. 604/1966). Law no. 92/2012 maintained reinstatement plus an indemnity for cases of discriminatory dismissals both above and below the staff threshold, and, limitedly to larger companies, in case the motivation adduced by the employer is found to be inexistent. Instead, the law sanctioned with a variable indemnity other cases of unfair dismissal in companies staffed with more than 15 employees. These rules are still applicable for contracts entered before the 7th of March 2015.

In fact, starting from that date, workers are applied the regime introduced by Legislative Decree no. 23/2015. The act further restricted cases under which workers have right to reinstatement. The main sanction for unfair dismissal became instead the indemnity and the act incentivised extra-judicial agreements solving disputes through monetary compensation. Under the goal of making costs for (unfair) dismissals foreseeable by the employer, the indemnity was set to range between 2 and 24 monthly pay according to scales strictly linked with seniority at work[4]. Law Decree no. 87/2018 increased these amounts to 6 and 36.

The severe limitation to the judicial scrutiny was ruled as unconstitutional by the judgement of the Constitutional Court no. 194/2018, linking the amount of the sanction also to other criteria, like the behaviour of the parties, provided its minimum and maximum values remain within the range of 6 and 36 monthly pay set by the law.

---

3 When considering the actual stability of open-ended contracts, the risk of dismissal on economic ground in outsourced activities shall be borne in mind, as the performance of contractors may be closely linked with the stability of their commercial contracts with the client company. In this respect, outsourcing was liberalised especially by Legislative Decree no. 276/2003, replacing the general principle of ownership of the means of production as a criterion to establish the lawful outsourcing of services, with the ownership by the contractor of autonomous organisation and business' risk, more difficult to identify. The act mitigated risks of abuses by sanctioning outsourcing implemented for the specific purpose of eluding mandatory rules set by law or collective bargaining (including minimum rates of pay set therein) (crime of fraudulent intermediation of manpower). This provision was abrogated by the Jobs Act (Legislative Decree no. 81/2015) and later reintroduced by Law Decree no. 87/2018, albeit this fraud is no longer considered a crime.

4 The range in companies staffed with 15 employees at most was instead between one and six monthly pay.

The new regime concerning unfair dismissal was accompanied in 2015 with a strong rebate on employers' social security contributions (8,060 euros per year for 36 months for those hired in 2015, meaning total exemption up to about 34,000 euros gross pay per year; a reduced incentive – with a maximum yearly amount equal to 3,250 euros was applied for 24 months to those hired with an open-ended arrangement in 2016).

2.2 Part-time contracts

Part-time contracts are contracts entailing a weekly working time lower than the standard duration set by law (40) or the possibly different one, set by collective agreements. Despite concerns over the misuse of part-time to conceal full-time employment relationship, part-time work has also been addressed by legislation as a tool to promote work-life conciliation or generational turnover. Therefore, the law does not set out particular restrictions for its use, focussing instead on measures meant to avoid discrimination of part-time workers, akin to Directive 97/81/EC.
Legislative Decree no. 81/2015 has introduced key changes to part-time work. While restricting the regulatory power mandated to collective bargaining, the act introduced a set of provisions strengthening the employer's discretion in changing time allocation of working hours and increasing the amount of the working week (so-called flexible clauses).
Trends in part-time shall also be read in the light of the 'concurrence' played by casual work and voucher-based work. These arrangements, introduced in 2003 to address relationships featuring low and unstable allocation of the working time, were gradually liberalised up to the restrictions introduced in 2013 for casual work, and in 2018 for voucher-based work to limit abuses. Both the schemes were even temporarily abrogated. Law no. 247/2007 abrogated provisions concerning casual work with effect from the 1st of January 2008, remaining only in the tourism and entertainment sector. Yet they were reinstated in June 2008 by Decree Law no. 112/2008. Law Decree no. 25/2017 abrogated voucher-based work in March 2017, while two months later Law no. 96/2017 reintroduced them under a set of more stringent conditions. The latest changes occurred with Law no. 96/2018, extending the possibility to use voucher-based work in accommodation activities.

2.3 Fixed-term contracts

Fixed-term contracts represent the main exception to open-ended contracts, being restricted until 2001 to seasonal activities or other cases established by collective bargaining. By means of Legislative Decree no. 368/2001, transposing Directive 1999/70/EC, fixed-term contracts were allowed in presence of 'technical, productive, organizational

or substitutive' reasons. The only limit in terms of duration entailed a maximum overall length of 36 months in case of prorogue of a fixed-term contract having an initial shorter duration. In addition, as anti-fraud rules, a stand-by period was introduced (10 days for contracts lasting less than 6 months, and 20 days for longer contracts) before a new fixed-term contract could be entered. Law no. 247/2007, setting a maximum length of 36 months even in case of consecutive fixed-term contracts, introduced stricter limits.
Among the main following changes to the wide set of rules governing fixed-term contracts, it is worth to mention:

- Law no. 92/2012: on the one side discouraging their use by introducing a 1.4% additional contribution to fund unemployment benefits; while, on the other side, eliminating the need to state justifying reasons for the first fixed-term contract;
- Law Decree no. 34/2014 as modified by Law no. 78/2014: removing justifying reasons and extending the number of possible prorogues to 5 over 36 months, while limiting the number of fixed-term contracts to a maximum threshold (20% of the permanent staff);
- Legislative Decree no. 81/2015, extending the maximum duration of fixed-term contracts to 36 months (considering also periods worked as a fixed-term temporary agency worker);
- Law Decree no. 87/2018: limiting the duration of fixed-term contracts to 24 months (considering also periods worked as a fixed-term temporary agency worker), reintroducing justifying reasons limitedly to renewals and introducing a single maximum threshold applying to both fixed-term contracts and temporary agency fixed term contracts (30% of the permanent staff).

2.4 Temporary agency work

Temporary agency work was firstly introduced in Italy by Law no. 196/97 under a number of conditions. Beside a system of public registration and authorisation for temporary work agencies to operate, the act allowed for fixed-term temporary agency contracts limitedly to: i) non-core firm tasks; ii) replacement of absent workers; iii) circumstances set forth by national collective bargaining agreements. The act balanced the risk of precarisation of temporary agency workers by imposing a 5% levy on gross pay meant to fund their vocational training, later reduced to 4%.
By means of Legislative Decree no. 276/2003, boundaries on activities for fixed-term temporary agency work were removed, allowing it in presence of a "technical, productive, organizational or substitutive" justifying reasons akin to fixed-term contracts. As a general rule, conditions applicable to fixed-term contracts were to apply also to fixed-term temporary agency work, including maximum duration and justifying rea-

sons, as modified over time by the reforms mentioned above. In addition, Legislative Decree no. 276/2003 introduced open-ended temporary agency work (staff leasing) for a limited set of activities[5].
Staff leasing was instead widely liberalised by Legislative Decree no. 81/2015, which abrogated sectoral limits while capping its use in reason of quantitative limits (i.e. 20% of the permanent staff).

2.5 Para-subordinate contracts

Para-subordinate employment, often presented as the Italian 'third status' between employment and self-employment, is grounded on provisions of the Code of Civil Procedure concerning a service performed personally and continuously by a worker, under the direction of the client.
The possibility to interpret the provisions as an ad-hoc type of employment relationship became clear in 1995, when Law no. 335/1995 created a special social security regime (the *Gestione Separata*) for para-subordinate collaborators and for self-employed workers included neither in private pension funds for liberal professionals managed by the professional association (e.g., for lawyers and architects) nor in public pension schemes for some categories managed by INPS (craftsmen, dealers and self-employed farmers). A very low 10% social security contribution rate was levied in 1996 from incomes of those enrolled in the *Gestione Separata*, much lower than the contribution rate entailed for employees (the standard being 32.7% until 2007 and 33% afterwards).
Legislative Decree no. 276/2003 further ruled para-subordinate work in the private sector by establishing a limited set of rights and specifying its legitimate use for:

- Activities linked to the implementation of a project, or a phase thereof, coordinated by the client but managed by the worker with a duty to deliver the result rather than a commitment in working time (project-based work);
- Occasional collaboration having a value lower than 5,000 euros and lasting no more than 30 days over a year (occasional para-subordinate employment), who are not obliged to pay pension contributions[6].

Law no. 92/2012 significantly restricted such rules. Acknowledging risks of abuse by companies, the act introduced:

- Stricter criteria for the use of project-based contracts, including a presumption of employment relationship in presence of work modalities similar to those of

[5] Akin to casual work, staff leasing was also abrogated by Law no. 247/2007 with effect from the 1st of January 2008, being reintroduced two years later by Law no. 191/2009.
[6] Consequently, this type of para-subordinate work is not covered in AD-SILC.

employees enrolled by the client or of performance of low-skilled and repetitive task;

- A link between the pay of project-based workers and minimum rates set for employees by collective agreements;
- A gradual rise of social security contributions attached to para-subordinate workers.

While ambitious in its approach and scope, the reform had quite a brief life to have an impact on the labour market, being soon replaced by Legislative Decree no. 81/2015. The act abrogated both project-based work and occasional para-subordinate employment. Instead, the act stated that para-subordinate workers shall be subject to protection of employees in case they are 'organised' by the client[7]. The meaning of the provision met different interpretations in courts, creating some uncertainties on the legitimate use of the contract, and, especially, on remedies applying in case of 'organisation' of the worker by the client, i.e. if the worker is to be classified as a subordinate worker (most likely permanent worker) or if he remains a para-subordinate worker while being applied (some) protections guaranteed to employees.

Provisions concerning the increase of social security contributions remained instead into force, albeit implemented on a slower pace than initially entailed by Law no. 92/2012. Para-subordinate collaborators are applied a 33% levy from 2018, after the (partial) alignment with social security contributions rates paid for employees was delayed by the 2014, 2015 and 2016 Stability Law.

For the purposes of social security law, para-subordinate work also includes a wide and heterogenous array of categories, which were not impacted significantly by the waves of reforms, such as administrators, statutory auditors, professionals. Limits and labour rights set by Legislative Decree no. 276/2003, Law no. 92/2012 and Legislative Decree no. 81/2015 have explicitly excluded these categories from their scope. Specific rules also exist for academicians labelled as para-subordinate (PhD students, postdoctoral fellows and physicians attending a postgraduate qualification course).

2.6 Self-employed workers

Provisions concerning self-employed work are grounded in article 2222 of the Civil Code, recognising the provision of services without a subordination link (i.e. without the exercise of direction and disciplinary power by the client) and mostly with the provider's own work.

[7] Provisions on para-subordinate work have been later revised by Law no. 81/2017 and by Law Decree no. 101/2019, without altering substantially the regulation established by Legislative Decree no. 81/2015.

Except for some occupation-specific rules, labour law did not address terms of employment of self-employed.
A remarkable exception was represented by Law no. 92/2012[8], deeming contracts with self-employed workers having a VAT code unlawful and to be converted into an open-ended contract in presence of two of the following three conditions[9]:

- The relationship with the same client lasts in total more than eight months a year within a period of two consecutive years;
- The amount paid to by the client represents more than 80% of the income earned by the worker within a period of two consecutive years;
- The worker has a fixed workplace within one of the client's units.

Yet, effects of the provisions were limited in scope and time. Beside a number of exceptions, e.g. the exclusion of professionals and of worker having a secondary education or higher degree, the rules were abrogated by Legislative Decree no. 81/2015. A limited set of rights against abusive contractual clauses (late payments, failure to grant a notice period for contract termination and unilateral change of contractual terms) was later established by Legislative Decree no. 81/2017, whereas Law no. 172/2017 introduced the right to a fair pay in line with parameters to be set by decree. This right is applicable only in relations with bank and insurance companies, large companies and public administration. The gap in the social security contribution rates represents a long-standing issue for its effects both on the 'concurrence' between employment contracts and (bogus) self-employed work and on the social protection coverage of self-employed workers.
While employees are applied a 33% rate (23.81% paid by the employer), plus variable levies bringing to total rate to about 40%, the self-employed are generally charged a 25% rate (plus a 0.72% levy to fund social benefits other than pensions). The rate, which is fully shouldered by workers, followed the same path of increases of para-subordinate workers until 2013, being halted at 27% before to decrease to 25% since 2017. Self-employed professionals who are members of orders (like lawyers, trade accountants, physicians, engineers and architects, notaries) shall instead contribute to social security funds established by their orders. These funds generally entail fixed annual fees independent from the declared income, plus a proportional levy usually similar to rates of *Gestione Separata*.

---

8 These provisions, entered into force on the 18th of July 2012, were adjusted by Law Decree no. 83/2012 with effect from the 12th of August 2012. For the sake of simplicity, the rules are presented as modified by the Law Decree no.83/2012.

9 From a legal point of view, the conversion was indirect. Bogus self-employed workers were to be considered project-based workers. Yet, in the very likely absence of criteria justifying project-based work, especially concerning the performance of a project, they should be subsequently considered as permanent workers.

2.7 Standard self-employed workers

For social security purposes, entrepreneurs implementing craft or trade activities primarily with their own work and with those of their family members are also considered self-employed workers and are subject to social security contributions.
In particular, crafters and traders are subject to a fixed contribution for the amount of business income below a conventional income, plus a share on the income exceeding this threshold, divided in two income brackets.
The minimum annual contribution, adjusted annually in reason of the consumer price index, is currently set at 3,836.16 euros for craftsmen and at 3,850.52 euros for dealers. Contributions over the conventional income 15,953 euros) are set at 24% up to € 47,379 and at 25% for the upper income bracket until a ceiling of 78,965 euros. These rates have been gradually increased starting from 2012 (Law Decree no. 201/2011), up from a 20% and 21% respectively.
Social security contributions are due also by self-employed farmers, i.e. small entrepreneurs habitually and usually involved in farming activities, and by agricultural entrepreneurs, provided most part of their working time is devoted to agricultural activities and they gain most of their overall labour income from these activities. These categories are subject to a 24% levy on conventional incomes depending on the value attached to their lands. The social security contribution rate of this category has been gradually increased and unified starting from 2012, previously lagging at 18,3% and at 15,3% in disadvantaged areas, with further reductions for those aged less than 21 (Law Decree no. 201/2011).

Conclusions

The various aforementioned categories may be distinguished in high detail in the AD-SILC dataset, thanks to the administrative sources. Indeed, on the one hand, INPS archives records for each working relationship the specific pension fund where the worker pays contributions. This allows us to distinguish employees, para-subordinate workers (further distinguishing collaborators and professionals enrolled to the *Gestione Separata*) and self-employed categories (craftsmen, dealers, self-employed farmers and the various categories of professionals who pay pension contributions to the specific private fund managed by their professional association). On the other hand, for what concerns the employees, INPS archives distinguish full- and part-time workers and (since 1998) those working with an open-ended or a fixed-term contract. Therefore, the AD-SILC dataset records in detail the contractual features of most individuals working in Italy. Among the categories and contracts presented above, at this time, AD-SILC does not cover occasional para-subordinate employment; being exempted from pension contributions, these workers were not tracked in administrative social security archives.

3. The duration of work arrangements

Introduction

This chapter aims at providing new findings on the duration of work arrangements. The focus is on the duration of precarious work history patterns, with some attention paid to open-ended contracts so as to investigate the actual stability of these arrangements. From our point of view, this aspect has not been thoroughly investigated within the labour market literature; we find that our preliminary evidence may invite further research on the topic.

The chapter is organised as follows. After a brief introduction of the methodology adopted, three analyses are presented, regarding, respectively, the duration of stable (open-ended) contracts, fixed-term contracts and para-subordinate contracts.

3.1 Methodological aspects

For the purposes of the analysis, we have considered data from 2004, given that from that moment we have been able to grasp the impact of the recent labour market reforms (from the entry into force of Law 30/2004, the so-called 'Biagi Law', onwards) in a 14-year-long career, from 2004 to 2017. From the INPS database, we have summarised work arrangements that are similar from the point of view of social security and employment contract features. In line with the description provided in the previous chapter, the following working categories are considered:

1. Private employees with open-ended contracts;
2. Private employees with fixed-term contracts;
3. Para-subordinate workers and self-employed professionals enrolled in *Gestione Separata* (see Chapter 2);
4. Other working categories (standard self-employed, i.e. craftsmen, dealers and self-employed farmers, professionals, i.e. architects, lawyers, etc.);
5. Unemployment or inactivity spells.

The latter status is identified with the receipt of unemployment benefits or with the absence of a contributory period.
Every individual has been associated with a sequence of non-overlapping working periods (among the working spells) with a specific length for each. A worker can experiment more than one atypical work period, we accounted bounce back and forth in every status on which the analysis has been focused.
We have distinguished three analysis referred to stable contracts, fixed-term contracts, and para-subordinate contracts. For each of them, we have investigated the determinants to leaving the status and those affecting an undesirable scenario, i.e. moving toward a less stable contract. In the case of para-subordinate workers, we have considered only workers with an income below 2,500 euro per month and a worsening of career prospects here is characterised by a transition to unemployment.
From a methodological standpoint, we have applied a survival analysis. This technique refers to the analysis of elapsed time to an event defined here according to every transition between temporary and permanent jobs. Retirement and death have been included as absorbing states, while workers that maintained the same contractual status until the end of the study do not experiment the event and therefore were considered as censored[1].
The multivariate analysis is approached with a proportional hazard regression, also called Cox regression. The response variable is the time between a time origin (in this case the beginning of an employment status, as defined above) and the end. The expected hazard, i.e. the rate of realization of the event in the next instant, is the product of the baseline hazard and the exponential function of the linear combination of the predictors. In this analysis, the outcome represents the exit from one of the working categories, and the hazard function measures the probability that if a person is still in the status in t, then the person will experience the event in the next instant. The estimated coefficients in the Cox proportional hazards regression model represent the change in the expected log of the hazard ratio relative to a one-unit change in a covariate, holding all other predictors constant. If the hazard ratio for a predictor is close to 1, then that predictor does not affect survival. If the hazard ratio is less than 1, then the predictor is protective and is associated with an improved survival of the duration of the working category. Finally, a hazard ratio greater than 1 implies that the predictor is associated with an increased risk of exit from the current working status[2].
Furthermore, the model we have applied takes into account repeated events for the same individual by introducing a frailty in the model. A frailty model is a random effects

1 Censoring is present when we have some information about a subject's event time, but we do not know the exact event time. Censoring might occur here because a subject does not experience the event before the study ends. See Kaplan &Meier, 1958.
2 For an accurate explanation of survival methodology, see Collett 1994.

model for time variables, where the random component has a multiplicative effect on the hazard. With this specification, we have described the influence of unobserved covariates in a proportional hazards model. In our assumption, we have used a frailty with Gamma distribution (Wienke 2010).
Concerning the set of regressions, aside from the overall hazard of the transition, we have explored in depth the specific transition to worse contractual arrangements. In this case, different contractual arrangements (corresponding here to a worsening, or improvement, of career prospects) can be possible and the occurrence of one can be precluded by another (Austin *et al.* 2016). In this case, survival models with competing risks have to be considered instead of a simple censoring, because they may give biased estimates of hazard ratios (Fine and Gray 1999). This way, the in-depth analysis has produced cause-specific hazard models for one event (a worsening of career prospects), taking into account the other (improvement).
The covariates included in the analysis are: age (under 30, 30-40, 40-50, 50-60 and above 60, the reference category), gender (female as reference), highest level of education attained (up to middle school, the reference, high school, university and over), residential area (North, the reference, Centre and South), income (in logarithmic scale), and three binary variables designed to capture the characteristics of the firm: the size (more than 200 workers) and the sector of activity (services and manufacturing).

3.2 The duration of stable contracts

The first analysis focuses on the duration of stable contracts. The research question aims at investigating the effective stability of open-ended contracts. The results of the analysis are reported in Table 3.1. Age plays a clear role: when keeping the other predictors constant, the mere fact of being senior protects against the risk of leaving contractual stability. This is probably due to the greater experience and know-how accrued by workers over time. Women, as shown in other papers (Fabrizi *et al.* 2012), have a lower risk of exiting from the stability area, keeping other variables constant, including income, therefore controlling for gender pay gap and unobserved covariates (frailty). It is not surprising that in Northern regions stable contracts are more lasting than in the other Italian areas, and that a high income (proxy of a better job) protects workers to remain with an open-ended contract. The educational attainment depends on different scenarios after the exit from the stability area. This aspect has been investigated in depth with the following analysis.
The focus then turns on the transition toward an unstable contract, considering, as a competitive risk, the opposite situation, i.e. an improvement of career prospects, characterised by a transition from contractual stability toward self-employment (i.e. self-employed professionals).

Table 3.1 Cox proportional hazards regression model with Gamma frailty on time to exit from a stable contract

	Coef	(SE)	P-value	HR	Lower.95	Upper.95
Age class (ref. Over 60)						
Under 30	0.069	0.027	0.011	1.072	1.016	1.130
30-40	-0.033	0.026	0.210	0.968	0.919	1.019
40-50	-0.137	0.026	0.000	0.872	0.828	0.918
50-60	-0.097	0.027	0.000	0.907	0.860	0.958
Man	0.101	0.012	0.000	1.107	1.081	1.134
Educational attainment (ref. Up to middle school)						
High school	-0.109	0.013	0.000	0.897	0.874	0.921
University	0.165	0.017	0.000	1.180	1.141	1.220
Macro areas (ref. Northern Italy)						
Central Italy	0.100	0.014	0.000	1.106	1.075	1.137
Southern Italy	0.349	0.015	0.000	1.418	1.376	1.460
Income (log)	-0.494	0.006	0.000	0.610	0.604	0.617
Large company	-0.018	0.017	0.270	0.982	0.950	1.015
Industry Sector	0.034	0.015	0.027	1.034	1.004	1.065
Services Sector	0.023	0.013	0.080	1.023	0.997	1.050

Note: Variance of random effect t = 1.188485; I-likelihood = -892788; P-value 0.000; N = 133018; Number of events = 81153.
Source: own elaborations from AD-SILC data

Table 3.2 shows that the cause-specific hazard of worsening career decreases with a higher educational level, keeping constant all of the other predictors in the model. This suggests that investing in human capital tends to have a high return in the Italian labour market given this protective function against undesirable job outlook scenarios. For the other predictors, the magnitude of relative effects on the incidence of adverse consequences is similar to the outcome observed for overall transitions.

Table 3.2 **Cox proportional hazards regression model with Gamma frailty on time to exit from a stable contract towards a less stable contract**

	Coef	(SE)	P-value	HR	Lower.95	Upper.95
Age class (ref. Over 60)						
Under 30	0.112	0.028	0.000	1.119	1.060	1.181
30-40	-0.011	0.027	0.690	0.989	0.938	1.043
40-50	-0.140	0.027	0.000	0.869	0.825	0.916
50-60	-0.092	0.028	0.001	0.912	0.863	0.964
Man	0.138	0.012	0.000	1.148	1.121	1.177
Educational attainment (ref. Up to middle school)						
High school	-0.202	0.013	0.000	0.817	0.796	0.839
University	-0.196	0.018	0.000	0.822	0.794	0.851
Macro areas (ref. Northern Italy)						
Central Italy	0.133	0.014	0.000	1.142	1.110	1.174
Southern Italy	0.381	0.015	0.000	1.464	1.421	1.509
Income (log)	-0.521	0.006	0.000	0.594	0.587	0.601
Large company	-0.023	0.018	0.210	0.978	0.944	1.013
Industry Sector	0.036	0.016	0.028	1.036	1.004	1.070
Services Sector	0.026	0.014	0.065	1.026	0.998	1.055

Note: Variance of random effect = 1.090077; I-likelihood = -736796.8; P-value 0.000; N = 133018; Number of events = 67173
Source: own elaborations from AD-SILC data

3.3 The duration of fixed-term contracts

The second analysis was set up to explore the duration of fixed-term contracts (Tables 3.3 and 3.4). The overall effect on the risk of leaving the status decreases with age. To better understand this aspect, we move to Table 3.4, where we can confirm that age plays a role and we observe a protection against the transition toward more precarious employment (para-subordinate or unemployment). However, the incidence of cause-specific hazard falls steadily until about age 50, and then rises again after that age. It means that the issue of precariousness is not limited to the younger component of the labour market.

For fixed-term contracts, similarly to what has been seen above, education helps preventing undesirable outcomes. Here we observe an even stronger effect of higher education in helping people to prevent transitions to unsecure jobs.

Table 3.3 Cox proportional hazards regression model with Gamma frailty on time to exit from a fixed-term contract

	Coef	(SE)	P-value	HR	Lower.95	Upper.95
Age class (ref. Over 60)						
Under 30	0.446	0.027	0.000	1.562	1.481	1.647
30-40	0.403	0.027	0.000	1.497	1.419	1.578
40-50	0.370	0.027	0.000	1.448	1.373	1.528
50-60	0.237	0.029	0.000	1.267	1.198	1.340
Man	0.184	0.011	0.000	1.203	1.177	1.229
Educational attainment (ref. Up to middle school)						
High school	-0.016	0.012	0.180	0.984	0.961	1.007
University	-0.052	0.016	0.001	0.949	0.919	0.980
Macro areas (ref. Northern Italy)						
Central Italy	0.004	0.013	0.780	1.004	0.978	1.030
Southern Italy	0.031	0.013	0.020	1.031	1.005	1.058
Income (log)	0.112	0.004	0.000	1.119	1.109	1.129
Large company	0.046	0.014	0.001	1.047	1.019	1.077
Industry Sector	-0.028	0.012	0.022	0.972	0.949	0.996
Services Sector	-0.027	0.010	0.010	0.974	0.954	0.994

Note: Variance of random effect = 1.090077; I-likelihood = -736796.8; P-value 0.000; N= 138161; Number of events = 120040
Source: own elaborations from AD-SILC data

Table 3.4 **Cox proportional hazards regression model with Gamma frailty on time to exit from a fixed-term contract towards a less stable contract**

	Coef	(SE)	P-value	HR	Lower.95	Upper.95
Age class (ref. Over 60)						
Under 30	0.318	0.039	0.000	1.374	1.274	1.482
30-40	0.292	0.039	0.000	1.339	1.241	1.444
40-50	0.188	0.041	0.000	1.206	1.114	1.306
50-60	0.465	0.039	0.000	1.592	1.476	1.717
Man	-0.039	0.016	0.015	0.962	0.933	0.993
Educational attainment (ref. Up to middle school)						
High school	-0.134	0.017	0.000	0.874	0.845	0.905
University	-0.247	0.024	0.000	0.781	0.746	0.818
Macro areas (ref. Northern Italy)						
Central Italy	0.043	0.020	0.030	1.044	1.004	1.084
Southern Italy	0.439	0.018	0.000	1.551	1.496	1.608
Income (log)	0.235	0.006	0.000	1.265	1.249	1.281
Large company	0.024	0.018	0.190	1.024	0.988	1.062
Industry Sector	-0.069	0.016	0.000	0.933	0.905	0.963
Services Sector	-0.049	0.013	0.000	0.952	0.928	0.978

Note: Variance of random effect = 1.418594; l-likelihood = -794789.3; P-value=0.000; N = 138161; Number of events = 73991
Source: own elaborations from AD-SILC data

3.4 The duration of para-subordinate contracts

In the last analysis, we have examined the most unstable area of the labour market. In most of the cases, indeed, para-subordinates work as dependent employees, with reduced social and welfare guarantees (Muehlberger and Pasqua 2009). Here we have very different results comparing overall hazard rate to exit from the status (Table 3.5) and the cause-specific hazard of a transition to unemployment (Table 3.6). Here, more than elsewhere, a composition effect is observed, depending on different scenarios after the exit from para-subordinate status. This is very clear if we compare the coefficients (both sign and magnitude) in Tables 3.5 and 3.6. This means that if we want to analyse precariousness in the labour market, it is not enough to look at the duration of the status (para-subordinate) because we have, using a terminology that has been consistently used in the labour market literature, both stepping-stone and dead-end scenarios (Alison *et al.* 2000). Because we are here interested in job fragility, we take a look directly at Table 3.6. Here we observe a strong age effect concentrated in the youngest part of the working force. In the worst scenario (from para-subordinate status to unemployment), men are less at risk than women. Once again, educational attainment plays a role and in the expected direction. The strongest protective effect is given by income, as a proxy of qualification and know-how.

Table 3.5 Cox proportional hazards regression model with Gamma frailty on time to exit from a para-subordinate contract

	Coef	(Se)	P-value	HR	Lower.95	Upper.95
Age class (ref. Over 60)						
Under 30	-0.010	0.033	0.760	0.990	0.927	1.057
30-40	0.108	0.033	0.001	1.114	1.043	1.188
40-50	0.182	0.033	0.000	1.200	1.124	1.281
50-60	0.179	0.035	0.000	1.195	1.116	1.281
Man	0.267	0.013	0.000	1.307	1.273	1.341
Educational attainment (ref. Up to middle school)						
High school	0.102	0.017	0.000	1.107	1.071	1.145
University	0.153	0.019	0.000	1.166	1.124	1.209
Macro areas (ref. Northern Italy)						
Central Italy	-0.096	0.015	0.000	0.908	0.882	0.936
Southern Italy	-0.235	0.017	0.000	0.791	0.764	0.818
Income (log)	-0.122	0.003	0.000	0.885	0.879	0.891
Large company	0.014	0.022	0.510	1.014	0.972	1.059
Industry Sector	-0.016	0.019	0.420	0.985	0.948	1.022
Services Sector	-0.011	0.016	0.490	0.989	0.957	1.021

Note: Variance of random effect = 0.2297699; I-likelihood = -382501.1; P-value 0.000; N = 43686; Number of events = 39209
Source: own elaborations from AD-SILC data

Table 3.6 **Cox proportional hazards regression model with Gamma frailty on time to exit from a para-subordinate contract towards unemployment**

	Coef	(SE)	P-value	HR	Lower.95	Upper.95
Age class (ref. Over 60)						
Under 30	1.162	0.087	0.000	3.196	2.693	3.793
30-40	1.030	0.088	0.000	2.801	2.359	3.326
40-50	0.720	0.089	0.000	2.054	1.725	2.445
50-60	0.399	0.094	0.000	1.490	1.239	1.792
Man	-0.113	0.026	0.000	0.893	0.850	0.939
Educational attainment (ref. Up to middle school)						
High school	-0.097	0.032	0.002	0.907	0.852	0.966
University	-0.078	0.036	0.029	0.925	0.862	0.992
Macro areas (ref. Northern Italy)						
Central Italy	0.012	0.030	0.700	1.012	0.954	1.074
Southern Italy	0.050	0.032	0.120	1.051	0.987	1.119
Income (log)	-0.263	0.007	0.000	0.769	0.759	0.779
Large company	0.013	0.043	0.760	1.014	0.931	1.103
Industry Sector	-0.004	0.038	0.920	0.996	0.925	1.073
Services Sector	-0.021	0.033	0.530	0.980	0.919	1.044

Note: Variance of random effect = 0.7423179; I-likelihood = -92528.5; P-value 0.000; N = 43686; Number of events = 9507
Source: own elaborations from AD-SILC data

Conclusions

In summary, duration models are useful to identify individual characteristics playing a role in transition among statuses. Furthermore, a competing risk model has been adopted to evaluate the incidence of cause-specific hazard. In the final report of the MOSPI project, we will improve the database with a complete career for every worker (works on the AD-SILC dataset are still in progress, as reported in Chapter 1) and cross-validation analyses will be performed. Nevertheless, the first findings we have obtained so far can be useful as a base for future research questions.

References

Austin P., Lee D., Fine J. (2016), Introduction to the Analysis of Survival Data in the Presence of Competing Risks, *Circulation*, 133, n.6, pp.601-609

Collett D. (1994), *Modelling Survival Data in Medical Research*, London, Chapman & Hall

Fabrizi E., Farcomeni A., Gatta V. (2012), Modelling work history patterns in the Italian labour market, *Statistical Methods & Applications*, 21, n.2, pp.227-247

Fine J.P., Gray R.J. (1999), A proportional hazards model for the subdistribution of a competing risk, *Journal of the American Statistical Association*, 94, n.446, pp.496-509

Kaplan E.L., Meier P. (1958), Nonparametric Estimation from Incomplete Observations, *Journal of the American Statistical Association*, 53, n.282, pp.457-481

Muehlberger U., Pasqua S. (2009), Workers on the Border Between Employment and Self-Employment, *Review of Social Economy*, 67, n.2, pp.201-228

Wienke A. (2010), *Frailty Models in Survival Analysis*, Boca Raton FL, CRC Press

Zhang Z. (2017), Survival analysis in the presence of competing risks, *Annals of Translational Medicine*, 5, n.3, pp.1-9

4. Earnings distribution in private employment: trends over the 2004-2018 period

Introduction

The economic literature and the economic policy debate have been increasingly concerned with the inequality in income and in standard of living of individuals and households (see for example OECD 2008, 2011, 2015 and the increasing number of academic articles and books focusing on this topic in recent years, e.g. Salverda *et al.* 2009, 2014; Stiglitz 2012; Piketty 2014; Atkinson 2015; Milanovic 2016).
The concern has arisen due to growing empirical evidence – made easier by the increased availability of proper data coming from different sources – which has proved that in most developed countries during the last decades, and also after the economic crisis that began in 2008 (OECD 2016), incomes and earnings have become more dispersed and also more and more concentrated in the hands of small and privileged segments of the society (the top 1 or 0.1%; Atkinson *et al.* 2011). The pace and the extent to which such trends have taken place have not been the same in all countries but, because of such trends, inequality is rather high and increasing in several developed and developing countries (OECD 2011; Atkinson 2015; Bourguignon 2018).
The economic literature (Canberra Group 2011) suggests that – to capture the distribution of economic wellbeing over a population – the best proxy of individuals' economic wellbeing is the equivalised disposable income, i.e. all net-of-tax incomes earned in the market by household members from whatever source (employment, self-employment, capital, land) including welfare state transfers and equivalised by dividing total income by the so-called *equivalence scale* in order to take into account differences in households' sizes.
However, for a better understanding of the mechanisms engendering disposable income inequality, the process that shapes equivalised disposable incomes may be depicted as a chain made of three links (OECD 2011; Raitano 2018). The first link refers to individual earnings inequality – the focus of this chapter – and is related to the labour market outcomes achieved by workers, which depend on hourly wages, working hours, the type and stability of the contractual arrangement and the duration of possible unemployment spells. All these factors are affected by mechanisms related

to labour market equilibria. The second link acts at the household level and refers to market income inequality related to earnings of all the household members (and, therefore, to how many of them are employed and, more in general, to household composition). Furthermore, this second link relates to incomes stemming from all other market sources, including capital income and gains, if any. The third and last link refers to the public redistribution through taxes and transfers (e.g. pensions, unemployment benefits, and minimum income schemes).

Although the analysis of the mechanisms underlying inequality requires looking at all the three links mentioned above and their interactions, it should be noted that the first link – i.e. the labour market – represents in all developed countries the place in which most of income inequalities are formed due to the crucial role that wages play in the formation of household income. For instance, by decomposing by source disposable income inequality in EU-15 countries, Raitano (2016) finds that labour income (employment and self-employment) has by far the largest proportionate inequality contribution, amounting as a share of total inequality to 89.7% in Nordic countries, 63.3% in Continental countries, 96.7% in the UK and 86.7% in Southern countries.

Starting from this evidence, this chapter investigates the evolution of the characteristics of the earnings distribution of employees in the private sector in Italy over a 15-year period (2004-2018), thus allowing us to observe whether changes in earnings distribution have occurred after the economic crisis. Nevertheless, since we only focus on private employees' earnings, it has to be pointed out that the evidence provided in this chapter does not provide an exhaustive picture of inequality trends in Italy. On the one hand, we do not take into account earnings distribution among public employees and the self-employed. On the other hand, a possible increase (reduction) in earnings inequality might be attenuated or more than compensated by the factors that act in the other links of the aforementioned chain (household composition and employment rates, non-labour market income, taxes and transfers). However, despite this caveats, we provide in this chapter a very detailed piece of descriptive evidence about what has happened in the first link of the 'inequality chain' over the last 15 years and what the main factors that may have affected the trend of earnings distribution are[1].

By carrying out decomposition exercises, we will investigate the role played as a driver of earnings inequality by basic individuals' characteristics – i.e. gender, age, education, citizenship – and features of the working activity – e.g. contractual arrangement, firm's size and sector. In particular, by distinguishing workers by their educational attainment, we will evaluate whether rising skill premia have played a crucial role for explaining observed trends of earnings inequality. Indeed, this is suggested by the strand of the

---

1 Naticchioni and Raitano (2019) using data on the universe of private employees in Italy carried out a similar analysis for the 1975-2017 period. However, due to data limits, these authors did not take into account workers' education.

economic literature that investigates drivers of income and earnings inequality by focusing mainly on wage gaps between workers with different levels of skills, usually represented by educational attainments in empirical analyses.

Most existing studies consider the rise in inequality as a consequence of processes that take place in the labour market. It is generally claimed in mainstream studies that structural phenomena such as skill biased technological change, digitalisation and globalisation, through their impact on labour supply and demand, have widened the productivity gap between differently skilled workers (e.g. Bound and Johnson 1992; Katz and Murphy 1992; Acemoglu and Autor 2011; Autor *et al.* 2013). According to this interpretation, in order to be productive, ICT technologies have to be coupled with high skilled workers, whose demand would therefore increase. On the contrary, new technologies along with the competition of low skilled workers in less developed countries fostered by the globalisation process (that also fostered the chances of off-shoring jobs in such countries) would push down the demand for low skilled workers in advanced countries, thereby compressing their wages. In a similar vein, more recently, various authors (e.g. Autor *et al.* 2006; Goos and Manning 2007; Acemoglu and Autor 2011; Cortes *et al.* 2014) have pointed out that technological progress and digitalisation mostly displaced medium skilled workers performing routine-jobs (e.g. clerical and administrative jobs) which are neither complementary to ICT (like non-routine cognitive tasks performed by high-skilled workers) nor neutral to it (like non-routine manual tasks performed by low-skilled workers). As a result, the labour market would have become polarised, i.e. most high-skilled workers and some low-skilled workers, but not medium-skilled workers, would enjoy higher wages and better occupational opportunities (Das and Hilgenstock 2018).

In light of the above, the aim of this chapter is twofold: i) presenting original and detailed descriptive evidence on trends of the earnings distribution among private employees by main characteristics in Italy over the 2004-2018 period; ii) carrying out decomposition exercises to distinguish to what extent observed trends of earnings inequality are explained by workers' and firms' characteristics. If other factors play a role, then a 'residual inequality' would emerge.

The chapter is structured as follows. In Section 4.1, we describe the data employed. In Section 4.2, the trend of the composition of the labour force by workers' characteristics and contractual arrangements is shown and discussed, while the distribution of worked weeks over the year is displayed in Section 4.3. Then, we present trends of mean earnings in Section 4.4 and main percentiles of the earnings distributions in Section 4.5. We devote a specific focus to low-paid workers in Section 4.6, which we call 'in work poor'. Finally, we explicitly show trends of earnings inequality in Section 4.7 and present findings of the decomposition exercises run in order to evaluate the role played by some workers' features as a driver of inequality in Section 4.8). Section 4.9 summarises and concludes the chapter.

4.1 Data

We use the innovative dataset AD-SILC, 3.0 which, as explained in Chapter 1, has been built by merging the 2004-2017 cross-sectional waves of IT-SILC – the Italian component of the European Union Statistics of Income and Living Conditions (EU-SILC) – with administrative social security records tracking private employees since their entry in the labour market up to the end of 2018 (managed by the INPS). As already mentioned, social security files report, on an annual basis, gross earnings, weeks of work and some worker's and firm's characteristics (e.g. the broad occupation – i.e. apprentice, blue-collar, white-collar, manager – the province of work, the industry), but do not record worker's education. On the contrary, IT-SILC waves record education, but cover a limited time span. Merging the longitudinal administrative record of all individuals interviewed in IT-SILC during the 2004-2017 period with IT-SILC information allows us to enrich social security records with the missing information about education.

In this chapter, we focus on employees in the private sector, since most studies about earnings distribution in Italy and abroad focus on this workers' category, which includes the large majority of workers (e.g. Card *et al.* 2013; Devicienti *et al.* 2018; Franzini and Raitano 2019)[2]. Moreover, the fact of relying on earnings data recorded in administrative files allows to greatly reduce measurement errors with respect to survey data even if it does not guarantee a full representativeness of top earners[3].

For each year of the 2004-2018 period, we consider all individuals with at least positive wage as a private employee in that year, excluding those aged less than 15 and more than 65[4]. This makes our definition of workers different from the definition of employed individual used in the Labour Force Survey, since we consider all individuals who received any salary in a certain year even if they spent most of the year without working. Our final dataset includes around 900,000 observations, including approximately 60,000 workers per year.

Apart from gender, employees are classified in various groups according to their characteristics: i) in four age groups (15-29, 30-39, 40-49 and 50-65); ii) in three groups according to the highest educational attainment (at most lower secondary, upper secondary and tertiary); iii) in three groups according to their citizenship (Italian, other European Union country and no EU country).

---

2 Furthermore, it is worth noting that earnings of public employees may still have some flaws in administrative archives due to the recent digitalisation of their information, while self-employed earnings are plagued by underreporting and by top and bottom coding in administrative archives.

3 Since our sample is based on the IT-SILC samples, it is not perfectly representative of top earners because of either underreporting by the very rich respondents or lower survey participation rates by the very rich households, or both (Burkhauser *et al.* 2018).

4 We do not include in our computations 'zero earners', i.e. those individuals without positive earnings in certain years (e.g. the long-term unemployed).

In what follows, we make use of three earnings variables (all variables are expressed in euros, converted to 2018 constant prices using the harmonised index of consumer prices)[5]:

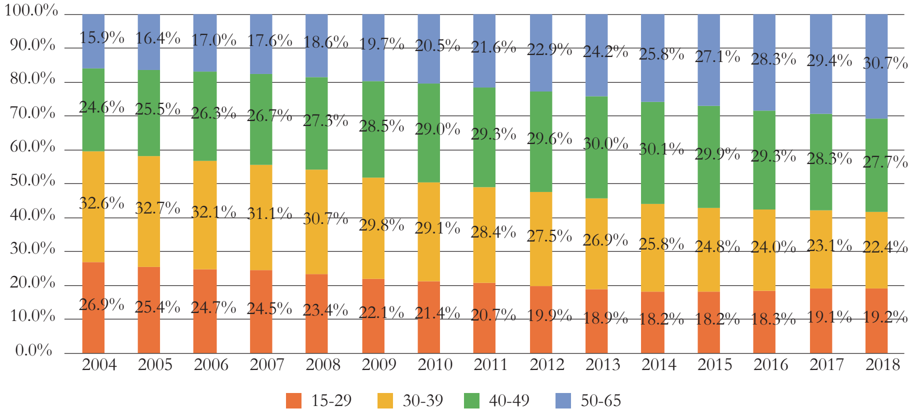
1. Annual earnings, the most suitable variable to capture the influence of labour market outcomes on workers' living standard. Annual earnings depend on unitary wages – i.e. hourly wages, not recorded in INPS archives – and hours of work in the year, which depend on the number of hours usually worked in a week (i.e. on overtime and part-time arrangements) and on the number of working weeks (i.e. on periods spent out of work in a year, which is also affected by temporary arrangements);
2. Weekly wage, not influenced by the varying number of weeks of work in a year; we consider the weekly wage of the 'main' working period during the year, which is obtained by dividing the most rewarded working period during the year by the corresponding worked weeks;
3. Weekly wage of 'strong workers', i.e. those individuals working the whole year (52 weeks) with a full-time arrangement. This is the most reliable proxy of hourly wage, usually considered by economists as the best measure of workers' productivity and as the outcome of the contractual bargaining.

4.2 Trend of the composition of the labour force

The composition of the workforce in the private sector in Italy significantly changed over the observed period. First, consistently with a greying process of the workforce linked to demographic trends, the share of employees in the youngest age class decreased from 26.9% to 19.2% from 2004 to 2018 while, conversely, the share of employees aged at least 50 rose from 15.9% to 30.7% in the same period (Figure 4.1). Apart from labour market dynamics, which affect mostly the chances of youngest workers to be hired through an employment arrangement rather than through an atypical para-subordinate arrangement (see Chapter 2), these trends may also be imputed, on the one hand, to the increasing participation rate in upper secondary and tertiary programmes by Italian youngsters, and, on the other hand, to the tightened conditions for having access to an old-age or an early-retirement pension introduced in the observed period, in particular after the 2011 reform (Jessoula and Raitano 2017).

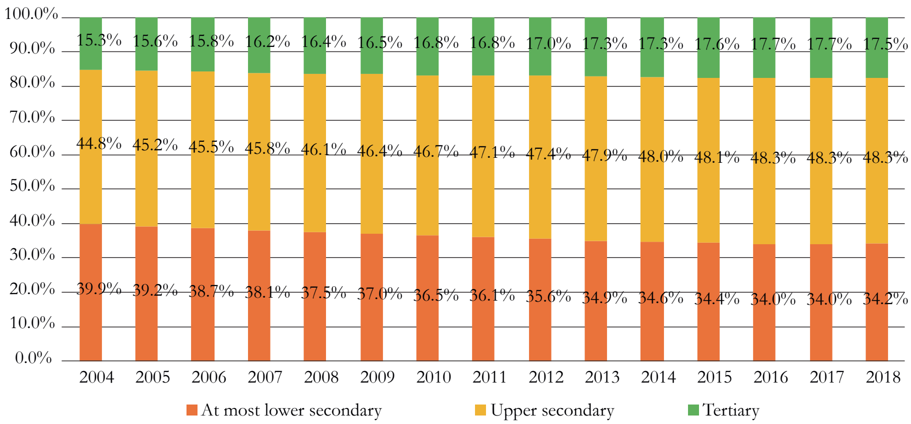
[5] When computing our earnings variables, we only refer to earnings corresponding to working periods, i.e. we do not include possible amounts received as indemnities for maternity, sickness or job suspension.

Figure 4.1 Employees distribution by age class



Source: own elaborations on AD-SILC data

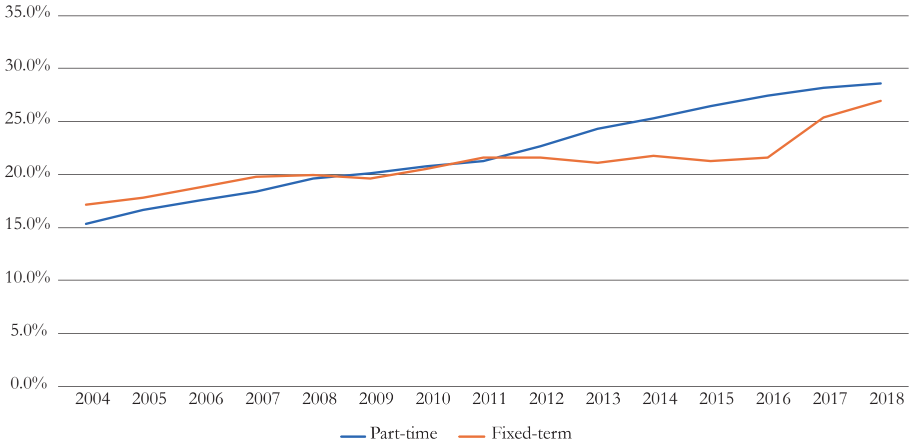
Figure 4.2 Employees distribution by education



Source: own elaborations on AD-SILC data

A process of structural change in the workforce composition also emerges when looking at the distribution of employees by education (Figure 4.2): the share of workers with at most a lower secondary degree dropped by 5.7% from 2004 to 2018, while the share of employees who attained a tertiary degree rose by 2.2% in the observed period. Major changes in the characteristics of the labour market also emerge when focusing on the spread of non-standard contracts. On the one hand, we consider the dichotomy between part-time and full-time arrangements, on the other, the one between open-ended and fixed-term arrangements (Figure 4.3). The spread of both types of atypical arrangements greatly rose in the observed period among private employees. Nevertheless, the two types of arrangements are characterised by different trends: the share of those who spent the most rewarded annual job relationship with a part-time arrangement rose from approximately 15% to 29% from 2004 and 2018 and, interestingly, the increase became faster starting from 2011, i.e. from the upsurge of the sovereign debt crisis in Italy.

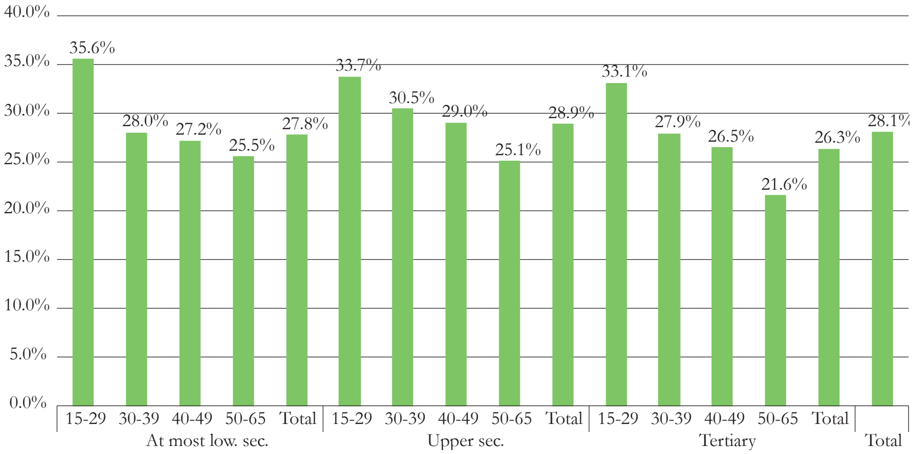
Figure 4.3 Spread of atypical contracts



Source: own elaborations on AD-SILC data

The share of employees working with a fixed-term contract rose from approximately 15% to 27% in the observed period. However, the time trend was characterised by a steep increase from 2004 to 2007, a sort of constancy up to 2016 and a further very steep rise starting from 2017, when the generous monetary incentives (i.e. social contributions reliefs) provided from 2015 to firms hiring with an open-ended arrangement were abolished. It has to be noticed that the slight constancy of the incidence of temporary arrangements during the crisis period might be due to a 'composition effect' that brought firms to not renew fixed-term contracts, thus reducing the spread of this type of arrangement when total employment was reduced.
Despite these increasing trends, the diffusion of atypical employment arrangements differs according to workers' education and age. The share of individuals working on a part-time basis in 2018 slightly differs by employees' education, with the highest shares emerging among the youngsters (Figure 4.4). Interestingly, the share of part-timers by age groups decreases with increasing age for each of the educational attainment level considered.

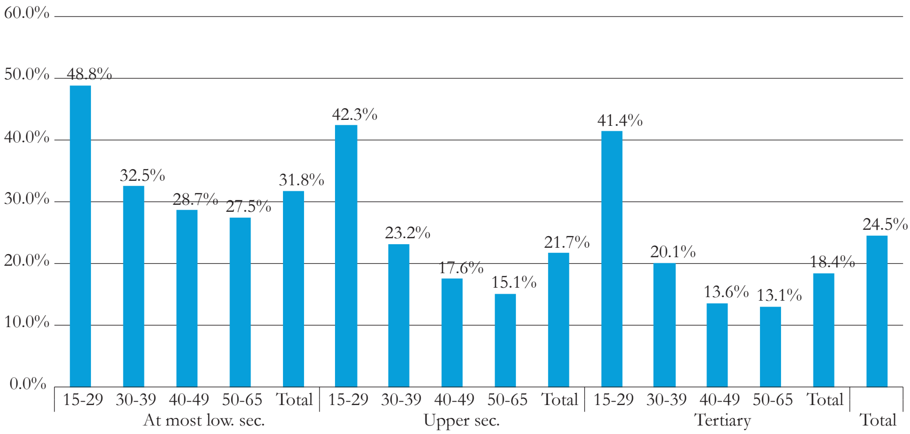
Figure 4.4 Share of employees working with a part-time arrangement by education and age group in 2018



Source: own elaborations on AD-SILC data

Sharper differences by age and education emerge with regard to the share of fixed-term employees in 2018 (Figure 4.5). This contractual arrangement mostly characterises individuals aged 15-29 and the low-skilled, even if not-negligible shares of fixed-term employees also emerge in all age classes and among the tertiary graduates.

Figure 4.5 Share of employees working with a fixed-term arrangement by education and age group in 2018



Source: own elaborations on AD-SILC data

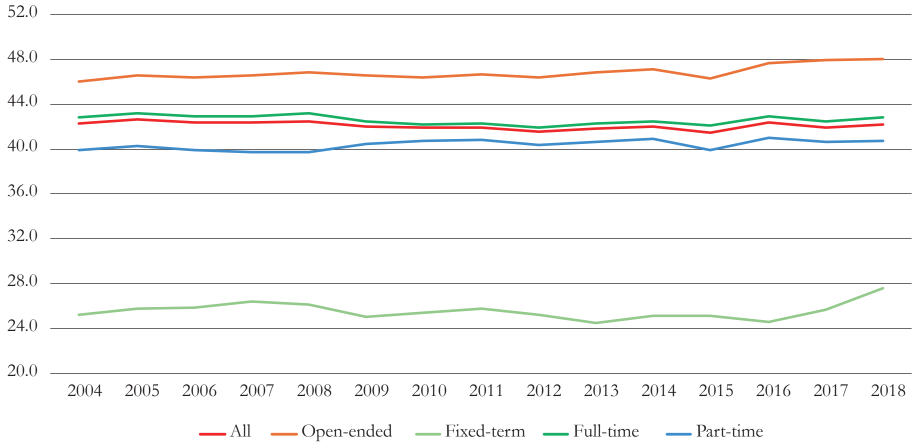
4.3 Trend of annual worked weeks

As pointed out, annual earnings inequality strictly depends on how many weeks a worker is able (or available in case of voluntary unemployment) to work in a year. This crucial driver of earnings inequality affects above all those working on a fixed-term basis, and in particular seasonal workers, who are often unable to work 52 weeks in a year. Figure 4.6 confirms this stylized fact, showing the trend of mean annual worked weeks by employees' contractual arrangement.

The average number of worked weeks decreased from 42.5 in 2005 to 41.5 in 2015, even though that figure has risen in recent years. As far as open-ended employees are concerned, the mean number of worked weeks has risen in recent years reaching approximately a value of 48. As to fixed-term employees, they worked on average 27.6 weeks in 2018, but lower values were recorded in past years (a minimum of 24.5 weeks was reached in 2013). No relevant differences in mean worked weeks emerge between full- and part-timers, even though part-timers work on average fewer weeks than full-timers (40.7 vs 42.8 in 2018).

As a further indicator of the distribution of worked weeks, we computed the incidence of 'weeks poverty' by employees' characteristics in 2018, where employees (with at least one week of work in a year) are defined as 'weeks poor' if they worked less than 27 weeks in a year (Figure 4.7)[6].

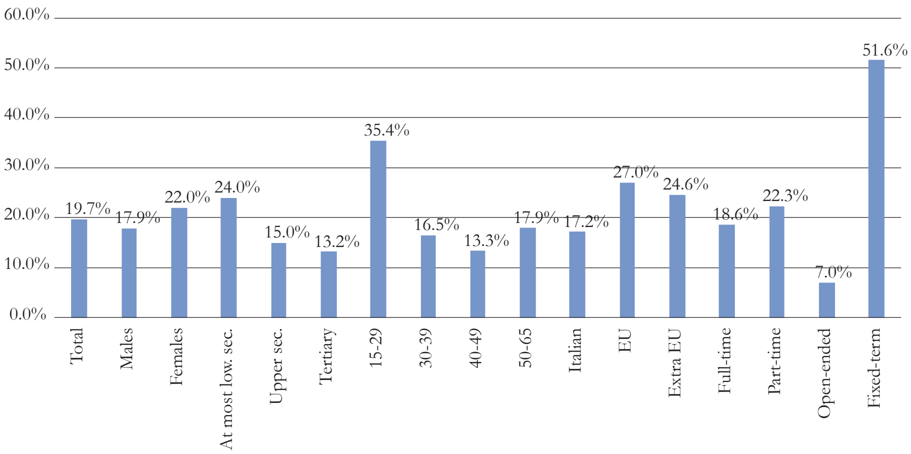
Figure 4.6 Mean number of worked weeks by contractual arrangement



Source: own elaborations on AD-SILC data

[6] Note that the incidence of weeks poor also includes those workers who retire having worked less than 27 weeks in the final year of their career.

Figure 4.7 Share of 'weeks poor' by employees' characteristics in 2018



Source: own elaborations on AD-SILC data

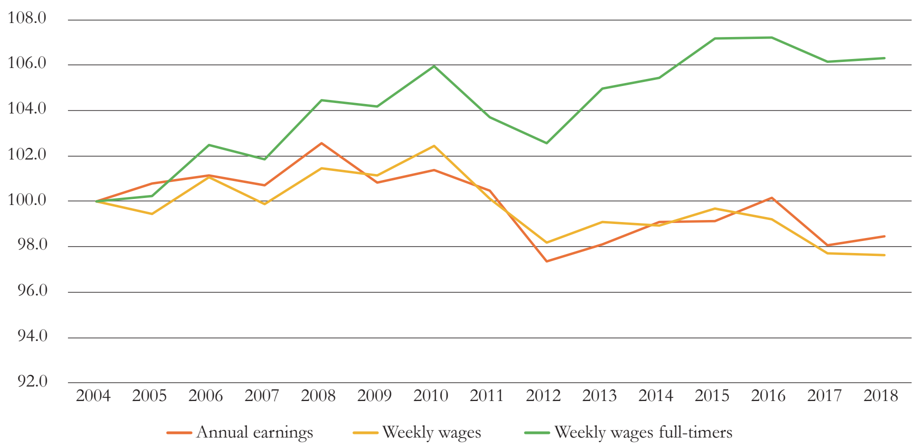
On average, 19.7% of employees are weeks poor, but this share is largely heterogeneous among workers' types according to the risk of interrupted careers. What matters most is clearly the risk of working on a fixed term-basis. Indeed, 51.6% of those who were fixed-term employees during the highest rewarded job relationship in 2018 were weeks poor. Furthermore, consistently with the high spread of temporary arrangements among the youngsters, 35.4% of those aged 15-29 were weeks poor in 2018.

4.4 Trend of mean earnings

Over the 2004-2018 period, we find different trends of real mean earnings according to the notion of wage considered (see Figure 4.8, which shows the rate of growth with respect to 2004). As far as the pre-crisis period is concerned, both annual earnings and weekly wages remained rather stagnant with a maximum cumulated increase amounting to 2.5% at most. On the contrary, we find a not-negligible increase in real mean weekly wages (6% from 2004 to 2010) when we focus solely on full-timers, suggesting that the stagnation was mostly due to the increasing spread of part-time arrangements. The picture changed since the upsurge of the sovereign debt crisis in 2011, where a large drop in mean values characterised the three series until 2013. After that, a recovery has taken place, but it has been much more intense for weekly wages of 'strong workers' than for the other two series. In real terms, mean annual earnings and weekly wages were 1.5 and 2.4% lower in 2018 than in 2004. In addition, it is worth noting that the picture regarding full-timers comes with a reassuring message, since Italy has been characterised by a large increase in the share of (often involuntary) part-time jobs over the last decades. This leads us to stress that our proxy of hourly wages (i.e. weekly wages of full-timers) is not representative of all workers but relates to a declining share of relatively advantaged employees.

When we distinguish workers by their main characteristics, differences emerge in both levels and trends of annual earnings (Table 4.1). The wage gender gap has remained rather constant over the observed period, while a not-negligible reduction in the gap between those aged 40-49 and those over 50 has occurred, perhaps due to the changing composition of the older workforce, which is also related to the effects of the increase in older workers' employment rates established by the pension reform process. On average, extra-EU citizens experienced a consistent drop in mean annual earnings during the crisis period, with a slight recovery from 2017 onwards. Nevertheless, both EU and non-EU citizens earn on average more than one-third less than Italian citizens. Interestingly, focusing on returns to education, an increase of the earnings premium for tertiary graduates with respect to those with at most a lower secondary or an upper secondary degree is found throughout the time span chosen. To put it in even simpler terms, studying more is a rewarding factor in the Italian labour market, for an increased wage premium overtime that is partially due to the lower probability high-skilled workers to work with an atypical contract scheme. Nevertheless, we are aware that the mere focus on mean returns is not enough to assess risks of investing in human capital adequately, which will be analysed in the following sections by looking at earnings heterogeneity among similarly educated workers.

Figure 4.8 Trend of mean earnings. Index number: 2004=100



Source: own elaborations on AD-SILC data

Table 4.1 Trend of mean annual earnings by employees' characteristics. Index numbers

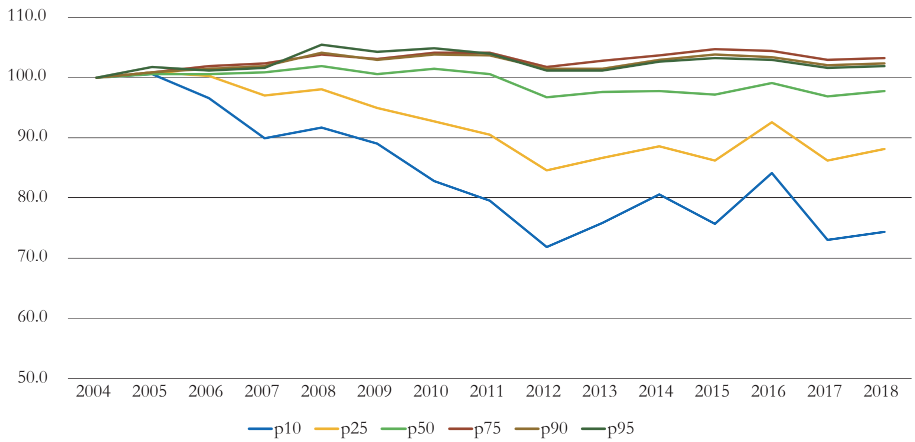
	Sex (Males=100)	Age (50-65=100)			Education (Tertiary=100)		Citizenship (Italian=100)	
	Females	15-29	30-39	40-49	At most low. sec.	Upper sec.	EU	Extra EU
2004	63.8	52.6	78.1	93.3	62.5	81.0	64.7	58.7
2005	63.8	52.2	77.7	92.8	62.6	80.5	63.1	59.1
2006	63.5	52.9	79.8	94.6	60.8	78.5	64.3	59.9
2007	63.9	53.2	81.9	96.2	62.2	80.3	54.9	61.1
2008	63.7	53.8	82.3	97.2	60.3	79.5	56.4	61.0
2009	64.7	53.7	81.5	95.2	59.3	78.8	55.1	56.0
2010	64.8	52.2	81.4	95.0	58.7	78.3	56.4	57.2
2011	64.7	52.2	82.7	95.7	58.4	78.2	57.5	58.3
2012	65.0	51.9	82.4	96.3	57.9	78.3	58.5	57.9
2013	65.5	52.1	81.5	96.5	58.2	78.3	59.3	57.7
2014	65.5	51.1	80.6	95.7	58.2	78.3	60.0	58.4
2015	65.2	50.9	81.2	96.9	58.1	78.4	62.3	60.0
2016	65.7	51.1	81.1	96.1	58.3	78.0	63.1	60.4
2017	65.2	49.8	80.8	96.5	58.2	77.7	62.9	61.3
2018	65.1	50.6	82.0	97.1	57.4	77.5	65.1	62.3

Source: own elaborations on AD-SILC data

4.5 Trend of percentiles of earnings distribution

A standard tool to provide a quick look at earnings distribution is the computing levels and trends of the main percentiles of distribution. Looking at percentiles of the annual earnings distribution (Figure 4.9, where percentage changes compared to 2004 are shown) a substantial increase in inequality emerges, since bottom percentiles clearly deteriorated throughout the whole observed period (despite a slight recovery from 2012).

Figure 4.9 Trend of percentiles of annual earnings. Index number: 2004=100



Source: own elaborations on AD-SILC data

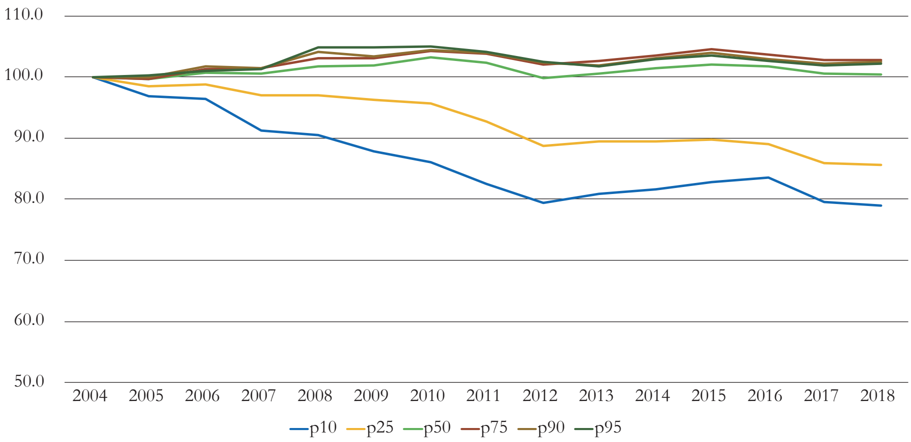
The median also reduced from 2011 onwards, resulting in a 2018 value still lower than the 2004 one. The highest percentiles have remained constant – or have been characterised by a slight increase – since 2008[7].

Interestingly a similar trend of percentiles emerges if we focus on weekly wages (Figure 4.10), thereby eliminating the influence of the different number of weeks worked during a year that, as remarked, mostly depended on the spread of fixed-term arrangements.

Differently, the picture completely changes when we look at the percentiles of the distribution of weekly wages of 'strong workers', i.e. those working with a full-time arrangement (Figure 4.11). Actually, all percentiles were characterised by a slight (but non-negligible) increase along the observed period, while only the poorest decile was characterised, in real terms, by a reduction from 2009 followed by a recovery from 2013. However, as clearly pointed out, trends of 'strong workers' (i.e. the proxy of hourly wage) are not representative of what has happened in recent decades in the Italian labour market, since this market has been characterised by a very intense growth of involuntary (an less paid) part-time jobs.

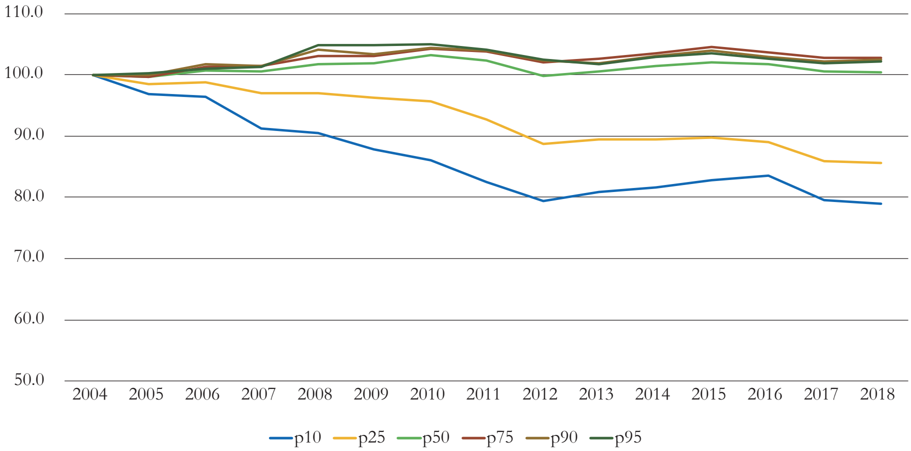
[7] Trends of bottom percentiles might underestimate trends of earnings inequality, since in our analyses we are only considering private employees, thereby excluding the so-called para-subordinate workers – i.e. dependent self-employed, those working as self-employed in legal terms but 'economically dependent'on a single client. Para-subordinate arrangements were often used until recent years as a low-cost type of contract for replacing low-paid employees (Raitano 2018).

Figure 4.10 Trend of percentiles of weekly wages. Index number: 2004=100



Source: own elaborations on AD-SILC data

Figure 4.11 Trend of percentiles of weekly wages of full-timers. Index number: 2004=100



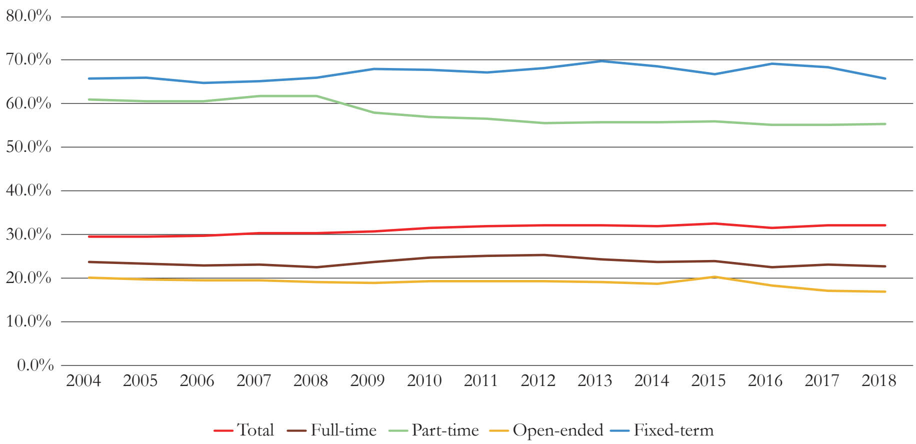
Source: own elaborations on AD-SILC data

4.6 Trend of the in-work poverty risk

The increasing spread of non-standard contracts affects mostly the lowest tail of earnings distribution. For a better assessment of what happened in the bottom tail, we computed an index of 'in-work poverty' risk that captures the incidence of low pay. In what follows, consistently with the standard definition of 'at risk of poverty', we define as 'in-work poor' those employees who earn in a year less than 60% of the median of annual gross earnings (approximately 18,300 euros in 2018, for an 'in-work poverty line' around 11,000 euros).

The share of employees earning less than our 'low pay' threshold slightly increased over the period from 29.6% in 2004 to approximately 32% (Figure 4.12). This is further evidence of the increasing inequalities in the Italian labour market. Not surprisingly, 'in-work poverty risks' are dramatically higher for part-timers (55.4% in 2018) and fixed-term employees (65.8% in 2018) compared to the remaining categories, even though the incidence for the latter is not negligible (22.7% and 17.0% for full-timers and open-ended employees in 2018, respectively). It is also worthy of mention that the incidence of in-work poverty risks among part-timers has reduced steadily from 2008. However, this descending trend, rather than signal an improvement in part-timers wages, might signal a changing composition of part-timers' overtime. The increasing spread of part-time contracts may be related to a rise in the number of part-time contracts with a higher number of contractual hours (and thus with an increase in wages). Unfortunately, we cannot test this hypothesis, since we do not have contractual working hours in our dataset.

Figure 4.12 Incidence of in-work poverty by contractual arrangement



Source: own elaborations on AD-SILC data

As expected, in-work poverty risks are higher for the less advantaged worker categories (see Tables 4.2 and 4.3), i.e. for females compared to males (42.3% vs 23.9% in 2018), youngsters (51.7% in 2018), non-Italian citizens (approximately 47% in 2018 vs 28.5% for Italian citizens) and low skilled ones (38.8% for those with a lower secondary degree at most compared to 26.2% and 22.6% for upper secondary and tertiary graduates in 2018, respectively). When workers are distinguished by education and age class (Figure 4.13, where data refers to 2018), one notices that low-paid risks are more frequent on average for those aged 15-29 than workers aged over 30. However, risks are not negligible at all ages, especially among those with lower secondary education at most. As already mentioned, in-work poverty risks correlate unambiguously with the type of contractual arrangement, since both part-time and fixed-term contracts expose non-standard employees to higher risks by affecting hours and weeks worked over a year. To shed light on the determinants of the in-work poverty status, we employed logit regression including worker characteristics as independent variables.

Table 4.2 Trend of in-work poverty by gender and age

	Sex		Age class			
	Males	Female	15-29	30-39	40-49	50-65
2004	20.7%	41.8%	41.2%	25.7%	22.9%	28.5%
2005	20.4%	42.0%	42.0%	25.8%	23.3%	28.0%
2006	20.6%	41.8%	42.2%	25.5%	23.6%	28.9%
2007	21.4%	42.2%	44.5%	25.8%	23.7%	29.2%
2008	21.2%	42.2%	44.3%	26.1%	23.5%	30.0%
2009	22.0%	42.1%	44.9%	27.2%	24.2%	30.1%
2010	23.0%	42.4%	47.4%	27.6%	24.5%	30.7%
2011	23.4%	42.5%	48.3%	27.6%	24.9%	31.4%
2012	24.2%	42.2%	49.4%	28.2%	25.3%	31.2%
2013	24.4%	41.7%	49.1%	28.7%	25.4%	31.0%
2014	24.1%	41.7%	49.9%	29.1%	25.1%	30.1%
2015	24.8%	42.2%	50.8%	29.7%	25.2%	31.1%
2016	23.7%	41.6%	50.6%	28.3%	24.6%	29.4%
2017	24.2%	42.2%	52.0%	29.0%	24.1%	29.7%
2018	23.9%	42.3%	51.7%	28.5%	24.2%	29.6%

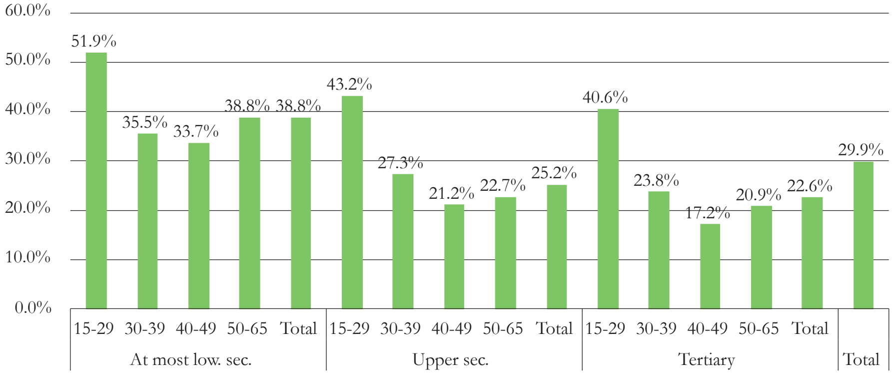
Source: own elaborations on AD-SILC data

Table 4.3 Trend of in-work poverty by education and citizenship

	Education			Citizenship		
	At most low. sec.	Upper sec.	Tertiary	Italian	EU	Extra EU
2004	33.8%	26.4%	28.0%	28.4%	52.0%	51.2%
2005	34.0%	26.3%	28.0%	28.3%	51.2%	50.3%
2006	34.4%	26.4%	27.3%	28.4%	47.6%	49.5%
2007	34.9%	27.3%	27.9%	28.7%	57.8%	49.2%
2008	35.7%	26.8%	27.0%	28.6%	56.3%	49.1%
2009	36.7%	27.0%	26.8%	28.4%	58.2%	54.8%
2010	37.8%	27.6%	27.0%	29.1%	58.2%	54.7%
2011	38.7%	27.7%	26.0%	29.5%	55.6%	53.1%
2012	39.3%	27.9%	26.2%	29.8%	54.9%	53.5%
2013	39.2%	27.7%	26.0%	29.5%	54.3%	53.8%
2014	39.1%	27.4%	25.3%	29.2%	54.1%	53.0%
2015	39.6%	27.7%	25.5%	29.8%	51.3%	51.5%
2016	38.6%	26.6%	24.0%	28.6%	50.4%	50.0%
2017	38.4%	27.0%	23.5%	28.8%	50.0%	49.2%
2018	38.8%	26.2%	22.6%	28.5%	47.1%	47.6%

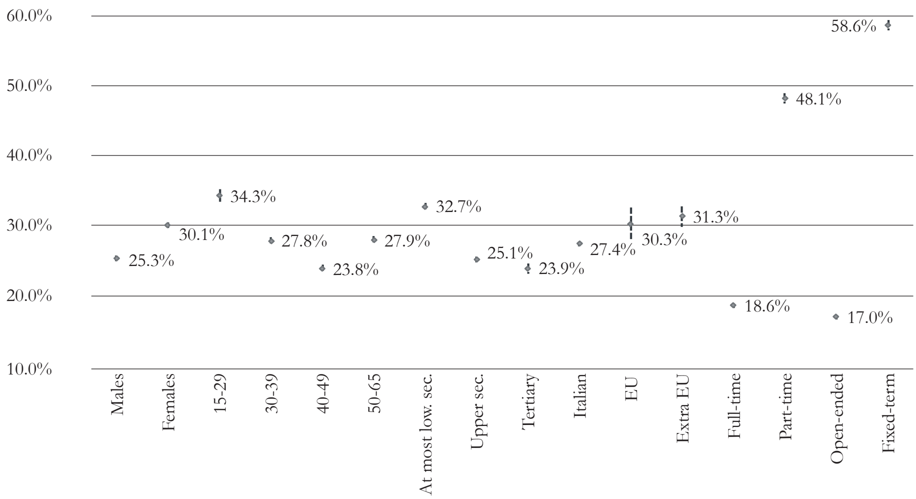
Source: own elaborations on AD-SILC data

Figure 4.13 Share of working poor by education and age class in 2018



Source: own elaborations on AD-SILC data

Figure 4.14 Predicted probability of being in-work poor in 2018 (95% confidence intervals)



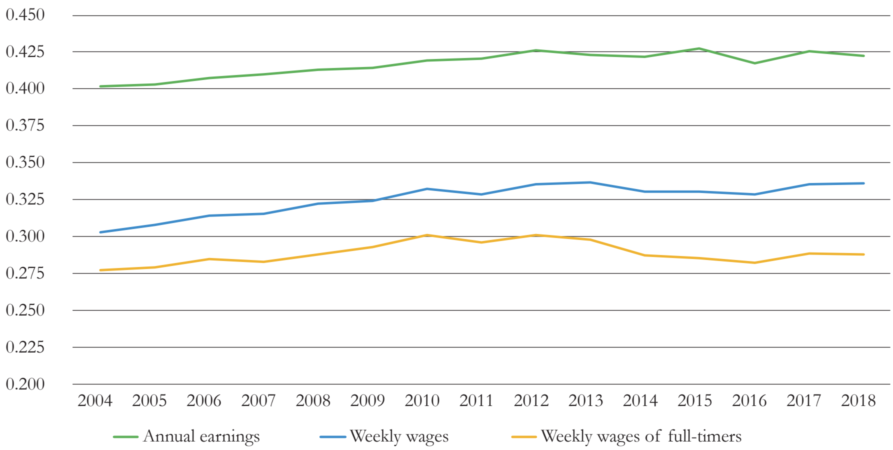
Source: own elaborations on AD-SILC data

Figure 4.14 shows the predicted probabilities of this regression (i.e. the average marginal effects, hereafter AME). All the estimated differences are statistically significant, even though, consistently with our remarks, the main driver of the in-work poverty condition is associated with the contractual arrangement. When controlling for workers' gender, age, education and citizenship, the estimated risk is 2.6 times higher for part-timers than for full-timers (estimated AMEs are 48.1% and 18.6%, respectively), and this risk is 3.5 times higher for fixed-term employees than for open-ended employees (estimated AMEs are 58.6% and 17.0%, respectively).

4.7 Trend of the Gini index of earnings inequality

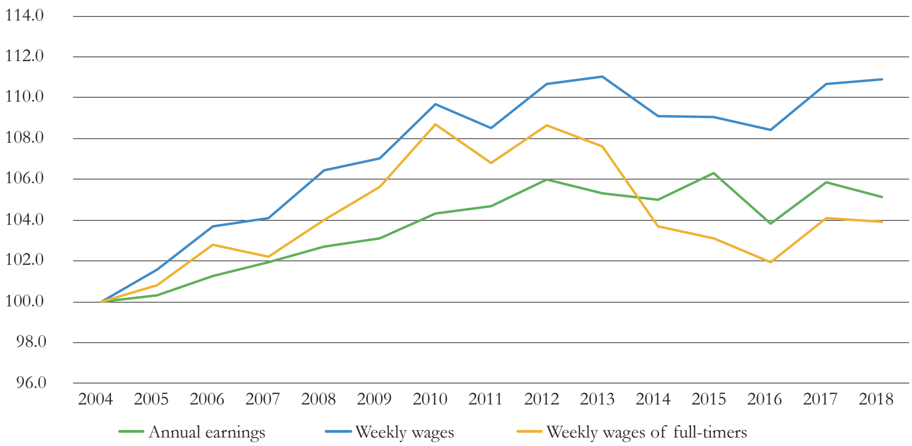
The Gini index of inequality largely rose until the upsurge of the crisis regardless of the earnings dimension considered (see Figure 4.15 and Figure 4.16, where trends compared to 2004 are shown). From 2012 onwards, as far as annual earnings and weekly wages are concerned, the value of the Gini has remained rather constant, while it presents a reduced trend when we focus on weekly wages of full-timers. This suggests that the trend of overall annual earnings inequality from 2014 onwards is mainly due to the further spreading of part-time arrangements. Notice that, not surprisingly, the Gini is higher for annual rather than weekly wages. In fact, the distribution of weeks of work has an amplifying effect on earnings inequality because those who earn lower wages are at greater risk of unemployment.

Figure 4.15 Gini index of gross earnings



Source: own elaborations on AD-SILC data

Figure 4.16 Trend of Gini index of gross earnings. Index number: 2004=100

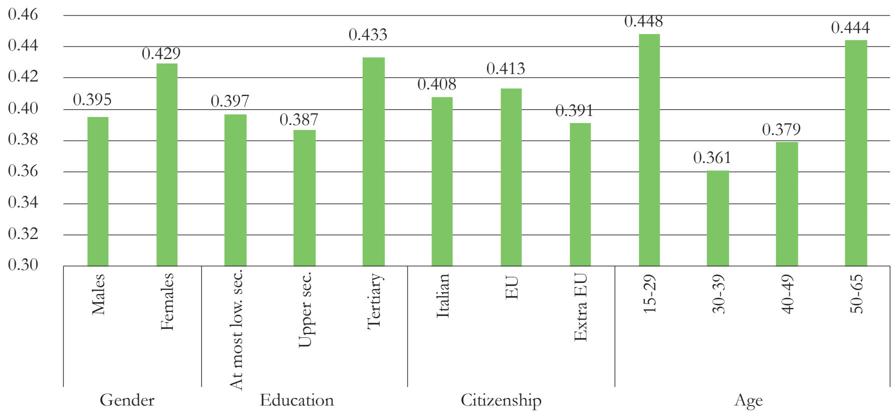


Source: own elaborations on AD-SILC data

As already stated above, earnings inequality is computed by taking into account only the subsample of individuals working during a certain year, thus without considering outflows from the labour force, e.g. due to the economic crisis that started in 2008. As a result, it must be stressed out that focusing solely on employees might alter the picture of the influence of the crisis on workers' living conditions. If the recession had

the effect of removing the least paid workers from the workforce, earnings inequality among the remaining ones would be reduced.
To take into account the whole effect of the crisis on earnings inequality in Italy, Raitano (2019) – using AD-SILC data updated to 2013 – computed the Gini of the annual earnings of private employees, including in computations those workers who were active in a certain year and then became unemployed (with zero earnings) in the following years. The findings are impressive: earnings inequality increased by 21% in Italy in the 2008-2013 period when the involuntary unemployed are taken into account, while earnings inequality within employees (excluding 'zero earners') rose by 2.5% from 2008 to 2013. Likewise, using EU-SILC longitudinal data, Raitano (2016) first focused on the values of gross labour income inequality (i.e. considering incomes both from employment and self-employment) on the sub-sample of individuals interviewed for the whole four-year period in the longitudinal EU-SILC and who were earning a positive labour income in 2008, and then moved to earnings inequality considering also individuals that became unemployed (i.e. earning a zero income) in the following years. Thanks to this approach, a large increase in labour income inequality in the period 2008-2011 emerges in all groups of EU-15 countries, especially in Nordic countries (+12.3%) and in Southern countries (+16.8%).
Finally, we computed the Gini index of annual earnings of different groups of workers in 2018 (Figure 4.17). Interestingly, wage inequality is higher within females than within males and within younger and older workers than within middle-aged employees, while no striking differences appear as to citizenship. Interestingly, inequality within tertiary graduates is the highest – it is also higher than the Gini computed on total workers – which shows the existence of vast heterogeneity in high skilled wages. In the next section, we will assess the role played by the various workers' features as a driver of earnings inequality.

Figure 4.17 Gini index of gross annual earnings by employees' characteristics in 2018



Source: own elaborations on AD-SILC data

4.8 Decompositions of trends of earnings inequality

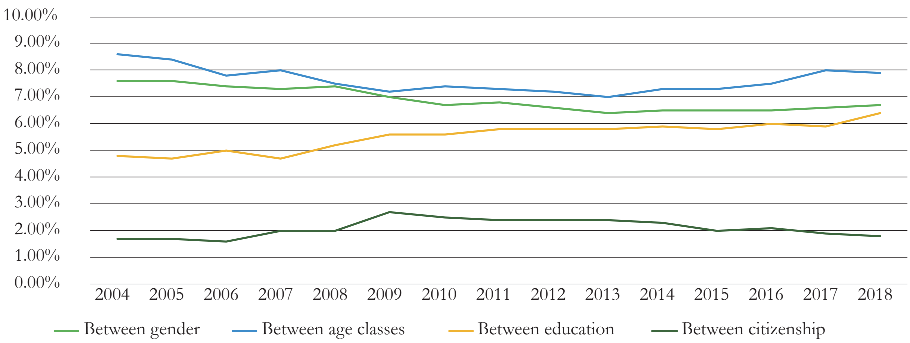
To better understand the role of workers' features as a driver of annual earnings inequality, we split the workers in subgroups according to their features and assess – through a decomposition exercise by subgroups of workers – the relative size of the inequality that emerges within and between the various groups. By doing so, we can measure the share of total inequality due to differences in mean earnings between the various groups – hereafter 'between inequality' – and the share of inequality not due to mean differences between subgroups, which refers to individuals belonging to the same subgroup – hereafter 'within inequality'.

To this aim, we use the Theil index of inequality. Unlike the Gini index, it is perfectly decomposable among groups, since it is expressed as the sum of between and within inequality. The between groups inequality is computed through a counterfactual distribution imputing the mean wage of the *j-th* group to all the individuals who fall in that specific group, while the within groups inequality is the weighted average of the inequality within each group[8].

By distinguishing workers by gender, four age groups, three educational groups and three citizenship groups, we find that no distinction alone can explain a relevant share of annual earnings inequality (Figure 4.18). For all types of subgroups, the share of inequality explained by the between component is always below 10%.

In more detail, the wage gender gap – which is however a crucial issue that has to be tackled by policymakers – explains no more that 7.5% of total inequality, and the share of the between gender inequality over total inequality has reduced overtime, with an increase in earnings inequality within both males and females.

Figure 4.18 Share of between inequality of gross annual earnings by employees' characteristics

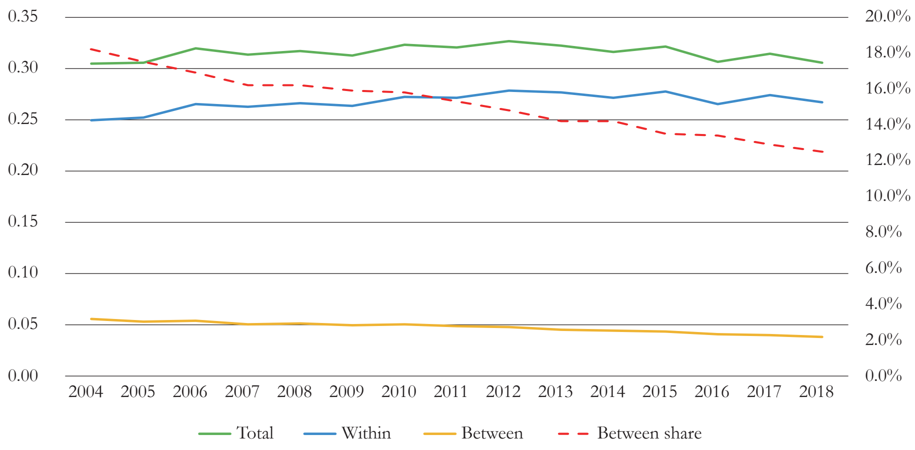


Source: own elaborations on AD-SILC data

8 In the Theil decomposition, inequality within each group is weighted by the relative income earned by each group (Cowell 1995).

The share of inequality explained by mean differences between differently educated workers is rather limited (around 5-6%), even though this share slightly rose during the time span chosen. However, before arguing that the role of education – even if limited – has increased overtime, consistently with the theoretical expectations expressed in the Introduction, further analyses are needed.
To better break down the role played by the intersection between education and age, we interacted the three educational groups and the four age groups together to obtain twelve distinct subgroups (Figure 4.19). As expected, with a higher number of subgroups, we find that the share of total inequality explained by mean differences between these subgroups is higher than the share explained by age or education alone. However, the share of total inequality explained by the 'between age and education component' has reduced overtime. As shown in Figure 4.19 (on the left axis), the trend of total inequality is entirely due to the trend of the within component. This result shows that inequality within workers with the same education and belonging to the same age class have risen overtime relatively more than mean differences between the various subgroups defined interacting education and age. Therefore, the discussed evidence clashes with the usual idea that inequality is increasing merely because of increasing skills and age divides.

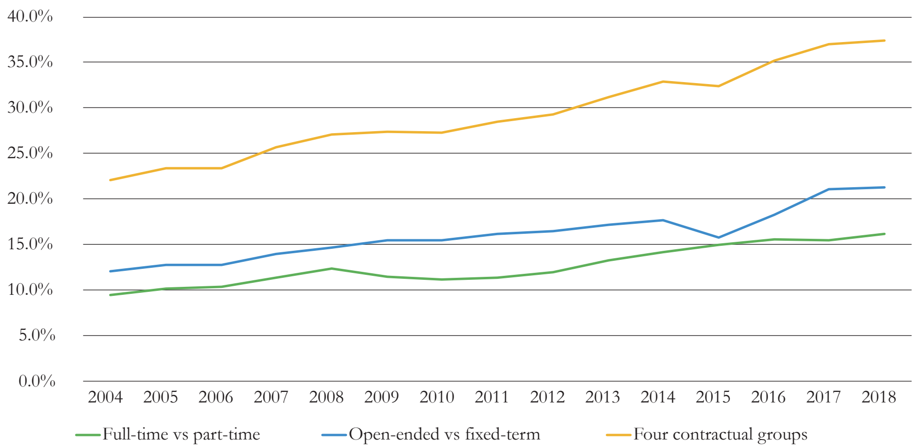
Figure 4.19 Theil decompositions by subgroups given by the combination of education and age groups: total values on the left axis; between share on the right axis



Source: own elaborations on AD-SILC data

To evaluate the role played by further possible drivers of earnings inequality, we focused on contractual arrangements, i.e. on dimensions strictly dependent on the type of labour market regulation (Figure 4.20). By splitting workers by the hourly dimensions (full- versus part-time) or by the contract duration (open-ended versus fixed-term arrangement), we find that the role of between differences has largely increased over time, which shows that a good portion of the within inequalities shown in previous figures is related to contractual segmentation among workers belonging to the same subgroup. Furthermore, when we split workers in four subgroups defined by interacting the two dummies – full- and part-time and open-ended and fixed-term arrangement – we find a massive increase in the share of annual earnings inequality attributable to the contractual dimension: the share of total inequality attributable to mean differences between the four contractual subgroups has indeed increased from 22.1% to 37.4% from 2004 to 2018.

Figure 4.20 Theil decompositions by subgroups of contractual arrangements: between inequality share



Source: own elaborations on AD-SILC data

This evidence suggests that labour market reforms, which contributed to increasing the spread of non-standard employment arrangements, might have contributed to the trend of earnings inequality much more than the determinants typically pointed out by the economic literature, i.e. workers' skills.

However, one may argue that individual skills are different from mere educational attainment, as well as that individual skills determine the type of contractual arrangement attained by the worker. Despite their impressiveness, our decompositions are not enough to argue that skills and human capital are not a major driver of the trend of earnings inequality in Italy. The AD-SILC dataset only includes information on educational attainment, which may be considered as a poor proxy of human capital. The dataset does not include further proxies of human capital and skills (e.g. the quality of the attained education, the specific worker's competences, motivation and effort), which remains unobservable in our empirical setting. Nevertheless, we might suppose that these unobservable skills are likely to become observable by employers and therefore can be captured also in terms of better hiring chances, contractual arrangements, occupation and promotion. Accordingly, we might argue that, among workers with the same educational attainment level, those with better skills should be hired in more rewarding industries and firms with better contractual arrangements (i.e. full-time instead than part-time), and should achieve a better occupational status (i.e. manager or white-collar rather than blue-collar).

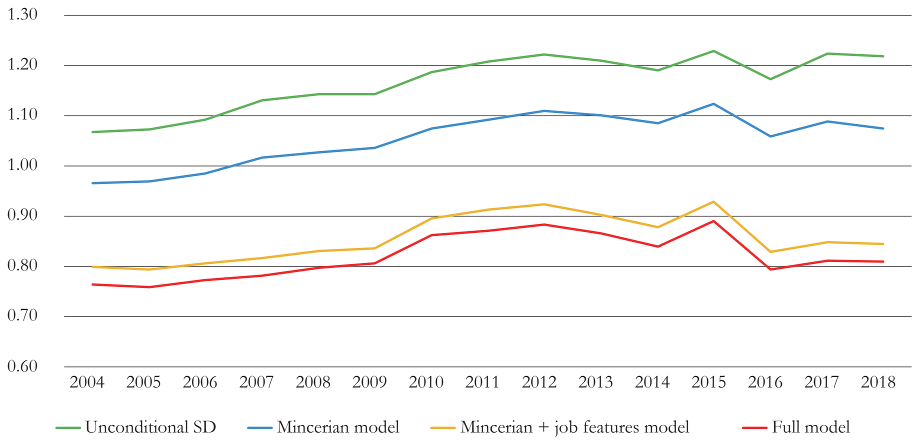
In other terms, workers' unobservable skills could affect several dimensions of the working career that are correlated with wages and, as a result, could affect wage inequality. Higher returns to human capital could then manifest themselves as an improvement in such dimensions, and the explanation of rising inequality based upon human capital and skill endowments could be vindicated.

To evaluate such hypothesis, following Card *et al.* (2013), we carried out OLS regressions of annual earnings and computed the standard deviation of the residuals (the RMSE) as an index of residual inequality in order to verify the extent to which individual wage gaps not explained by observed characteristics might be associated with workers' skills other than education (see Figure 4.21, where we also show the unconditional standard deviation of annual earnings to allow comparisons).

Estimated models in Figure 4.21 include an increasing number of covariates: in the first 'Mincerian' model, we only check for basic demographic characteristics (gender, age, citizenship and education); then we add job features (occupation and contractual dummies); finally, we also add dozens of covariates about firms' characteristics and province fixed effects in the 'Full model'. Compared with the unconditional standard deviation of annual earnings, the higher the number of covariates, the lower the standard deviation of the residuals. However, the reduction in the RMSE is rather limited even in the 'Full model' where we include around hundreds of covariates. More importantly, and consistently with the trend of between education inequality shown previously, the trend of RMSE is increasing over the whole observation period until 2016. Furthermore, it is very similar to the trend of the unconditional standard deviation of annual earnings, pointing out that also these further covariates that might mask unobservable components of workers' skills and human capital do not explain the increasing trend of earnings inequality in Italy since 2004[9]. Such evidence does not lend any support to the idea that the size and the dynamics of earnings inequality mainly depends on workers' skills, as proxied by educational attainments and by other workers' and firms' characteristics that might capture unobservable components of workers' human capital.

9 Franzini and Raitano (2019) carried out a similar exercise for the 1990-2013 period.

Figure 4.21 RMSE of OLS estimates of annual earnings[1]



[1] The following covariates are sequentially included in the various model: dummies on gender, education and citizenship in addition to age and age squared in the 'Mincerian model'; dummies on occupation, full-time contract and open-ended contract in the 'Mincerian + job features' model; Nace 2-digit and province fixed effects in addition to dummies on firms' types and a cubic polynomial on firm's size are added in the 'Full model'.
Source: own elaborations on AD-SILC data

However, a different picture emerges for 2017 and 2018, since residual inequality flattens when we also check for education, job and firms features. In future works, we will try to test whether the stagnancy of residual inequality from 2017 will be a persistent feature of the Italian labour market and what the determinants of this possible stagnancy are.

Conclusions

Following suggestions by Franzini and Raitano (2019), we can conclude that education is clearly rewarded in Italy, as confirmed by mean earnings by education, but large differences within similarly educated workers also emerge and the drivers of these 'within education' differences need to be more carefully investigated by the economic literature.

The trust on explanations of inequality based on individual skills is in some sense reassuring and provides clear policy implications. Markets efficiently pay workers according to their skills, although the latter are unevenly distributed in the population, mostly because of the different human capital accumulation (which is crucially constrained by the family background). Thus, policies should mainly favour the skill endowment of those coming from less advantaged backgrounds. However, in our opinion, the empirical and theoretical literature should pay more attention to drivers of the 'within inequality' (e.g. unobservable abilities, better quality of the education, luck, soft skills, social connections), especially at the top of the income ladder, in order

to assess the acceptability of earnings gaps within similarly educated individuals in terms of efficiency and equality of opportunity.
Earnings inequality depends on the interactions between several possible drivers and all the processes behind these interactions should be carefully analysed. In particular, an aspect to be investigated more in future research concerns the influence on within inequality of institutional factors, with specific emphasis on the deregulation of the labour market that started in Italy in the mid-1990s and led to an increasing share of precarious workers, even among tertiary graduates.
Finally, as remarked by Franzini and Raitano (2019), it is of crucial importance to identify what determines unexplained wage inequality. Firstly, we could learn more about the actual functioning of labour markets and uncover significant differences, e.g. across countries, regions or industries. Secondly, we could enrich our understanding of how acceptable inequalities are and of their economic and social consequences. Human capital is typically considered acceptable as a cause of inequality if access to education is open to everybody. Its limited importance as a driver of wage inequality would suggest that wage inequality could be much less acceptable than normally believed. Third, knowing more about unexplained wage inequality could help design better policies to cope with inequality.

References

Acemoglu D., Autor D. (2011), Skills, Tasks and Technologies. Implications for Employment and Earnings, in Ashenfelter O., Card D. (eds.), *Handbook of Labor Economics Vol. 4A,* Amsterdam, Elsevier, pp.1043-1171

Atkinson A.B. (2015), *Inequality. What can be done?,* Cambridge MA, Harvard University Press

Atkinson A.B., Piketty T., Saez E. (2011), Top Incomes in the Long Run of History, *Journal of Economic Literature*, 49, n.1, pp.3-71

Autor D., Dorn D., Hanson G. (2013), The China Syndrome. Local Labor Market Effects of Import Competition in the United States, *American Economic Review*, 103, n.6, pp.2121-2168

Autor D., Katz L., Kearney M. (2006), The Polarization of U.S. Labor Market, *American Economic Review*, 96, n.2, pp.184-194

Baccaro L., Howell C. (2017), *Trajectories of neoliberal transformation. European industrial relations since the 1970s*, Cambridge, Cambridge University Press,

Bound J., Johnson G. (1992), Changes in the Structure of Wages in the 1980s. An Evaluation of Alternative Explanations, *American Economic Review*, 82, n.3, pp.371-392

Bourguignon F. (2018), World changes in inequality: an overview of facts, causes, consequences and policies, *CESifo Economic Studies*, 64, n.3, pp.345-370

Burkhauser R., Hérault N., Jenkins S., Wilkins R. (2018), Survey under-coverage of top incomes and estimation of inequality: what is the role of the UK's SPI adjustment?, *Fiscal Studies*, 39, n.2, pp.213-240

Canberra Group (2011), *Handbook on Household Income Statistics*, New York and Geneva, United Nations

Card D., Heining J., Kline P. (2013), Workplace Heterogeneity and the Rise of West German Wage Inequality, *The Quarterly Journal of Economics*, 128, n.3, pp.967-1015

Cortes M., Jaimovich N., Nekarda C., Siu H. (2014), The who and how of disappearing routine jobs, VoxEU.org, 2 October <https://bit.ly/3eMNtuW>

Das M., Hilgenstock B. (2018), *The Exposure to Routinization. Labor Market Implications for Developed and Developing Economies*, IMF Working Paper n.135, Washington, International Monetary Fund

Franzini M., Raitano M. (2019), Earnings inequality and workers' skills in Italy, *Structural Change and Economic Dynamics*, 51, issue C, pp.215-224

Goos M., Manning A. (2007), Lousy and lovely jobs. The rising polarization of work in Britain, *The Review of Economics and Statistics*, 89, n.1, pp.118-133

Jessoula M., Raitano M. (2017), Italian pensions from "vices" to challenges: assessing actuarial multi-pillarization twenty years later, in Natali D. (ed.), *The New Pension Mix in Europe. Recent Reforms, their Distributional Effects and Political Dynamics,* Bern, Peter Lang Publishing

Katz L., Murphy K. (1992), Changes in Relative Wages, 1963-87. Supply and Demand Factors, *Quarterly Journal of Economics*, 107, n.1, pp.35-78

Milanovic B. (2016), *Global Inequality. A New Approach for the Age of Globalization*, Cambridge MA, Harvard University Press

Naticchioni P., Raitano M. (2019), Le tendenze di lungo periodo della distribuzione personale dei redditi individuali da lavoro in Italia, in INPS (ed.), *XVIII Rapporto Annuale*, Roma, INPS

OECD (2008), *Growing Unequal*, Paris, OECD Publishing

OECD (2011), *Divided we Stand. Why Inequality Keeps Rising*, Paris, OECD Publishing

OECD (2015), *In It Together. Why Less Inequality Benefits All*, Paris, OECD Publishing

OECD (2016), *Income inequality remains high in the face of weak recovery*, Income Inequality Update, Paris, OECD Publishing.

Piketty T. (2014), *Capital in the Twenty-First Century*, Cambridge MA, Harvard University Press

Raitano M. (2016), Income inequality in Europe since the crisis, *Intereconomics-Review of European Economic Policy*, 51, n.2, pp.67-72

Raitano M. (2018), Italy. Para-subordinate workers and their social protection, in OECD (ed.), *The Future of Social Protection: What works for non-standard workers?*, Paris, OECD Publishing, pp.145-170

Raitano M. (2019), Trends and structural determinants of income inequality. An overview, in European Commission (ed.), *Addressing Inequalities. A Seminar of Workshops*, Luxembourg, Publications Office of the European Union

Salverda W., Nolan B., Checchi D., Marx I., McKnight A., Tóth I., van de Werfhorst H. (eds.) (2014), *Changing Inequalities in Rich Countries. Analytical and Comparative Perspectives*, Oxford, Oxford University Press

Salverda W., Nolan B., Smeeding T. (eds.) (2009), *The Oxford Handbook of economic inequality*, Oxford, Oxford University Press

Stiglitz J. (2012), *The Price of Inequality*. New York, Norton

5. Career patterns and notional accumulation of pension contributions in the NDC scheme

Introduction

Since the beginning of the 1990s, the Italian pension system has been characterised by an intense and long-lasting reform process, concerning both public and private schemes (Jessoula and Raitano 2017)[1]. The reform process focused on three main issues. First, it changed the computation formula of the public scheme, which moved from an Earnings Related (ER) formula to an actuarial Notional Defined Contribution (NDC) formula. Second, the retirement age steeply increased and an automatic link between changes in life expectancy and in retirement age was introduced. Third, the reform process established the architecture of the supplementary private pillar that was underdeveloped until 1993.

Most of all, a crucial structural change was introduced by the 'Dini reform' in 1995. By replacing the ER scheme with the NDC, this reform modified the logic of functioning of the pay-as-you-go public pension pillar. As is known, following actuarial rules, NDC benefits are computed on the basis of contributions paid along the whole working life and of the retirement age, guaranteeing on the notional accumulation of contributions a notional annual rate of return (equal to the average nominal GDP growth of the previous five years in Italy)[2]. Upon retirement, the (notionally) accumulated amount of contributions is converted in an annuity applying the so-called transformation coefficients[3]. Note also that the means-tested minimum pension (*integrazione al minimo*)

---

1 Until 1992, the Italian public pension system was considered rather generous in cross-country comparison. Workers could retire when aged 60 (males) and 55 (females) and an early-retirement option – seniority pension – was available when the worker had accrued 35 years of contributions, independently of age. Pension benefits were based on an Earnings Related formula – the benefit amount was based on the number of years of contribution and on final wages (5 for private employees, the last salary for public employees) – and, hence, workers who retired with 35/40 years of contributions received a pension amounting to around 70/80% of the final wage, respectively.

2 The accrual of contributions is notional since – being the Italian public pension scheme pay-as-you-go financed – contributions are used to finance current pension spending.

3 Transformation coefficients are parameters based on the average unisex life expectancy at a certain age, depend on the retirement age (the later the retirement, the higher the pension) and are updated every two years taking into account changes in life expectancy.

is no longer provided in the NDC scheme, and poor elderly people are only entitled to income-tested social allowances (*assegno sociale* and *pensione di cittadinanza*).

However, the phasing-in of the NDC formula has been extremely gradual, since only individuals who started to work from 1996 will receive a benefit entirely based on such formula. Those who had been working for less than 18 years at the end of 1995 will receive a mixed benefit (ER for working years up to 1995 and NDC afterwards), whereas those who had been working at least 18 years at the end of 1995 remained in the ER scheme (excluding possible working years from 2012, which are taken into account according to the NDC formula).

Apart from the main need of taking under control public pension spending thanks to the sort of automatic adjustment in the intertemporal pension budget guaranteed by a proper NDC formula (Holzmann and Palmer 2006), the introduction of the NDC in Italy was also due to the willingness of deleting some manifest inequities related to the ER formula. ER pensions were related to final wages, independently of the retirement age and of the amount of contributions paid during the whole career, thereby giving rise to very different rates of return on contributions paid into the pension system, favouring self-employed workers (who paid a reduced contribution rate), those that had steeper age-earnings dynamics, and those who retired at a younger age. On the contrary, in the NDC scheme, as mentioned, the benefit is strictly related to the amount of contributions paid over the whole career, its value depends, through actuarial mechanisms, on the retirement age, and all individuals earn the same rate of return on contributions, which is linked to GDP growth rate. Therefore, the technicalities behind the NDC scheme remove both the incentives to early retirement and the inequities that existed in the previous ER scheme.

However, – differently from the previous ER scheme, which, linking pensions to final wages, ensured against risks occurred in a long phase of the working lives – being in the labour market for many years no longer automatically guarantees adequate pensions in the NDC. Due to the interactions of possible adverse events during the working career (i.e. low contribution rates, periods with involuntary part-time jobs, low earnings and frequent unemployment spells without being entitled to notional contributions), also individuals active for a long span of their lives might receive modest NDC pension incomes when retired. Actually, apart from safety nets guaranteed by means of tested minimum income schemes paid to poor elderly, the actuarial rules behind the NDC scheme exclude instead any form of redistribution within pensioners in Italy. Therefore, the Italian public pension system is becoming merely a mirror of individuals' outcomes in the labour market.

It should be pointed out that the risk of modest pensions is not merely brought about by the NDC formula (at the current retirement age, until the 1980s the NDC would have paid benefits higher than under ER rules), but it depends on the coexistence of strict actuarial rules (preventing from any kind of risk sharing among the population), the low

GDP growth rates that have characterised Italy since the beginning of the 21st century and a labour market characterised by increasing inequalities, as shown in Chapter 4.
Nevertheless, the steep increases in retirement age established by the 2009-2011 reforms should allow workers with full careers (i.e. without serious interruptions of the working activity) to receive adequate benefits since, in an NDC system, the higher the retirement age, the higher the pension benefit. Therefore, the steep increase in retirement age established by these reforms should increase the expected public pension benefits of those workers characterised by successful careers, thereby greatly reducing their incentives to contribute to supplementary funds voluntarily. Actually, simulations on expected pensions (e.g. Raitano 2017) confirm that individuals able to spend a long working career (lasting more than 40 years) will receive upon retirement a benefit adequate to their previous wage, with a relatively high replacement rate, around 80%. However, the same simulations also show that individuals with unsuccessful careers – due to low wages and/or frequent unemployment spells – risk receiving, upon retirement, a rather low pension with a value not close to the social assistance benefit (*assegno sociale*).
Therefore, the main issue to be inquired concerns how many individuals will be characterised by unfortunate working careers. At the moment, we are not able to answer this crucial question, since the NDC scheme was introduced in 1996 and no individuals entirely belonging to that scheme have so far spent a whole working career. In this project, we aim at coping with this issue by carrying out dynamic microsimulations of future working careers to compute the notional accumulation of contributions over the whole working life. Relatedly, in this chapter, making use of the AD-SILC dataset, we present the descriptive evidence about the actual labour market outcomes achieved by the individuals belonging entirely to the NDC scheme in the beginning phase of their working career.
The AD-SILC dataset allows us to observe in a great detail the working histories of individuals entered in the labour market since 1996, i.e. of the cohorts of workers whose pension benefits will be entirely based on the NDC formula. In particular, since it records, for every working relationship (or periods spent receiving allowances and indemnities) experienced during a year, duration (in weeks), total gross earnings (also including possible allowances for maternity, sickness or job suspension and an unemployment benefits, which are guaranteed by notional pension contributions) and the specific public pension fund which the worker pays contributions to, it permits to exactly measuring the amount of contributions paid since pension contribution rates differ across workers' categories (i.e. between employees, self-employed and para-subordinate workers). Hence, observing the working histories of individuals belonging to the NDC scheme from entry in the labour market up to 2018 helps us understand how many individuals risk receiving poor pension benefits in the future, if the pattern of their career will not improve in the next decades.

In this sense, it seems very fruitful to carry out the analysis of possible inadequacy risks by separately studying the pattern of the three factors that may determine low pension prospects, i.e., low wages (also due to involuntary part-time arrangements), reduced contribution rates – which characterise the self-employed and, until 2018, para-subordinate workers (Raitano 2018) –, high frequency of job interruptions, especially when notional contributions for periods spent without working (i.e., for unemployment, sickness, maternity or job suspension) are not guaranteed. Finally, as a summary measure of the success of the part of the working career that we are able to observe in our dataset, it is worth observing the whole amount of contributions paid by individuals enrolled to the NDC scheme up until the end of 2018.

In more detail, we carry out two sets of analyses on two different AD-SILC subsamples. First, we consider the sample of 8,528 individuals entered in the labour market in the 1996-1998 three-year period and we observe the first 20 years of their working career after the entry year (i.e., in 1997-2016 for those entered in 1996, in 1998-2017 for those entered in 1997 and so on). Second, we consider all cohorts of workers who entered in the labour market from 1996 to 2013 – the sample is composed by 43,112 individuals – and observe their labour market outcomes in the 5-year period after entry (i.e., in 1997-2001 for those entered in 1996, in 1998-2002 for those entered in 1997 and so on)[4]. Note that, being interested in assessing the adequacy of future pensions, we do not include non-Italian citizens in our samples since we do not know if and how much pension contributions they have accrued abroad.

The two sets of analyses have complementary advantages. On the one hand, observing the first cohorts belonging to the NDC scheme allows us to observe the working career pattern and the (notional) accumulation of contributions in a 20-year period, which is a very large span of the working career. On the other hand, observing 5-year career patterns of almost all cohorts entered in the labour market since 1996 allows us to assess whether labour market outcomes in the first phase of the working career changed across cohorts, penalising, in particular, those cohorts of individuals who became active during the recession period that started in 2008.

When analysing the working career pattern over the observed phase of the working career, as mentioned, we focus on three main dimensions of disadvantages:

i. Career fragmentation, i.e. accruing contributions spells for fewer weeks than potential worked weeks due to non-working spells not guaranteed by notional contributions;
ii. Periods spent working through disadvantaged contractual arrangements, which we associate to those arrangements characterised by a reduced contribution rate (an NDC formula proportionally disadvantages indeed those workers burdened by a lower contribution rate); currently, while public and private employees pay a

4 We identify the entry year as the first year with at least 13 worked weeks at an age no higher than 40.

33% contribution rate on their gross wage, self-employed (craftsmen and dealers) and freelance para-subordinate workers are burdened by a 24% and 25% rate, respectively, while para-subordinate collaborators now pay the same contribution rate of the employees, after decades in which their contribution rate was much lower (e.g. 10% in 1996-1998; Raitano 2018);

iii. Risks of low pay, which we summarise with the 'in-work poverty' risk, i.e. the risk of receiving annual earnings lower than 60% of median earnings of private employees (i.e. those paying the full contribution rate) working the whole year with a full-time arrangement[5].

The occurrence of the three listed risks determines the success of the working career and the worker's capacity to pay proper contributions in the NDC scheme. Hence, we can consider the total notional accumulation of pension contributions at the end of the observed phase of the working career as a synthetic measure of the success of the career, which depends on these three risks[6].

To evaluate the spread of the three types of risks and to assess the adequacy of the notional accumulated contributions amount we make use of various indicators[7].

In more detail, to capture risks of career fragmentation, we compute the distribution of the ratio between total weeks spent in the period receiving actual or notional contributions and total possible worked weeks over the period (1040 weeks in the 20-year-period, 260 in the 5-year period).

To capture risks related to disadvantaged contractual arrangements, we compute the ratio between weeks spent working as a para-subordinate worker (i.e. both as a free-lance or a collaborator) and total possible worked weeks over the period, and the ratio between weeks spent working with an arrangement characterised by a reduced contribution rate (i.e. as a para-subordinate worker or a self-employed) and total possible worked weeks over the period.

To summarise risks of in-work poverty, we compute the number of years after the entry in the labour market spent receiving annual labour incomes lower than 60% of the median earnings of private employees employed full-time for the whole year.

---

5 Median annual gross wage of full-timers employed the whole year in the private sectors amounted approximately to 25,000 euros in 2018. Therefore, the in-work poverty line in 2018 was approximately 15,000 euros.

6 Note that workers can 'disappear'for an entire year from INPS archives for various reasons not recorded in the administrative archives: long-term unemployment not covered by notional contributions; voluntary inactivity; suspension of the working activity for studying; entry into undeclared work; transfer abroad. Although the lack of information on the reason for the absence from the archives does not allow judgments on the voluntariness and seriousness of this absence, the absence implies – except for cases of emigration abroad – a contribution gap for social security purposes and, therefore, a risk of insufficient accumulation of contributions at the end of the career. Similarly, any 'gray'work periods, in which part of the remuneration is hidden from the INPS, for example through forms of fictitious part-time, even if they increase the income available to workers in the active phase, involve the absence of contributions and, therefore, more limited future pensions in the NDC scheme.

7 Note that the entry year is not considered to compute these indicators.

Finally, to summarise the success of the working career over the whole observed period, we compute the accumulated contribution amount and, to have an idea of the adequacy of that amount, we present the ratio between the individual accumulation and the accumulation of a reference worker, where we take as a reference an individual always employed as a private employee and earning each year the median wage of full-timers private employees.

5.1 Career patterns along the 20-year period

In this section we use administrative data on working histories in the first 20 years of career (excluding the entry year) of workers who entered the labour market in the period 1996-1998, i.e. the first 3 cohorts of workers fully enrolled to the NDC scheme and find rather concerning results.

Table 5.1 shows mean values at the end of the 20-year period of the indicators explained in the Introduction. As concerns worked weeks (also including weeks giving rise to notional contributions), on average, individuals in our sample paid contributions for approximately 16 out of 20 years, even though, as expected, the ratio between actual worked and potential worked weeks (i.e. 1040) is higher for males than for females and for those with a higher education.

On average, few weeks have instead been spent as a para-subordinate worker, even though, noteworthy, the risk of working with this type of contract with reduced welfare guarantees rises with workers' education. Considering also periods spent working as a 'standard self-employed', on average our sample spent approximately 2 years (i.e. 10.8% of possible total weeks) as a para-subordinate or a self-employed.

Low gross labour incomes (which include also the share of social contributions paid by the worker) are largely diffused within our subsample. Considering an in-work poverty line approximately equal to 15,000 euros in 2018 nominal values, we find that the mean number of years spent as an in-work poor along the 20-year period is around 7 (i.e. the ratio is 35.2%), and it is even higher than 8 and 10 for females and the low-skilled, respectively.

Finally, we find that, on average, at the end of the observed period, workers belonging to our subsample have accrued a notional amount of contributions equal to 82% of the amount of contributions that would have been accumulated in the same period by an employee continuously employed with a wage approximately equal to 26,000 euros in 2018 nominal values (thus paying each year, in 2018 nominal terms, approximately 8,500 euros as pension contributions), whom we will henceforth call 'median employee'. As expected, being related to the different spread of the three aforementioned risks occurring during the working career, our favoured synthetic indicator of the multi-year career success is correlated with individual characteristics and is much higher for males (91.7%) than for females (70.8%), and for tertiary graduates – who, confirming the existence of a mean positive reward for human capital investment accumulate at the end of the period more than the reference 'median employee' (106.9%) – than for those with an upper secondary degree (78.8%) or at most a lower secondary degree (59.5%).

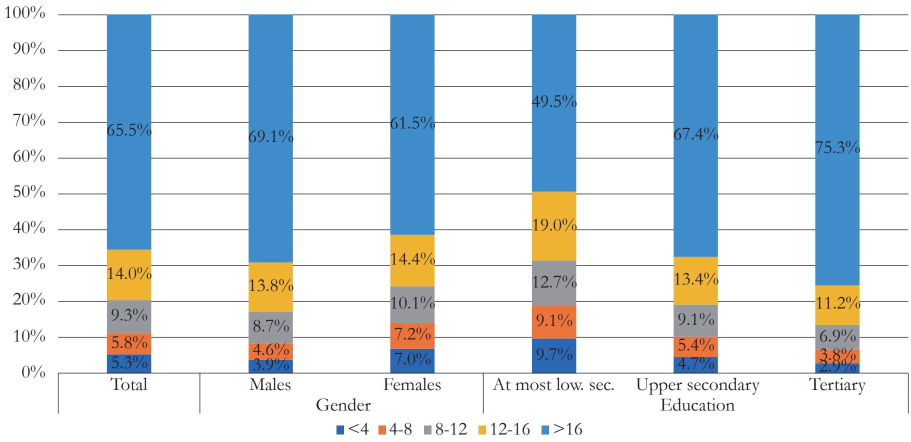
Table 5.1 Summary indicators of the 20-year career pattern: mean values

	Share of worked weeks with respect to total weeks in the period			In-work poverty	Contribution accumulation at the end of the period
	Worked weeks	Weeks as a collaborator or a freelance	Weeks paying a reduced rate	Years spent as an in-work poor	Contrib. w.r.t the reference employee
Total	79.0%	1.9%	10.8%	35.2%	82.0%
Gender					
Males	81.5%	2.2%	12.8%	28.5%	91.7%
Females	76.2%	1.5%	8.5%	42.9%	70.8%
Education					
Low. ed.	69.5%	0.4%	6.2%	50.1%	59.5%
Mid. ed.	80.1%	1.7%	10.0%	32.8%	78.8%
High ed.	84.9%	3.4%	16.4%	27.5%	106.9%

Source: own elaborations on AD-SILC data

However, despite the useful insights provided by Table 5.1, focusing on mean values only is not enough to evaluate the spread of risks of unsuccessful careers – and, then, not adequate contributions accrual – for the cohorts of workers belonging to the NDC scheme. To capture the spread of these risks better, we then show descriptive statistics of the distribution of the aforementioned indicators of the risks in the 20-year working career.

Figure 5.1 Distribution of the number of actual contribution years in the 20-year period



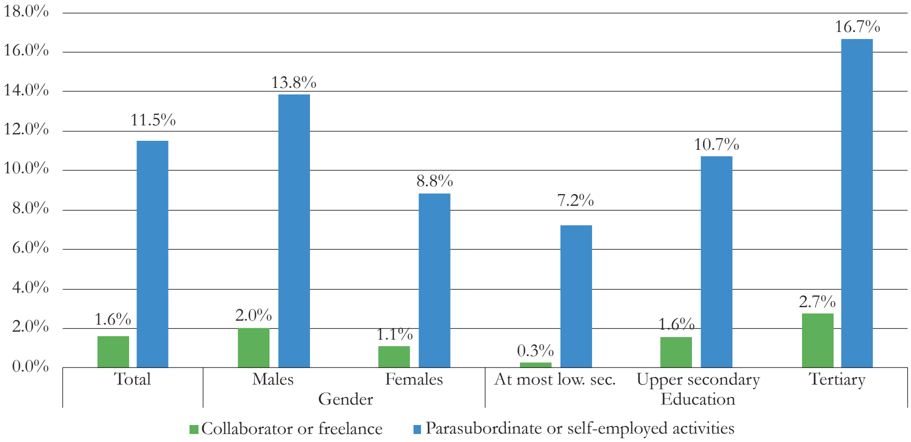
Source: own elaborations on AD-SILC data

First, it has to be noted that a limited share of workers belonging to our subsample (4.5%, 2.7% and 6.7% among males and females, respectively) have a short working career, as they are absent in administrative archives after the 10th year after entry into activity. Remember, however, that, as mentioned, we do not have information in our dataset to disentangle the reason for the disappearance from the workforce registered in INPS archives.

As concerns the spread of the risk of career interruptions (Figure 5.1), we find that a non-negligible share of workers is characterised by a large number of missing contributory periods over the 20-year period: 20.4% of workers have indeed paid contributions for at most 12 out of 20 years. Moreover, as expected, the risk of career fragmentation is higher for females and the low-skilled (24.3% and 31.5% of them, respectively, have paid contributions for no more than 12 out of 20 years).

Contrasting the common wisdom that often indicates the spread of para-subordinate atypical arrangements as the main determinant of future risks of inadequate pensions, Figure 5.2 shows that working as a para-subordinate is not usually a persistent status (1.6% of our sample worked at least 8 out of 20 years as a para-subordinate free-lance or collaborator), whereas – being standard self-employment and professional jobs largely diffused and persistent in Italy – the share of those who worked through contracts characterised by a reduced contribution rate at least 8 of 20 years rises to 11.5% (16.7% within tertiary graduates).

Figure 5.2 Share of workers who paid contributions at a reduced rate for at least 8 out of 20 years



Source: own elaborations on AD-SILC data

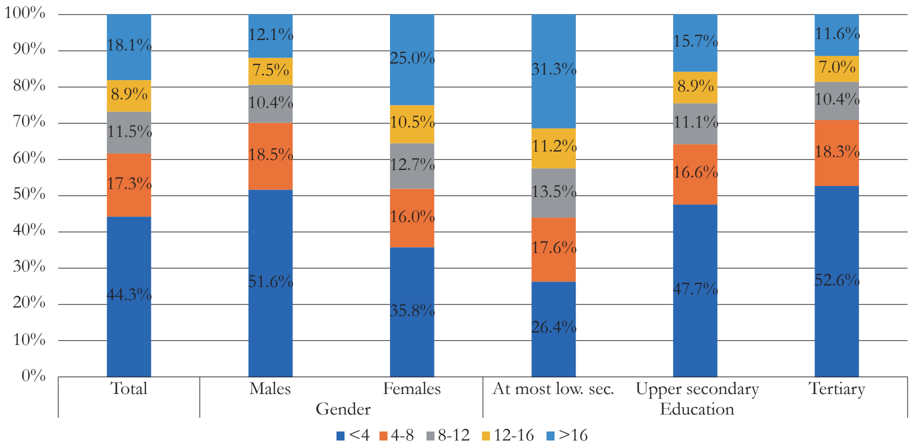
Low wages, also due to the increasing spread of part-time and fixed-term arrangements shown in Chapter 4, characterise for many years a large share of workers who entered the labour market in 1996-1998 (Figure 5.3). Overall, 38.4% of workers earned an annual labour income lower than the in-work poverty line for at least 8 out of 20 years, and this share dramatically increases within females (48.2%) and low-skilled workers (56.0%)[8].

Summing up, risks of future pension inadequacy derive from three factors: low wages, jobs interruptions and low contribution rates. All these elements might engender a limited accumulation of contributions in the NDC scheme. As clarified, as a synthetic risk indicator, it is interesting to show how new entrants compare in terms of accumulation of pension contributions in the observed phase of the working career to the representative 'median employee'.

Hence, comparing the actual accumulation of individuals enrolled in the NDC scheme to the potential accumulation of this hypothetical median worker helps to show how many individual have accumulated, so far, few pension contributions in the initial phase of their career and face the risk of becoming poor pensioners if the career pattern does not improve in next years.

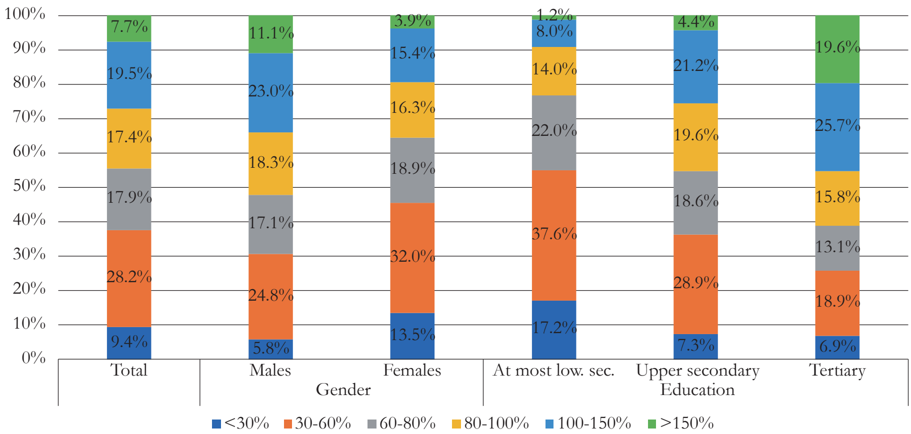
[8] Note that the in-work poverty risk refers both the risk of earnings in a year a positive income but lower than the poverty line or having zero income in that year.

Figure 5.3 Distribution of the number of years spent as an in-work poor over the 20-year period



Source: own elaborations on AD-SILC data

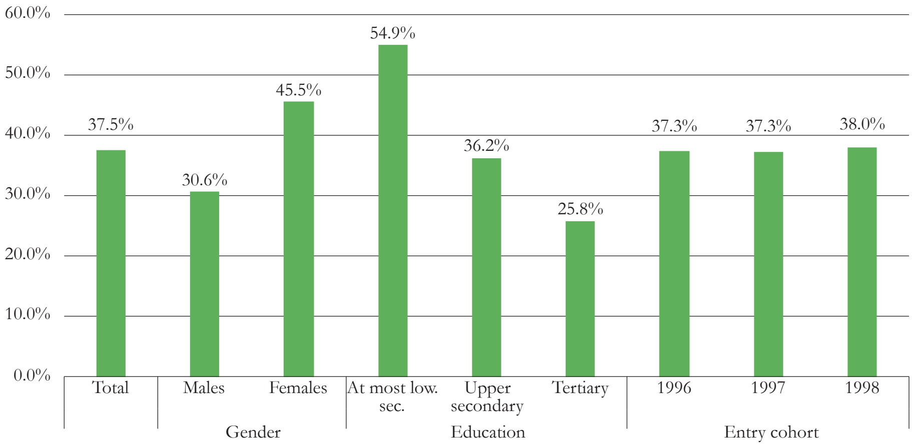
Figure 5.4 Distribution of notional pension accumulation in the 20-year period, with respect to the reference employee



Source: own elaborations on AD-SILC data

Worryingly, we find a large heterogeneity in contributions accrual over the 20-year period (Figure 5.4). Looking at all workers, while 27.2% of them have accrued more than the reference median employee (but only 19.3% and 9.2% among females and the low skilled, respectively), 37.5% have accrued less than 60% of the accumulation of this representative employee. Education is on average rewarded over the long run, since 45.5% of tertiary graduates in our sample has accumulated more than the median employee in the observed period, but, consistently with the picture of a large heterogeneity also within workers with similar characteristics, the share of tertiary graduates with a reduced contributions accrual is not negligible at all (approximately 1 out of 4 has accumulated no more than 60% of the median employee over the 20-year period).

Figure 5.5 Share of workers who accumulated less than 60% of the reference employee in the 20-year period

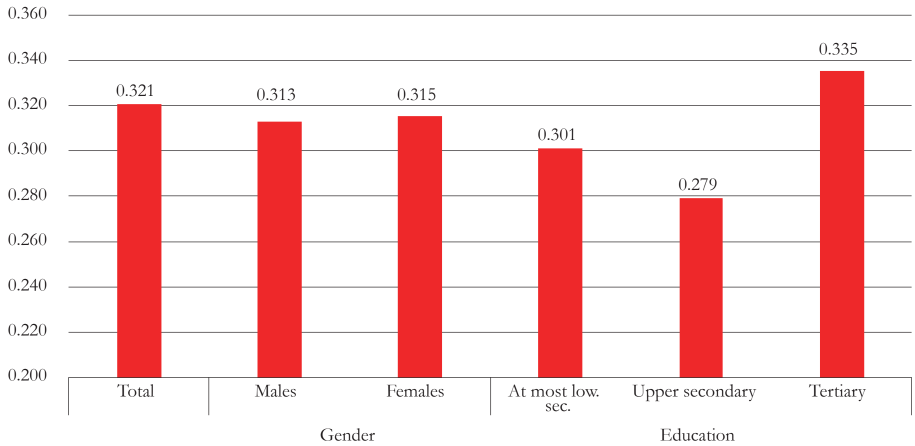


Source: own elaborations on AD-SILC data

Figure 5.5 summarises the risk of low contribution accumulation showing the share of workers – distinguished by gender and education – who accumulated less than 60% of the accumulation of the median employee in the 20-year period. Therefore, these workers, in line with the relative poverty approach, may be considered at risk of poverty when elderly if their career pattern will not significantly improve in the second half of their career.

The picture of a large heterogeneity in career patterns is also confirmed when – as an indicator of multiyear labour income inequality – we compute the Gini index of the accumulation of contributions at the end of the 20-year period (Figure 5.6). Considering that an usual rule of thumb suggests that inequality is high when the Gini value exceeds 0.30, the Gini amounts to 0.321 and, confirming the large heterogeneity within tertiary graduates, the highest value emerges within the highly skilled (0.335).

Figure 5.6 Gini index of the accumulated amount at the end of the 20-year period



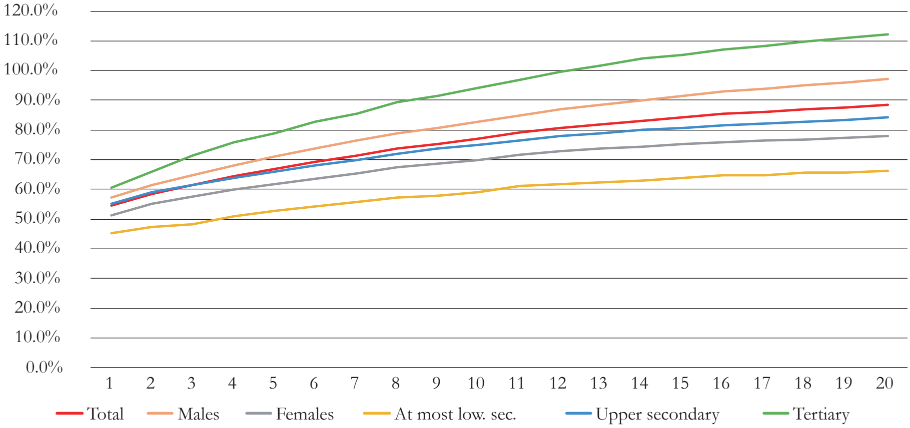
Source: own elaborations on AD-SILC data

However, due to the characteristics of our data, we are able to observe only a limited (even if long) span of the working career while career patterns are usually upward sloped due to increasing employment opportunities and wage premia related to work experience. Hence, one may argue that this worrying picture might largely change in future years when the whole career is observed. To this aim, as mentioned, dynamic microsimulations allow us to complete individual career patterns. However, to take into account the influence of experience patterns over the 20-year period, it is interesting to depict how the mean accumulated amount – expressed with respect to the accumulation of the median employee – evolved at the end of each year from the entry in the labour market (Figure 5.7, where notice that the reference employee, always earning median wage, is by definition characterised by a flat wage pattern over the 20-year period)[9]. As expected, all groups of workers are characterised by an increasing career pattern, showing that risks of inadequate accrual of contributions are clearly overestimated when workers are observed at an overly early stage of their career. However, it is also noteworthy pointing out the heterogeneity of the career patterns (Rubinstein and Weiss 2006). Indeed, the rate of growth of contributions accrual over the 20-year period is much steeper for tertiary graduates and males than for females and the low-skilled, thereby suggesting that the increase in work experience per se might not be enough to bring less advantaged workers out of risks of inadequate pensions when retired.

---

[9] Note that values in Figure 5.7 at the 20th period differ from values shown in Table 5.1 since in Figure 5.7 we only show mean values for workers who actually worked in that year, while in the computations of Table 5.1 we also include those workers who disappeared from the administrative archive before the 20th year.

Figure 5.7 **Mean accumulated amount at the end of each year of the 20-year period, as a share of the accumulation of the reference employee**



Source: own elaborations on AD-SILC data

5.2 Career patterns along the 5-year period

Observing career patterns of individuals who entered the labour market in 1996-1998 has the limit of not allowing us to compare the evolution of these patterns across cohorts, and to highlight the heterogeneity of the impact of the economic crisis across cohorts.

To this aim, in this section we compare the indicators of the career pattern observing the dynamics of the 5-year period after the entry in the labour market of 1996-2013 cohorts. Therefore, if, on the one hand, this comparison does not allow us to provide proper insights about future pension risks, on the other hand it allows us to observe whether a deterioration of early career patterns have emerged in past years, especially due to the upsurge of the economic crisis. Therefore, even if we present summary indicators along the 5-year periods also distinguishing workers by gender and education, in this section we focus on differences by entry cohort.

Table 5.2 Summary indicators of the 5-year career pattern: mean values

	Share of worked weeks with respect to total weeks in the period			In-work poverty	Contribution accumulation at the end of the period
	Worked weeks	Weeks as a collaborator or a freelance	Weeks paying a reduced rate	Years spent as an in-work poor	Contrib. w.r.t the reference employee
Total	75.4%	3.5%	13.0%	48.6%	60.2%
Gender					
Males	76.3%	3.3%	14.0%	43.3%	65.1%
Females	74.6%	3.7%	11.9%	54.1%	55.1%
Education					
Low. ed.	68.1%	0.9%	7.2%	61.5%	48.1%
Mid. ed.	76.3%	2.6%	11.0%	46.9%	59.6%
High ed.	79.0%	6.9%	20.6%	42.6%	69.6%
Entry cohort					
1996	77.4%	5.6%	17.9%	43.4%	62.9%
1997	78.0%	2.5%	14.4%	42.7%	64.1%
1998	78.2%	2.3%	11.1%	43.1%	64.6%
1999	78.3%	2.9%	12.1%	42.6%	65.2%
2000	77.3%	2.4%	11.6%	44.6%	63.7%
2001	77.1%	3.1%	12.6%	45.2%	62.1%
2002	75.2%	3.8%	12.2%	48.9%	59.5%
2003	75.8%	3.5%	12.8%	48.1%	61.0%
2004	76.5%	3.9%	13.5%	47.2%	61.7%
2005	77.7%	3.6%	13.1%	46.8%	62.2%
2006	76.8%	3.8%	13.4%	46.7%	61.5%
2007	72.9%	3.5%	11.8%	55.1%	54.3%
2008	71.4%	3.6%	11.4%	56.1%	54.8%
2009	69.4%	4.5%	14.4%	57.9%	51.5%
2010	70.5%	3.8%	11.7%	56.4%	54.9%
2011	70.8%	3.5%	12.6%	57.1%	54.3%
2012	70.4%	3.7%	13.2%	58.5%	53.8%
2013	74.6%	3.3%	14.2%	55.1%	56.9%

Source: own elaborations on AD-SILC data

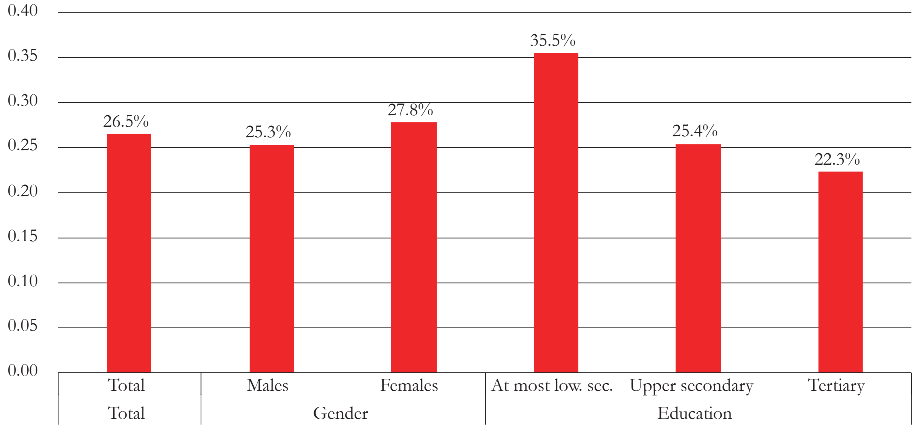
Mean values of all the indicators used in this chapter clearly depict a picture of worsening condition in the first phase of the career for the youngest cohorts, and in particular for the individuals who entered the labour market in the 2007-2012 period and were then more exposed to employability risks due to the effects of the crisis in the starting phase of their career (Table 5.2). For instance, it has to be highlighted that the mean share of contributory weeks on total weeks in the 5-year period (i.e. 260) decreased by approximately 10 p.p., comparing cohorts of workers who entered the labour market at the end of the 20th century and those who entered at the end of the first decade of the 21st century. Likewise, it is striking to compare the mean accrual of contributions in the 5-year period of the 1999 entry cohort (65.2% with respect to the median employee) and the 2009 cohort (51.5%). Therefore, the worrying picture shown in Section 5.1 might further worsen in future years if the career pattern of the cohorts more exposed to the macroeconomic risks engendered by the economic crisis does not greatly improve in future years.

It has also to be noted that the share of workers 'absent' in the 5th year is around 10% in most cohorts, but, as expected, it increases during the crisis from the 2004 cohort (5th year in 2009 – 13%) up to nearly 17% for cohorts 2008-2010 (16.8%, 16.5% and 17.2%, respectively).

A very worrying picture of worsening conditions for the youngest cohorts – even if limited to this short early phase of the career – also emerges when we look at the spread of the indicators of career risks that we consider in this chapter.

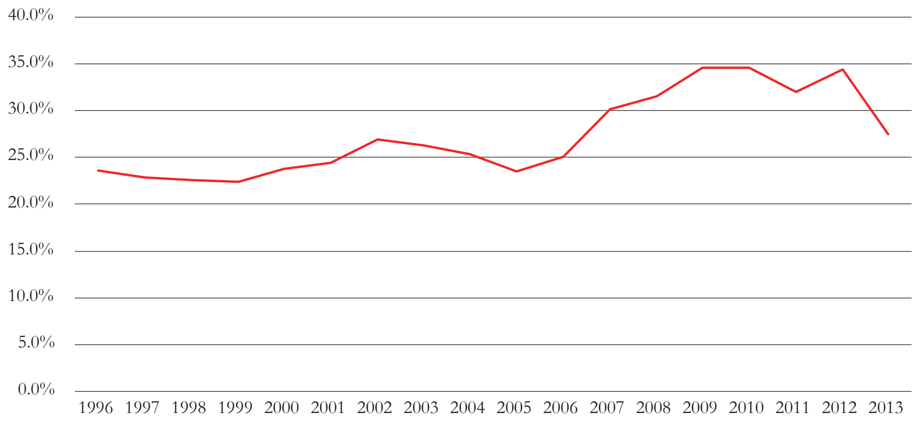
The share of workers who accrued less than 3 years of contributory periods since the year after the labour market entry is 26.5% on average across all cohorts (Fig. 5.8) and rose from approximately 25% up to 35% for the cohorts that started working during the crisis (Figure 5.9). No large differences across cohorts emerge instead as concerns the share of workers who worked for at least 2 out of 5 years with a reduced contribution rate arrangement (Figure 5.10, where a spike in 1996 emerges due to the entry in activity in 1996 of the *Gestione Separata*, which is the fund where para-subordinate freelance and collaborators are obliged to pay pension contributions; Raitano 2018).

Figure 5.8 Share of workers who paid contributions for less than 3 years in the 5-year period



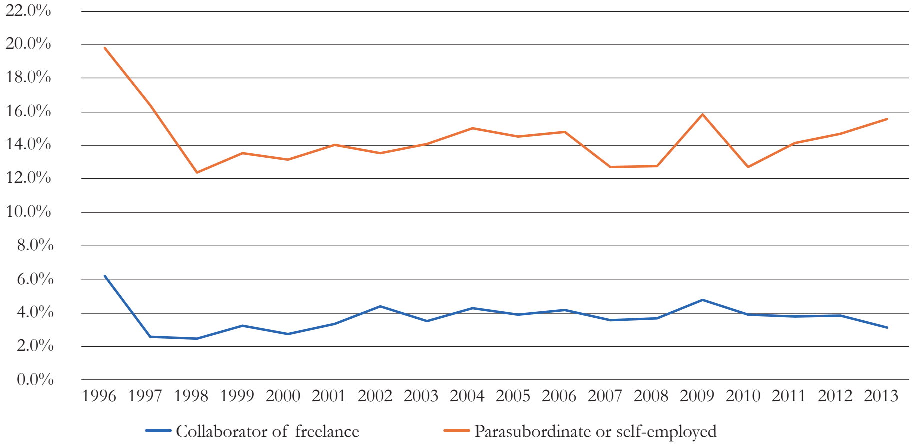
Source: own elaborations on AD-SILC data

Figure 5.9 Share of workers who paid contributions for less than 3 years in the 5-year period, by entry cohort



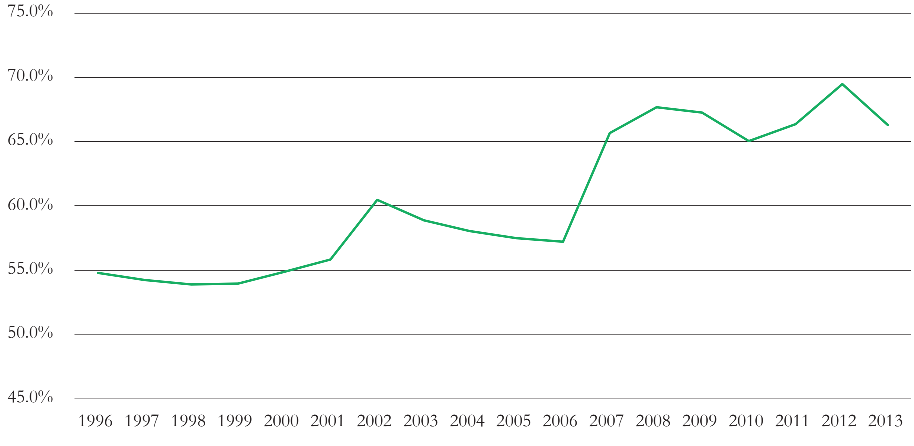
Source: own elaborations on AD-SILC data

Figure 5.10 Share of workers who paid contributions at a reduced rate for at least 2 out of 5 years, by entry cohort



Source: own elaborations on AD-SILC data

Figure 5.11 Share of workers who were in-work poor for at least 3 out of 5 years, by entry cohort



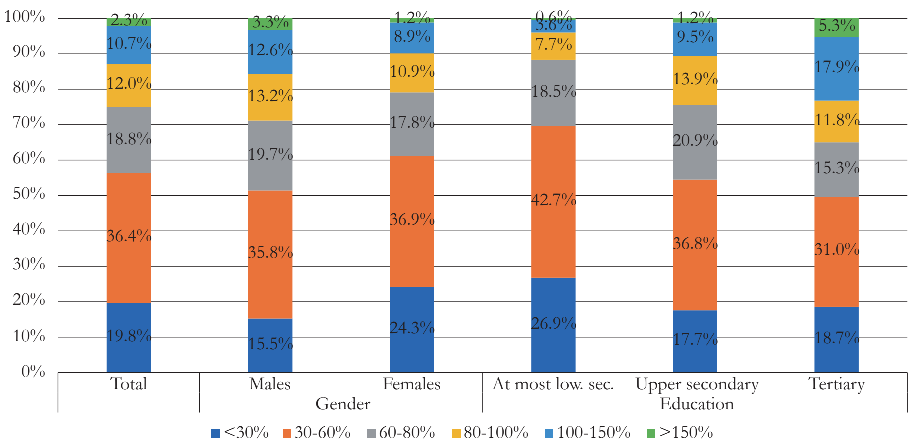
Source: own elaborations on AD-SILC data

In-work poverty risks are largely diffused in the early phase of the career. Nevertheless, the spread of this risk dramatically increased across cohorts (Figure 5.11): the share of workers who spent at least 3 out of 5 years as an in-work poor for at least 3 out of 5 years rose from approximately 55% to 70% moving from the oldest cohorts to the youngest cohorts.

Finally, as concerns the accumulation of pension contributions, assessed with respect to the median employee, we find that among all cohorts, 46.2% of workers have accrued less than 60% of the accumulation of the median employee in the 5-year period (Figure 5.12), but this 'future poverty risk when retired' indicator dramatically rose overtime (Figure 5.13), signalling a reduced capacity of many of those workers belonging to the youngest cohorts to pay an adequate level of contributions in the early phase of the career.

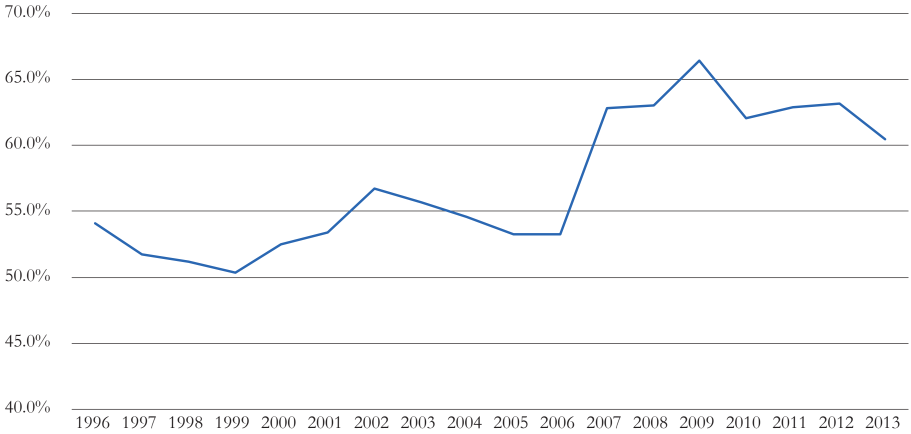
Note also that the heterogeneity within workers has also risen overtime, as summarised by the Gini index of contribution accrual over the 5-year period (Figure 5.14).

Figure 5.12 Distribution of notional pension accumulation in the 5-year period, with respect to the reference employee



Source: own elaborations on AD-SILC data

Figure 5.13 Share of workers who accumulated less than 60% of the accumulation of the reference employee in the 5-year period, by entry cohort



Source: own elaborations on AD-SILC data

Figure 5.14 Gini index of the accumulated amount at the end of the 5-year period, by entry cohort

Source: own elaborations on AD-SILC data

Conclusions

In this chapter, exploiting the richness of the longitudinal AD-SILC dataset, we observed labour market outcomes over a long phase of the working career of the first cohorts of individuals entirely belonging to the NDC pension scheme. We found that – despite the individual career pattern is on average improving overtime – a non-negligible share of individuals is characterised by unsuccessful careers that bring about a limited accrual of contributions in the NDC. Furthermore, comparing for a more limited time span individuals who started working before and during the aftermath of the economic crisis, we found that the exposition of labour market risks has largely increased for younger cohorts. Of course, we are unaware of the future development of the career of these cohorts of workers, which will be nevertheless simulated in this project by exploiting the potentialities of the T-DYMM microsimulation model. However, we can argue that, if the career patterns do not improve in next years, a non-negligible share of workers belonging to the NDC scheme – where pension benefit almost entirely mirrors the outcome of the whole working career – might face the risk of becoming poor pensioners.

Given this context – characterised in Italy by persistent weak macroeconomic performances and increasing inequalities in the labour market – one may thus evaluate from an equity perspective possible limits of a pension computation formula as the NDC, where pensions strictly depend on contributions.

From a normative perspective, those who consider fair – over than actuarially correct – an NDC scheme are implicitly assuming that all individuals have the same opportunities to pay contributions in their working life, regardless of possible circumstances disadvantaging them (e.g. skills, health, the capacity to continue working at older ages, the luck related to the business cycle or to the contractual arrangement obtained). The increasing inequalities among workers – regarding wages, employment opportunities, contractual arrangements – in EU countries and Italy (see Chapter 4.2 and Raitano 2019) might instead bring us to question the idea that pension contributions only depend on individual efforts and only these efforts should be rewarded by the pension system.

Therefore, even if it has the main advantage of stabilising the intertemporal pension budget and correcting some regressive inequities of the previous scheme, the NDC might not be able to correct further types of inequities emerging in the labour market and disadvantaging some groups of workers. Following Roemer (1998), to improve social justice, individuals should be compensated at retirement for their unequal outcomes obtained in the labour – independently of their merits – through a progressive redistribution that directs relatively more resources to those who paid lower contributions during the working life. Apart from safety nets guaranteed by means tested minimum income schemes paid to poor elderly, the actuarial rules behind the NDC scheme exclude instead any form of redistribution within pensioners in Italy.

Therefore, as already remarked, the Italian public pension system is becoming merely a mirror of individuals' outcomes in the labour market.
Furthermore, deviations from actual fairness might also depend on non-casual gaps in life expectancy among individuals and both the amount of the NDC pension and the requirements for having access to old-age pensions or early retirement benefits depend in Italy on the average life expectancy.
Actuarial fairness implies that the internal rate of return that equals the actual value of contributions paid and pensions received over the whole life course is the same for all individuals (Borsch-Supan 2006). However, if life expectancy differs between two individuals and the pension system does not take into account this difference, these individuals will receive a different pension wealth even if they pay the same contributions and the individual with a lower longevity would be penalized. Thus, a social security system, as the Italian NDC, which computes pension benefits only according to the accrual of paid contributions – without considering differences in expected longevity within groups of individuals – might not be considered as actuarially fair, since it might clash with accepted criteria of solidarity because it redistributes from those living less to those living more[10].
As a consequence, if people with a better socioeconomic status and higher income live longer – as shown by all empirical studies on the socio-economic gradient of health and mortality (e.g. Marmot 2015) –, they will have a relatively higher pension wealth (with respect to the accumulated amount of contributions) during their lifetimes compared to people with lower socioeconomic status and incomes and a higher risk of death at any early age. Thus, without an appropriate compensation, the NDC scheme risks engendering a regressive redistribution over the course of an individual's life (i.e. from the less well-off to the more well-off individuals), and this kind of redistribution clearly clashes with both actuarial fairness and more substantial concepts of distributive justice.
A further disadvantage might also emerge in Italy, penalising those coming from a worse socioeconomic condition, since, as pointed out, the retirement age rises when the average life expectancy increases, independent of individual expected longevity.
Therefore, from a policy perspective, without disregarding the many advantages due to the introduction of the NDC formula, one may evaluate the possibility of introducing possible corrections in the NDC formula to provide a certain guarantee to individuals who would be characterised by a long-lasting unsuccessful working career (Raitano 2017). The dynamic microsimulations of labour market outcomes and pension benefit distribution over the long run carried out in this project will thus provide further robust evidence about the need of possibly re-introducing some redistributive tool in the actuarial based Italian public pension scheme.

---

[10] The redistributive direction is further complicated by the fact that the Italian NDC also pays survivor pensions and, thus, on average, redistributes from singles to couples.

References

Borsch-Supan A. (2006), What are NDC systems? What do they bring to reform strategies?, in Holzmann R., Palmer E. (eds.), *Pension reform: issues and prospects for NDC schemes*, Washington DC, World Bank Publishing, pp.35-55

Holzmann R., Palmer E. (eds.) (2005), *Pension Reform: Issues and Prospect for Non-financial Defined Contribution (NDC) Schemes*, Washington DC, World Bank Publishing

Jessoula M., Raitano M. (2017), Italian pensions from "vices" to challenges: assessing actuarial multi-pillarization twenty years later, in Natali D. (eds.), *The New Pension Mix in Europe*, Brussels, Peter Lang Publishing, pp.39-66

Marmot M. (2015), *The health gap. The challenge of an unequal world*, London, Bloomsbury Publishing

Raitano M. (2017), *Poveri da giovani, poveri da anziani? Prospettive previdenziali e vantaggi della pensione di garanzia*, Quaderni della coesione sociale n.1, Reggio Emilia, OCIS

Raitano M. (2018), Italy. Para-subordinate workers and their social protection, in OECD (ed.), *The Future of Social Protection: What works for non-standard workers?*, Paris, OECD Publishing, pp.145-170

Raitano M. (2019), Trends and structural determinants of income inequality: an overview, in European Commission (eds.), *Addressing Inequalities. A Seminar of Workshops*, Luxembourg, Publications Office of the European Union

Roemer J. (1998), *Equality of opportunity*, Cambridge MA, Harvard University Press

Rubinstein Y., Weiss Y. (2006), Post schooling wage growth: investment, search and learning, *Handbook of the Economics of Education*, 1, pp.1-67

6. Wealth data analysis (SHIW)

Introduction

In this chapter, we provide some evidence on wealth distribution, decomposition and its determinants in Italy based on the Survey on Household Income and Wealth (SHIW). According to the Bank of Italy (hereafter BI), net wealth is divided into three main components, i.e. real wealth (mostly houses), financial activities and financial liabilities. Each of these is made of several sub-components that are described in detail in the rest of the chapter.

Although this analysis focuses on the last two decades of the Italian economic history, it is worth to underline that the stock of net wealth has been increasing in the last 50 years, with the 80s being the only decade of stagnation. This phenomenon is strictly related to the massive increase in the share of households owning a house. In the 70s, the share of homeowner households amounted to 45% – almost equal to that of tenant households – while in recent years the share has risen to 70%[1].

In what follows, particular emphasis is placed on the analysis of the evolution of financial wealth and its main determinants by focusing on portfolio choices. As well-known by practitioners, sample surveys are not the most trustworthy sources of data when dealing with individual and household wealth due to under-reporting issues[2]. In particular, as far as financial wealth is concerned, this is a recognised issue even when employing SHIW data, which has been tackled by the use of correction methods as put forward by D'Aurizio *et al.* (2006). However, it has been pointed out that this issue is more relevant when looking at the absolute level of wealth than in the case of relative distribution analysis[3].

1 See D'Alessio (2012) and the bi-annual BI reports based on SHIW for a comprehensive coverage of the history of Italian wealth (https://www.bancaditalia.it/pubblicazioni/indagine-famiglie/index.html).

2 Under-reporting refers to respondents that do not declare (intentionally or not) the ownership of a movable or immovable property, as well as when the declared value does not correspond with the true one in the case of declaration.

3 According to Cannari and D'Alessio (2018), the adjustment made to take account of under-reporting does not substantially change the trend in inequality, whereas it may affect absolute inequality levels.

The chapter is structured into two parts. In Section 6.1, we provide some evidence on the distribution of wealth and wealth ratios such as the ratio of financial wealth on net total wealth as well as the ratio of risky financial assets on total financial assets. The time span of the descriptive analysis covers the 2000-2016 period, enabling us to show the differences in trend before and after the 2008 financial crisis and the sovereign debt crisis of 2011. Furthermore, we use the aforementioned ratios as dependent variables for the study of the determinants of financial portfolio choices (Section 6.2). The estimates are based on the bi-annual waves for the 2004-2016 period. In particular, we look at financial literacy, non-standard employment and atypical contracts in relation to financial investment decisions. Finally, we cover the topic of supplementary pension schemes, which have become a crucial component of household wealth in the last years. Both analyses are conducted at the household level, while individual characteristics refer to the household head.

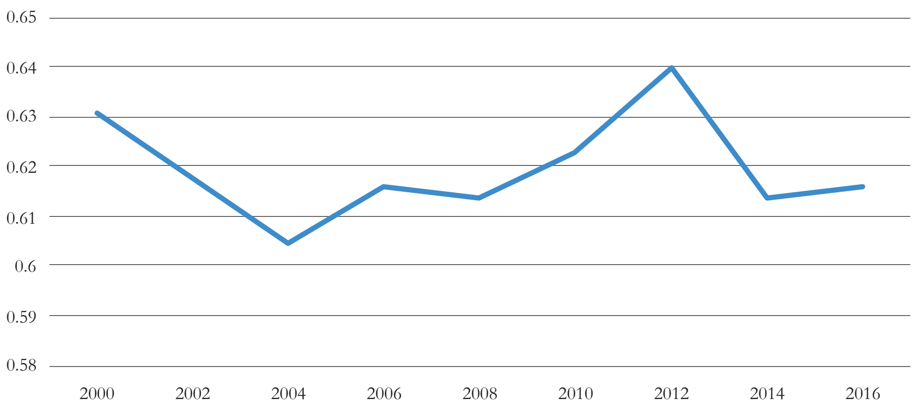
6.1 Descriptive analysis

We first focus on evidence emerging from the SHIW data.

6.1.1 Inequality

In this first part of the descriptive analysis, we provide an overview on the wealth distribution in Italy focusing on inequality. This is done in analogy with what is shown in Section 3.7 for the inequality of gross earnings. As to the inequality analysis that follows, we adopted two of the most popular inequality measures, namely the Gini index and the p75/p25 ratio.

Figure 6.1 Gini of net wealth



Source: own elaborations from SHIW 2000-16

The trend of the Gini index in Figure 6.1 is in line with the BI study by Cannari and D'Alessio (2018)[4]. The decrease in wealth inequality as measured by Gini in the beginning of the 2000s is a well-known evidence in Italy. The most common interpretation of this decline has been linked to an increase in the prices (and therefore market values) of houses (Cannari and D'Alessio 2018, Figure 2). This interpretation is also confirmed by decomposing the Gini index of net wealth into its three main components (see Table 6.1).

Following the decomposition method proposed by Lerman and Yitzhaki (1985), the share of Gini explained by the real wealth component has been increasing from 2000 onwards. Another concurrent explanation to the reduction in inequality at the beginning of the century relies on the decrease in stock prices following the 2000-2001 dot-com crisis.

Table 6.1 Shares of the Gini index of net wealth

Wealth component	2000	2002	2004	2006	2008	2010	2012	2014	2016
Real wealth	.848	.883	.923	.910	.918	.904	.888	.875	.857
Financial activities	.179	.131	.100	.106	.095	.108	.120	.130	.154
Financial liabilities	-.026	-.014	-.023	-.016	-.013	-.012	-.008	-.005	-.011

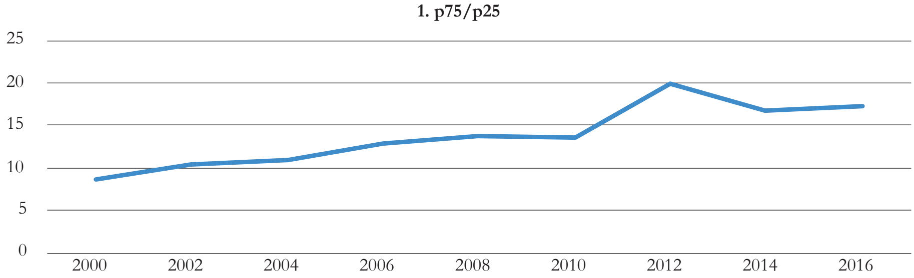
Source: own elaborations from SHIW 2000-16

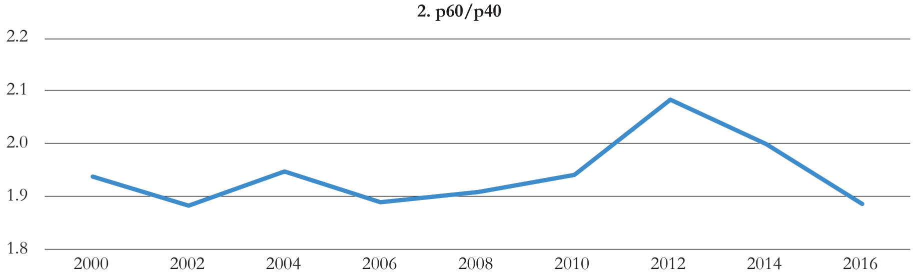
The other relevant fact emerging from the Gini trend is the spike occurring in 2012. It shows an upward trend during the 2008-2012 period reflecting the combined effect of the 2008 crisis followed by the 2011 sovereign debt crisis on the Italian economy. Unlike the Gini index for the 2000-2016 period, the p75/p25 ratio presents an overall ascending trend overtime (see Figure 6.2.1).

This can be explained by the fact that the Gini index is more sensitive to changes in the middle of a distribution rather than in the tails (Atkinson 1970). The trend is even more pronounced when looking at the p90/p10 measure (not shown here), while the p60/p40 ratio (see Figure 6.2.2), as expected, has a trend that is more similar to the Gini index, with an initial slight decrease at the beginning of the observed time period, followed by an upsurge corresponding with the financial crisis that culminates with the 2012 wave of SHIW.

4 We refer to this study for a more complete coverage of the topic.

Figure 6.2 Percentile ratios of net wealth:





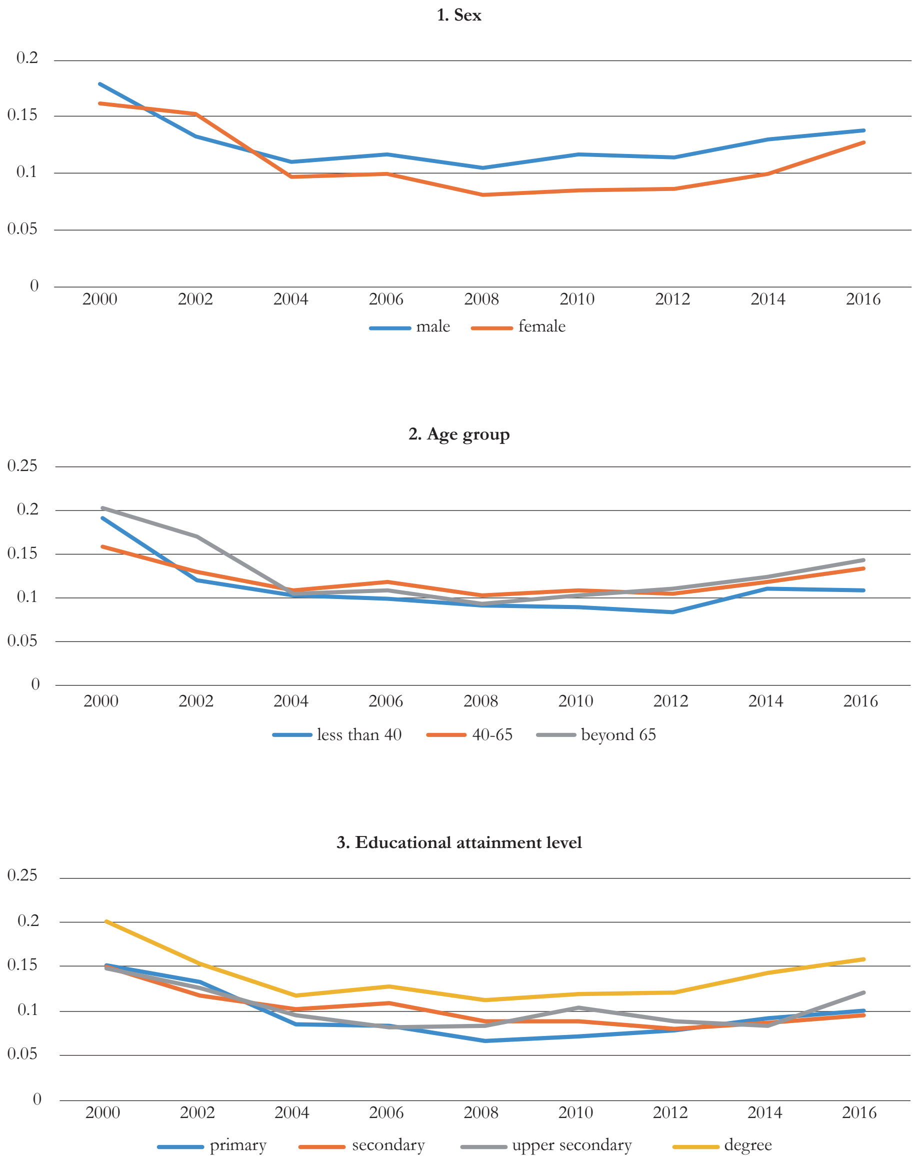
Source: own elaborations from SHIW 2000-16

6.1.2 Financial wealth ratios

In the second part of the chapter, we study the overtime evolution of financial wealth aggregates. We focus on the ratio of financial wealth on net total wealth (hereafter *f/w*) and on the ratio of risky financial assets on total financial assets (hereafter *rf/f*). We decompose these aggregates by sex, age and educational attainment level. We define three age groups: up to 40 years old, from 40 to 65 and beyond 65. For the educational attainment level, we distinguish among four groups of household heads: at least primary education, lower secondary, upper secondary and tertiary.

The first three graphs (see Figure 6.3.1-6.3.3) refer to *f/w* and present a decreasing trend over the time span considered with a slight recover in the last four years – similar to a U-shaped trend. When decomposing by household heads' characteristics, no significant discrepancy emerges except for: i) a tendency for males to hold a higher share of financial assets than female from 2004 onwards, even if the gap appears narrowed in more recent years; ii) a slight negative gap in *f/w* for those aged less than 40 in 2016, while their initial share in 2000 was almost equal to that of people aged more than 65 (the category with the highest *f/w* for both years); and iii) a positive gap throughout the whole period for tertiary graduates with respect to the other three categories.

Figure 6.3 f/w by household heads' characteristics:

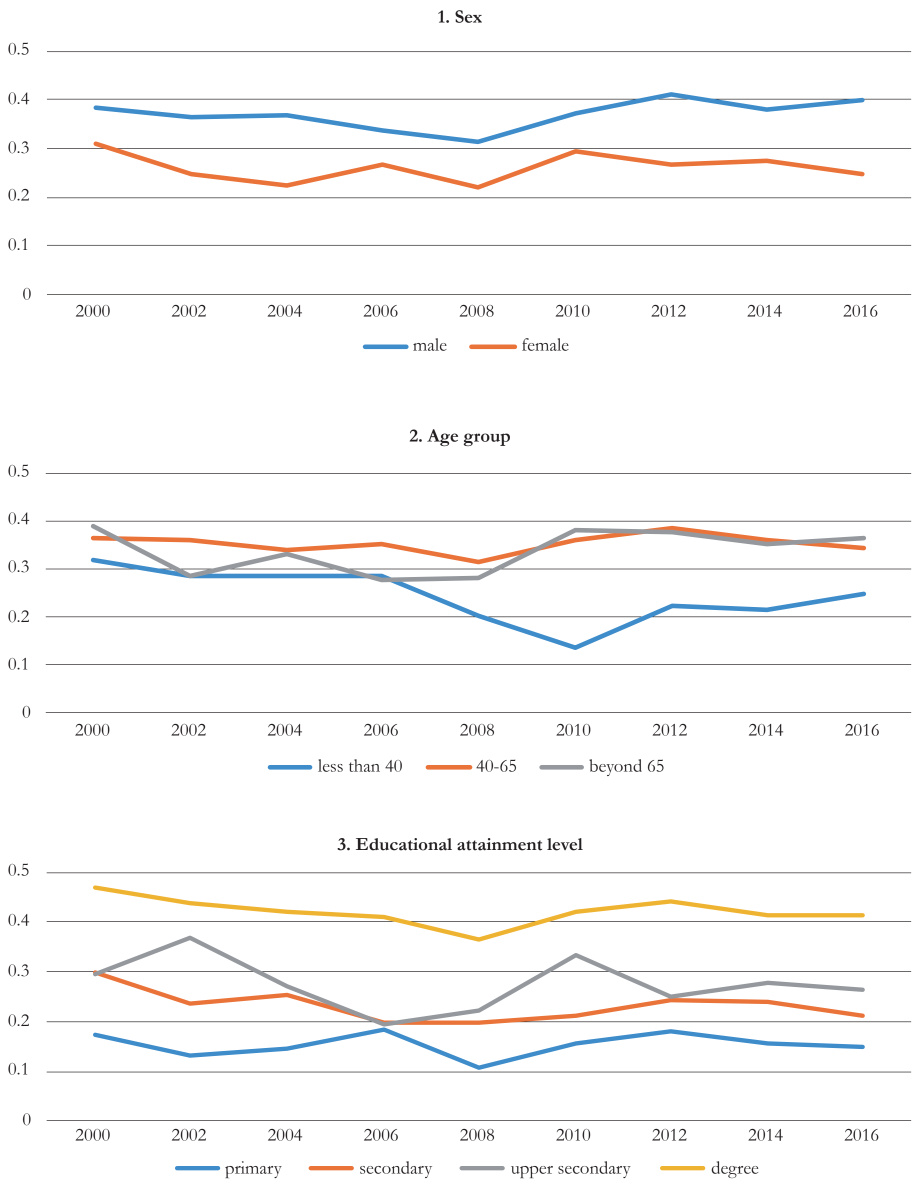


Source: own elaborations from SHIW 2000-16

According to the BI definition, risky assets concern private sector bonds, equity shares and other participating interests, managed investment schemes, securities issued abroad and other financial assets (e.g. deposit, certificates of deposit, repurchase agreements).

As far as *rf/f* is concerned (see Figure 6.4.1-6.4.3), unlike *f/w*, the overall trend was quite stable throughout the 2000-2016 period.

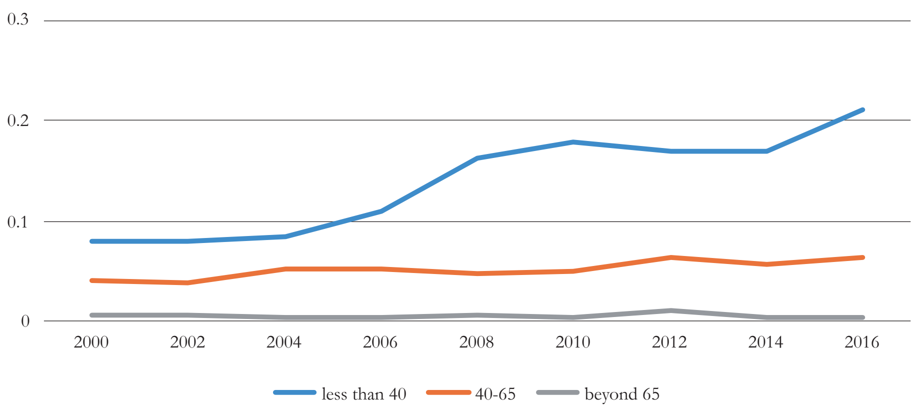
Figure 6.4 rf/f by household heads' characteristics:



Source: own elaborations from SHIW 2000-16

In this case, decomposing by sex, age and education attainment level provides interesting discrepancies among groups. Female household heads tend to hold a lower share of risky assets with respect to male. A large body of literature on behavioural and experimental economics has been investigating gender differences in risk and ambiguity aversion[5] (see, among others, Borghans *et al.* 2009; Schubert *et al.* 1999), as well as in financial risk taking (Powell and Ansic 1997). Although these studies differ in the approach to and interpretation of the results made, women are found to be more risk and ambiguity averse than men and are characterised by a lower financial risk-taking. The only age group that has significantly reduced investments in risky assets just before and after the 2008 financial crisis is that of the younger household heads. Parallel to this evidence, as it arises in Figure 6.5, this group increased its financial liabilities on net wealth during the same time span (mostly mortgages). A possible explanation is that the weight of mortgages, exacerbated by liquidity shortages during the crisis, deterred households in this age group from investing in risky assets.
Finally, the gaps in the risky investments are highly significant when looking at different educational attainment levels. Household heads with primary education or less tend to invest in risky assets on average about 15% of their financial portfolio, whereas individuals with tertiary education invest on average around 45%.

Figure 6.5 Financial liabilities on net wealth by age group of the household head



Source: own elaborations from SHIW 2000-16

[5] The term *ambiguity (uncertainty) aversion* is defined as "*the decision maker's preference to bet for and against a chance event whose probability is known, rather than on an equivalent with vague probabilities*" (Di Mauro 2008).

6.2 Determinants of financial choices

In the second part of the analysis, we try to disentangle some of the determinants of financial portfolio choices. We do it by looking at both the aggregates described in the previous section, i.e. *f/w* and *rf/f*. In the first paragraph, we devote our attention to financial literacy. In the second one, we concentrate on working conditions and in particular on non-standard employment arrangements and atypical contracts. Finally, in the last paragraph, we look at supplementary pension scheme choices.

6.2.1 Financial literacy

Financial literacy has been a novel topic in the literature on financial choices. Lusardi and Mitchell (2014) have put one definition forward: *"by financial literacy we mean peoples' ability to process economic information and make informed decisions about financial planning, wealth accumulation, pensions, and debt"*. In this paragraph, we want to verify whether the level of financial literacy has an influence on the propensity of Italian households to invest in risky assets. Arrondel *et al.* (2015) in France and Liao *et al.* (2017) in China found recent evidence on the topic.

In the SHIW questionnaire for the 2016 wave, three questions on financial literacy were included regarding nominal and real interest rates and risk diversification. Thanks to these questions, we constructed an ordered categorial variable that captures the level of financial knowledge varying from 0 to 3 (i.e. 0: illiterate; 1: lowly literate; 2: lowly-to-moderately literate; 3: literate) depending on the number of correct answers. In Table 6.2, the distribution of financial literacy by sex, age group and education attainment level is shown. Overall, the sample is well-distributed across the four categories into which we stratified financial literacy.

First, it should be noted that Italian households tend to answer correctly to the financial-related questions. Male household heads present a higher level of financial literacy, while the middle aged (between 40 and 65 years old) are the most financial literate age group. Furthermore, not surprisingly, the level of financial literacy is highly correlated with the educational attainment level, with almost 50% of those with tertiary education responding correctly to all the three questions. These descriptive outcomes are in line with the results of Baglioni *et al.* (2018) in a work based on a survey questionnaire in which the authors find that being a woman or being less educated is associated to less familiarity with the financial environment. The issue of a gender gap in financial literacy has been addressed in detail also by Hasler and Lusardi (2017).

In the rest of the paragraph, we deepen the role of financial literacy jointly with other relevant variables listed and discussed below as a driver for household investment choices. In particular, we focus on risky investments using as dependent variable the ratio of risky financial assets on total financial assets – what we defined for simplicity' sake as *rf/f*. Unlike the previous analyses, we use SHIW data for the year 2016 since it is the first and only wave (at the time of writing) in which questions regarding the level of financial literacy were asked. The analysis is conducted through a Tobit model that has been chosen in order to properly deal with the great amount of values equal to zero in the dependent variable (only a small percentage of the households present positive values of risky assets in their portfolio).

Table 6.2 Financial literacy by household heads' characteristics

Financial literacy	Male		Female		Total
(0) illiterate	17.62		31.02		24.13
(1) lowly literate	18.45		20.11		19.26
(2) lowly-to-moderately literate	31.90		28.43		30.21
(3) literate	32.03		20.44		26.40
	Less than 40	40-65	Beyond 65		Total
(0) illiterate	20.67	16.00	33.18		24.13
(1) lowly literate	20.50	18.87	19.43		19.26
(2) lowly-to-moderately literate	29.50	33.82	26.62		30.21
(3) literate	29.33	31.32	20.76		26.40
	Primary	Lower sec.	Upper sec.	Tertiary	Total
(0) illiterate	47.18	21.99	11.50	6.44	24.13
(1) lowly literate	20.68	22.11	16.84	12.52	19.26
(2) lowly-to-moderately literate	22.10	32.68	33.67	33.29	30.21
(3) literate	10.04	23.22	37.99	47.75	26.40
Total	100.00	100.00	100.00	100.00	100.00

Source: own elaborations from SHIW 2000-16

The explanatory variables included in the model are the following: level of financial literacy, sex, age, age squared, marital status, employment status, self-employed (binary variable), tertiary education (binary variable), macro areas (binary variables), wealth quintiles and number of wealth earners.
Results are shown in Table 6.3. As expected, an increase in the level of financial literacy corresponds to an increase in *rf/f*. The employment status has a positive impact on risky financial choices, and the same happens when the household head holds a university degree. The Northern regions of Italy have a higher propensity to invest in risky assets. Last, the level of the ratio grows with increasing net wealth showing that richer households tend to invest more in risky financial assets than poorer ones.

Table 6.3 Tobit model for rf/f

rf/f	Coefficient	S.e.
Financial literacy	0.2768***	(0.0213)
Female	-0.0778**	(0.0378)
Age	-0.0057	(0.0088)
Age squared	0.0000	(0.0001)
Married	0.0517	(0.0434)
Employed	0.1583***	(0.0549)
Tertiary education	0.1869***	(0.0435)
Centre	-0.3795***	(0.0435)
South	-1.0921***	(0.0693)
Wealth quintile_2	0.9576***	(0.1334)
Wealth quintile_3	1.1429***	(0.1321)
Wealth quintile_4	1.3981***	(0.1318)
Wealth quintile_5	1.6660***	(0.1333)
No. Earners	-0.0292	(0.0285)
Self-employed	0.0060	(0.0541)
Constant	-2.0644***	(0.2912)
Observations	5,834	
Left-censored obs.	4,909	

Source: own elaborations from SHIW 2000-16

6.2.2 Non-standard employment and atypical contracts

In the second paragraph, we analyse the role of some specific working conditions on household investment choices. Following the results obtained in Section 3.8, in which the type of contractual arrangement emerged as a key determinant of earning inequality, we dwell on the possible impact of this dimension on the ownership of financial assets. In particular, we focus on non-standard workers (para-subordinate and temporary agency workers) and employees with fixed-term contracts (about 8% of the sample).

We model separately the relationship between the head's status of non-standard worker and two endogenous ratio variables:

1. Head's status of non-standard employment and the ratio of financial wealth on net total wealth (*f/w*) investigated through a Heckman two-step selection model (Heckman 1979) in order to take into account the role of selection separately from the outcome. The instrument for the exclusion restriction is the level of head's parental educational attainment.

2. Head's status of non-standard employment and ratio of risky financial assets on total financial assets (*rf/f*) analysed by means of a Tobit model.

The dependent variable for the Heckman two-step selection model is a modified version of f/w discussed above. This modified ratio aims at capturing a static data generation process for what concerns f/w, with some corrections that are listed and motivated in what follows:

a. The numerator of this ratio is equal to $max\ (0, f - acs)$ with $acs = max\ [0, (yd)/12]$. The term *acs* represents a household specific proxy of monthly disposable income net of non-monetary wage components (e.g. fringe benefits) and imputed rents from real estate properties. By doing so, we aim at capturing current and past propensity to save in terms of financial wealth in excess of a measure of average current account balance.
b. The denominator is $max\ (0, w)$, which is total net wealth if the value of assets is greater than liabilities, otherwise it is zero. This prevents the ratio from being negative.
c. The selection variable to estimate the control function (inverse Mills ratio) is equal to one if the modified ratio is greater than zero, while it is zero if the modified ratio is nil. The idea is to consider those households who declared a positive value of financial wealth below a certain threshold of bank (or postal) account savings – which depends on the average monthly flow of disposable income – as if they had no financial wealth. As a result, those households who have a positive value of the modified ratio hold a stock of financial wealth – also in the form of liquid assets – in excess to the household-specific threshold.

As to control covariates, we include: gender, age, age squared, marital status, employment status, tertiary education (binary variable), macro areas (binary variables), real wealth quintiles, number of wealth earners, sector of employment (i.e. agriculture, industry, trade, transport, banking and insurance, public administration).

We use pooled SHIW data for the 2004-2016 period and control for time effects.

Results are shown in Table 6.4. Starting from the bottom, it is worth noting that the estimated coefficient for inverse Mills ratio – in a two-step version estimator – is significant at 1% level, while the Wald test, which relates to the ML estimates – rejects the null hypothesis of independence between equations. These tests support the choice of modelling the selection process in a separate fashion.

As to the type of contractual arrangement, the table shows that fixed-term job contracts and non-standard employment status are both significant predictors of a lower probability of holding financial assets (both liquid or not), above all for the latter predictor. These variables are more relevant to model the selection (i.e. *Pr* (modified ratio)>0, see the second column) rather than the outcome (i.e. the value of the modified ratio, see column 1). Not surprisingly, the status of retirement explains positively both the probability and the outcome due to the high share of household financial wealth held by pensioners. Furthermore, tertiary education, the number of wealth earners within the household and self-employed household heads are also positive predictors in both equations.

Table 6.4 Ratio of financial wealth on net total wealth (f/w) with Heckman correction

	Modified_ ratio (1)	S.e.	Selection (2)	S.e.
Father's education			0.0270*	(0.0140)
Non-standard employment	-0.0176	(0.0116)	-0.3813***	(0.0727)
Fixed-term	-0.0128**	(0.0063)	-0.2453***	(0.0369)
Female	-0.0131***	(0.0035)	-0.0165	(0.0249)
Age	0.0048***	(0.0012)	-0.0105	(0.0078)
Age squared	-0.0000***	(0.0000)	0.0001**	(0.0001)
Employed	-0.0035	(0.0110)	0.3660***	(0.0771)
Work experience	-0.0006*	(0.0003)	0.0004	(0.0023)
Retired	0.0227***	(0.0074)	0.4220***	(0.0475)
Married	-0.0032	(0.0039)	-0.0720***	(0.0262)
Tertiary education	0.0524***	(0.0038)	0.3350***	(0.0272)
No. earners	0.0057***	(0.0020)	0.0270*	(0.0157)
Self-employed	0.0207***	(0.0045)	0.0827**	(0.0357)
Centre	-0.0180***	(0.0040)	-0.1760***	(0.0294)
South	-0.0596***	(0.0037)	-0.4051***	(0.0251)
Real Wealth quintile_2	-0.3357***	(0.0083)	0.5636***	(0.0327)
Real Wealth quintile_3	-0.4870***	(0.0068)	0.6385***	(0.0336)
Real Wealth quintile_4	-0.5067***	(0.0069)	0.8143***	(0.0348)
Real Wealth quintile_5	-0.5284***	(0.0071)	0.8741***	(0.0390)
year = 2006	0.0086	(0.0055)	-0.0931**	(0.0405)
year = 2008	0.0010	(0.0054)	-0.2012***	(0.0396)
year = 2010	0.0072	(0.0057)	-0.1946***	(0.0402)
year = 2012	0.0005	(0.0054)	-0.3806***	(0.0395)
year = 2014	-0.0048	(0.0054)	-0.2877***	(0.0381)
year = 2016	-0.0090	(0.0060)	-0.2182***	(0.0410)
Sector dummy_industry	0.0157	(0.0099)	-0.0077	(0.0730)
Sector dummy_trade	0.0011	(0.0102)	-0.0718	(0.0762)
Sector dummy_transport	-0.0016	(0.0123)	-0.0445	(0.0904)
Sector dummy_banking	0.0510***	(0.0125)	-0.0403	(0.1050)
Sector dummy_ public admin.	0.0101	(0.0095)	-0.0605	(0.0720)
Mills	-.255***	(0.03602)		
Constant	0.4517***	(0.0316)	0.0670	(0.2023)
No. observations	36,183		36,183	
Wald test	(rho = 0): chi2(1) = 6.27	Prob > chi2 = 0.0123		

Source: own elaborations from SHIW 2000-16

Table 6.5 Tobit model for risky-to-total financial assets ratio (rf/f), SHIW, 2004-16

rf/f	Coefficient	S.e.
Non-standard work	-0.1047*	(0.0541)
Fixed-term	-0.0754***	(0.0234)
Female	-0.1025***	(0.0124)
Age	0.0165***	(0.0029)
Age squared	-0.0002***	(0.0000)
Work experience	0.0012	(0.0013)
Employed	-0.1088**	(0.0495)
Retired	0.0241	(0.0296)
Married	-0.0053	(0.0132)
University Degree	0.3116***	(0.0136)
Centre	-0.2700***	(0.0131)
South	-0.7469***	(0.0181)
Wealth quintile_2	0.6437***	(0.0290)
Wealth quintile_3	0.6496***	(0.0290)
Wealth quintile_4	0.8846***	(0.0285)
Wealth quintile_5	1.1480***	(0.0292)
year = 2006	-0.0827***	(0.0187)
year = 2008	-0.1413***	(0.0190)
year = 2010	-0.0762***	(0.0189)
year = 2012	-0.0822***	(0.0191)
year = 2014	-0.1083***	(0.0192)
year = 2016	-0.1698***	(0.0204)
No. of earners	0.0094	(0.0080)
Sector dummy_industry	0.1305***	(0.0436)
Sector dummy_trade	0.1055**	(0.0449)
Sector dummy_transport	0.0130	(0.0533)
Sector dummy_banking	0.3898***	(0.0492)
Sector dummy_ public admin.	0.0914**	(0.0432)
Self-employed	-0.0538***	(0.0166)
Constant	-1.3669***	(0.0865)
Observations	41,498	

Source: own elaborations from SHIW 2000-16

A geographical pattern emerges with respect to the probability and the outcome, both decreasing as one moves from northern to southern parts of the country. Notably, a strong non-linearity emerges when considering the role of being in a certain real wealth quintile. In particular, holding real wealth is a significant positive predictor of holding financial assets, but at the same time, it affects negatively the ratio of financial wealth on net total wealth (*f/w*). A U-shape effect emerges for the selection over time (i.e. year binary variables) while no trend is found for the outcome. Finally, the employment sectors considered seem not to play any relevant role except for a positive effect of working in the banking sector.

The analysis regarding the ratio of risky financial assets on total financial assets (*rf/f*) is conducted through a Tobit model for the same reasons explained above in the previous paragraph. The explanatory variables included in the model are: non-standard work (binary variable), fixed-term contract (binary variable), sex, age, age squared, marital status, year (binary variables), employment status, work experience, self-employed (binary variable), tertiary education (binary variable), macro areas (binary variables), wealth quintiles, sector of employment (i.e. agriculture, industry, trade, transport, banking and insurance, public administration) and number of wealth earners within the household.

The results are shown in Table 6.5. Being employed in a non-standard job has a negative and significant impact on *rf/f*, and the same holds for fixed-term workers. Female household heads tend to be less hazardous in terms of financial investments. Those with tertiary education show a higher tendency to invest in risky assets. Living in the northern part of Italy increases the probability of investing in risky assets. The *rf/f* ratio increases with respect to the net wealth owned by a household. The year binary variables denote a negative trend of *rf/f* across the time span observed. As expected, working in the banking sector has a significant positive impact on the probability of investing in risky assets.

6.2.3 Supplementary pensions

In this final part of the section, we concentrate on supplementary (private) pension schemes, since their increased incidence in wealth portfolios. We study the determinants of the choice of having a supplementary pension through a logit model. As in the previous case, we use pooled SHIW data for the 2004-2016 period.

The explanatory variables included in the model are: sex, age, age squared, marital status, year (binary variables), employment status, work experience, self-employed (binary variable), tertiary education (binary variable), macro areas (binary variables), wealth quintiles, sector of employment (agriculture, industry, trade, transport, banking and insurance, public administration), non-standard work (binary variable), fixed-term contract (binary variable) and number of wealth earners within the household.

Table 6.6 Logit model for participation in supplementary pension schemes

Supplementary pension	Coefficient	S.e.
Female	-0.0700	(0.0511)
Age	0.1717***	(0.0176)
Age squared	-0.0020***	(0.0002)
Work experience	0.0027	(0.0052)
Employed	0.2956*	(0.1759)
Retired	-0.0530	(0.1299)
Married	0.0487	(0.0528)
University Degree	0.4112***	(0.0533)
Centre	-0.2888***	(0.0553)
South	-0.6911***	(0.0578)
Wealth quintile_2	0.6888***	(0.0834)
Wealth quintile_3	0.9123***	(0.0837)
Wealth quintile_4	0.9974***	(0.0821)
Wealth quintile_5	1.3583***	(0.0856)
year = 2006	0.0032	(0.0858)
year = 2008	0.0683	(0.0860)
year = 2010	0.8558***	(0.0799)
year = 2012	0.7308***	(0.0800)
year = 2014	0.7623***	(0.0807)
year = 2016	0.6468***	(0.0884)
No. of earners	0.2870***	(0.0295)
Sector dummy_industry	0.8266***	(0.1516)
Sector dummy_trade	0.5705***	(0.1588)
Sector dummy_transport	1.0291***	(0.1751)
Sector dummy_banking	1.5272***	(0.1739)
Sector dummy_ public admin.	0.4350***	(0.1521)
Self-employed	-0.5824***	(0.0633)
Non-standard	-0.3031	(0.1843)
Fixed-term	-0.0999	(0.0796)
Constant	7.5298***	(0.4316)
Observations	48,748	

Source: own elaborations from SHIW 2000-16

Results are shown in Table 6.6. Participation in supplementary pension schemes increases with age and is strongly influenced by holding a university degree. As to the other predictors, private pensions are more commonly spread in the northern regions of Italy. Not surprisingly, the choice to contribute to a supplementary pension scheme is positively correlated to the level of household net wealth. As far as year binary variables are concerned, it emerges that there has been an increasing trend in participation with respect to the reference year (2004). Finally, it is interesting to stress out that being a self-employed worker is negatively correlated with the probability of being enrolled in a supplementary pension scheme. This evidence could be explained in a context of overall risk management, where the self-employed – who bear a higher risk in their private working activities than other workers – prefer not to add extra sources of risk in their life-cycle investment choices.

Conclusions

In this chapter, we provided some evidence on the distribution of wealth among households in Italy based on the Survey on Household Income and Wealth (SHIW) released by the Bank of Italy. A decrease in the wealth inequality at the beginning of the 2000s is a well-known piece of evidence in Italy, and a common interpretation of this downturn links it to an increase in the prices of the houses whose ownership is widespread among Italian households. Another concurrent explanation to the reduction in inequality at the beginning of the century relies on the decrease in the stock prices following the dot-com crisis. The other relevant fact emerging from the descriptive analysis is a spike in the Gini index of net wealth occurring in 2012. It follows an upward trend that started in 2008 and it reflects the combined effects on the Italian economy of the 2008 worldwide financial crisis followed by the European sovereign debt crisis in 2011. Unlike the non-monotonic trend shown by the Gini index, inter-quantile ratios show an overall increase in wealth inequality in the 2004-2016 period. We then tried to disentangle some of the determinants of financial portfolio choices by looking at the ratio of financial wealth on net total wealth (*f/w*) and at the ratio of risky financial assets on total financial assets (*rf/f*). We find a significant positive role of financial literacy on the holding of risky financial assets, while non-standard employment arrangements or atypical positions in the labour market are associated with lower investment in risky assets. Finally, the analysis of participation in supplementary pension schemes confirms and quantifies a critical issue for social security wealth and pensions' long-term adequacy in Italy. In particular, those individuals and households who would most benefit from a supplementary pension scheme – i.e. those with low accumulated financial and real assets and/or with fragmented careers in the labour market – are also those who show the lowest participation in the private pension pillar.

References

Atkinson A.B. (1970), On the Measurement of Inequality, *Journal of Economic Theory*, 2, n.3, pp.244-263

Arrondel L., Majdi D., Savignac F. (2015), Stockholding in France: the role of financial literacy and information, *Applied Economics Letters*, 22, n.16, pp.1315-1319

Baglioni A., Colombo L., Piccirilli G. (2018), On the anatomy of financial literacy in Italy. Economic Notes: Review of Banking, *Finance and Monetary Economics*, 47, n.2-3, pp.245-304

Borghans L., Heckman J.J., Golsteyn B.H., Meijers H. (2009), Gender differences in risk aversion and ambiguity aversion, *Journal of the European Economic Association*, 7, n.2-3, pp.649-658

Cannari L., D'Alessio G. (2018), *La disuguaglianza della ricchezza in Italia: ricostruzione dei dati 1968-75 e confronto con quelli recenti*, Questioni di economia e finanza n.428, Roma, Banca d'Italia

D'Alessio G. (2012), Ricchezza e disuguaglianza in Italia, in D. Checchi (ed.), *Disuguaglianze diverse*, Bologna, Il Mulino

D'Aurizio L., Faiella I., Iezzi S., Neri A. (2006), The under-reporting of financial wealth in the Survey on Household income and Wealth, Temi di discussione n.610, Roma, Banca d'Italia

Di Mauro C. (2008), Uncertainty Aversion Vs. Competence. An Experimental Market Study, *Theory and Decision*, 64, n.2, pp.301-331

Hasler A., Lusardi A. (2017), *The gender gap in financial literacy. A global perspective*, Global Financial Literacy Excellence Center, The George Washington University School of Business <https://bit.ly/2NS3Z12>

Heckman J.J. (1979), Sample selection bias as a specification error, *Econometrica*, 47, n.1, pp.153-162

Lerman R.I., Yitzhaki S. (1985), Income inequality effects by income source. A new approach and applications to the United States, *Review of Economics and Statistics*, 67, n.1, pp.151-156

Liao L., Xiao J.J., Zhang W., Zhou C. (2017), Financial literacy and risky asset holdings: evidence from China, *Accounting & Finance*, 57, n.5, pp.1383-1415

Lusardi A., Mitchell O.S. (2014), The economic importance of financial literacy. Theory and evidence, *Journal of economic literature*, 52, n.1, pp.5-44

Powell M., Ansic D. (1997), Gender Differences in Risk Behaviour in Financial Decision-Making. An Experimental Analysis, *Journal of Economic Psychology*, 18, n.6, pp.605-628

Schubert R., Gysler M., Brown M., Brachinger H. (1999), Financial Decision-Making. Are Women Really More Risk-Averse?, *American Economic Review*, 89, pp.381-385

7. Microsimulation modelling: T-DYMM 3.0

Introduction

The world of work is ever changing. Discontinuity in working careers and the challenges of adequate social protection for non-standard workers are amongst the most relevant topics in this fast-evolving environment of today. To address these and other issues, the new version of T-DYMM, the Treasury Dynamic Microsimulation Model of the Italian Ministry of Economy and Finance, is a unique opportunity to investigate the medium and long-term effects of current policies and alternative scenarios. In a perspective of ever-increasing market income inequalities, it is of crucial importance to foster the discussion on how to rebalance post-market outcomes. Within the MOSPI project, we believe that T-DYMM's improved simulation capabilities will be of great help for such purposes. In this Chapter, a brief taxonomy of microsimulation models will be presented in Section 7.1. By doing so, our intention is to provide an overview on the methodological fundamentals and main uses of microsimulation modelling in the field of economics. A distinction between models focused on short-term effects rather than medium- and long-term effects of a given policy change will be taken as reference in the discussion, and a preliminary look at ways in which behavioural adjustments to policy changes can be embedded within dynamic models will be offered. Our aim is to outline T-DYMM's features in a comparative context, giving account of its role in the modelling community. Section 7.2 turns the attention to the developments of the latest version of T-DYMM, while offering a perspective on its previous releases (Caretta *et al.* 2013; IESS 2016). Finally, Section 7.3 outlines the research topics that the new version of the model will attempt to tackle.

7.1 The long road of microsimulation modelling: a taxonomy of models

Within the broad spectrum of techniques and approaches that we usually refer to as *tools for the evaluation of public policies*, the use of simulation models finds a concrete and well-defined space. Unlike in natural sciences, where randomised experiments are the cornerstone of a better understanding of the surrounding environment, they are hardly ever practicable for the study of socio-economic matters. The difficulties encountered in defining the ideal conditions for the conduction of the experiments made researchers decide towards simulating the environment under study. For example, as in our case, starting from a representative sample of the population and simulating individual characteristics and the tax-benefit system over the long run, one can study the effect of a change in pension entitlement rules on the adequacy of the pension system. The presence of simulated individuals within the model allows for an in-depth analysis of the effects of a given policy on specific subgroups of the population as well as of the overall distributive impact.

The first simulation models in the economic field were born in the 1930s (Baldini and Toso 2009). Generally defined as *macroeconometric models*, they focus their attention on the functioning of the economic system as a whole (simplified by a system of equations) and on the aggregate variables that better define it. From this purely macroeconomic view, the focus later moved on to the inclusion of a microeconomic foundation in *computational models of general equilibrium* (CGE), joining general equilibrium theory with evidence provided by aggregate administrative data. The need of taking account of the full heterogeneity of the population led to the birth of what can now be defined as *microsimulation models* around the end of the 1950s, even though their extensive use was made possible starting from the beginning of the 1970s thanks to the developments in software computational capabilities. The shift of attention from macroeconomic to microeconomic units can be traced back to the work of Orcutt (1957), who laid down the foundation of the field in 'A new type of socio-economic system', coming back ten years later (Orcutt 1967) to highlight the importance of merging microsimulation techniques to macroeconomic models. The wealth of information now available on individuals and households has finally allowed the detailed simulation and projection of the multitude of dimensions and interactions that economic agents experience.

In the broad field of economics, microsimulation models are tools used in the simulation activity based on sample or administrative data of micro units of analysis such as individuals, households or firms. The modelling activity consists of the use of computational software in order "to capture the impact of changes in (*e.g. tax-benefit*) policy parameters and/or changes in the behaviour of economic agents" (Dekkers 2015b). Unlike macroeconomic models, individual heterogeneity is completely preserved and modelled.

Over the years, a vast literature has described uses, methodological choices and design of microsimulation models (among the many: Merz 1991; Bourguignon and Spadaro 2006; Dekkers 2015b). For our purposes, we will make use of a distinction that sees static models and labour-supply models on one hand, where the effect of a specific policy change is evaluated in the short term, opposed to dynamic models, which are tools for the study of medium- and long-term effects. What is here argued does not have the ambition to cover the entire range of applications in microsimulation analyses. Many other uses within the economic field such as spatial models (O'Donoghue *et al.* 2014; Tanton 2018), health models (Rutter *et al.* 2010), macro-micro models (Peichl 2016) and firm level models (Buslei *et al.* 2014) are not treated in the following.

7.1.1 Focus on short-term effects: static and labour supply models

The peculiar trait of these models is the focus of analysis on the short-term temporal horizon. Static models simulate the complex system of taxes and benefits of a specific country – or for a multitude of countries as in the case of EUROMOD (Sutherland and Figari 2013). For instance, by imposing an increase in the tax rates of the personal income tax, what the model does let us calculate are the *first-order effects* (otherwise called *direct effects*) of the policy change simulated, that is the redistribution of burdens and benefits and total budgetary costs of such a reform without taking account of behavioural responses.

To put it in even simpler terms, we do not know how many hours one would be willing to offer in the labour market following a tax shock. As a result, this would impact on the level of work-earned income supposing that the increased or decreased number of hours are fully absorbed by the demand side. This does not mean that static models are pointless in the study of labour supply responses: direction of change can be grasped by looking at the average tax rate per income groups or at the effective marginal tax rate of specific taxpayers. Their use is often the starting point of a more elaborated analysis that first requires the simulation of the underlying tax-benefit system.

The great majority of the models that simulate tax-benefit rules do it on cross-sectional survey data representative of the population of reference (Martini and Trivellato 1997; Ceriani *et al.* 2013). Because of the likely mismatch between gross income aggregates and corresponding administrative totals, the implementation of a *net-to-gross algorithm* is required in order to adjust gross values on which the simulation is based (Immervoll and O'Donoghue 2001). The cases where the model is built up just on administrative data are more limited, even though the combination of the two data sources has become a common practice. Furthermore, social-demographic variables and income values on which the model is built up refer to a tax year (e.g. *t-3*) that rarely correspond with the tax-benefit system simulated (*t*). This implies the update of income values to national account aggregates (e.g. using the Consumer Price Index) as well as the *static ageing* of the population, where individuals maintain their original characteristics as initially collected in the survey but a new sample weight is assigned to capture socio-demographic changes in the transition *t-3* → *t*. Taxpayers' weight by income groups of the reference (gross) distribution is often reweighted before simulating the tax-benefit system as of *t*. The last step consists in the macroeconomic validation of the model, where total simulated income earners/taxpayers and income aggregates (such as total gross income subject to taxation, total taxes due and total benefits received) are compared with official values to assess model accuracy (Sutherland 2018). The modelling framework highlighted above can be further extended by including labour supply responses. Individuals react to policy changes by adjusting the number of hours worked or changing consumption patterns. Capturing these effects means studying the *second-order effects* (or *indirect effects*) of a policy change in the tax-benefit structure.

A simplified framework borrowed from Aaberge and Colombino (2014) might be of help in defining such models. Suppose that a representative sample of individuals or households ($i=1, \dots, N$) at time t faces a budget set β_i whose elements (x_i) include labour supply choices such as hours of work and sector of occupation as well as consumption behaviours. To simplify matters, $\beta\iota$ can be expressed in the most elementary way as follows:

$$B_i=\{(c,h) : c \leq f(wh,I)\}$$

where h are hours of work; c is the net income available; w indicates the hourly gross wage rate; I stands for unearned income and $f(\,)$ is the tax-benefit system. Assuming that preferences do not change with a shock imposed in the tax-benefit system, the *i-th* agent will act so as to maximise a constrained utility function $U_i\,(x)$ that gathers individual preferences. By letting h varying within a predefined weekly range (e.g. $h = 35, \dots, 45$ with h integer number) or continuously over time, our agent will choose the new x_i following the policy change on the basis of $x_i=max\ U_i\,(x)$ with $\in\beta_i$.

For a more extended explanation of how these models work, see Aaberge and Colombino (2018), where a distinction between *discrete choice models* and *continuous labour supply models* is provided.

The discrete approach seems to have gained more popularity in the microsimulation community both due to its more realistic definition of what labour supply choices are, where individuals are more likely to face limited working options (e.g. part-time or full-time job), and because of the computational advantages in the estimation of individuals' preference functions (Creedy and Kalb 2005).

A list of the static and labour supply models still in operation in the Italian context and main features is given in Table 7.1.

Given the above description, T-DYMM shares with static models the simulation of the tax-benefit system. The new version of the model will benefit from a further development of the *Taxation module* made possible by the use of individual tax returns, as well as from a more comprehensive simulation of social protection measures.

Table 7.1 Static and labour supply models still in operation in Italy

Name	Reference	Type	Data	Main feature
BETAMOD	Albarea *et al.* (2015)	S	IT-SILC	Reliable estimates of individual tax evasion rates
BIMic	Curci *et al.* (2017)	S	SHIW	Bank of Italy's model; reliable estimates on movable and immovable properties
Colombino's model	Colombino (2015)	LS	SHIW	Extensively used for the study of a universal policy of income support in Italy
EUROMOD (IT)	Sutherland and Figari (2013)	S	IT-SILC	Cross-country comparability and full simulation of benefits
FaMiMod	Cozzolino and Di Marco (2015)	S	IT-SILC	ISTAT's model; broadness of the policies simulated
ITaxSIM	Baldini *et al.* (2015b)[1]	S	IT-SILC	Department of the Treasury's model; broadness of the policies simulated
Pacifico's model	Pacifico (2009)	LS	SHIW	Allows for errors in predicted wages, unobserved heterogeneity in preferences and unobserved monetary fixed costs
MAPP©	Baldini *et al.* (2015a); Boscolo (2019)	S	IT-SILC	In-cash and in-kind transfers; detailed simulations of income components exempt from progressive taxation
MEF (FI-NANCE)	Di Nicola *et al.* (2015)	S	IT-SILC/ ITR	Department of Finance's model; built on an exact match between survey data and personal income tax returns
MicroReg	Maitino *et al.* (2017)	S	IT-SILC	Indirect taxes and in-kind benefits
Pellegrino's model	Pellegrino *et al.* (2011)	S	SHIW	Housing taxation
SM2	Betti *et al.* (2011)	S	IT-SILC	Net-to-gross algorithm used in EU-SILC for Italy
TABEITA	D'Amuri and Fiorio (2006)	S	IT-SILC	Emphasis on net-to-gross procedure and validation issues
TREMOD	Azzolini *et al.* (2017)	S	ICFT/ ITR	Regional model (Province of Trento) based on the EUROMOD platform
UPB	Gastaldi *et al.* (2017)	S	SHIW	Parliamentary Budget Office's model; broadness of the policies simulated

Note: S is 'static model', while LS is 'labour supply model'. IT-SILC stands for 'European Union Statistics on Income and Living Conditions' for Italy; SHIW is the 'Survey on Household Income and Wealth' of the Bank of Italy; ICFT is the 'Survey on Living Condition in the Province of Trento' (*Indagine sulle condizioni di vita delle famiglie trentine*); while ITR means individual tax returns. For a more complete review of static models in Italy, see Azzolini *et al.* (2017). For an international discussion, see Li *et al.* (2014c).

[1] The paper indicated seems to be the only one where ITaxSIM was used. Daniele Pacifico built the model.

7.1.2 Focus on medium- and long-term effects: dynamic models

Dynamic models are tools that simulate and forecast agents' behaviours over time. They have been used for many purposes such as projections of socio-demographic dimensions under current or alternative policies; for the sustainability and equity assessment of social welfare programmes intended to have substantial impacts in the long run; to combine microsimulation with macroeconomic analysis, where dynamic models serve as inputs to CGE models (Colombo 2010; Buddelmeyer *et al.* 2012; Peichl 2016). As in the previous case, the starting point is a representative sample of the population at a given time t or a mixed dataset combining survey data with administrative and/or census data at the micro level. The case where only hypothetical data are used is more rare[1]. Behaviours are modelled and simulated using regression-based analyses combined with Monte Carlo simulation techniques. Unlike static models, the ageing of the population takes place by simulating future individual and household characteristics (*dynamic ageing*). Suppose that the likelihood of event τ (e.g. to give birth, to die, to get married, to be an employee rather than a self-employed worker) is known, either because it has been calculated with a probabilistic model or because we have external information providing the likelihood of event τ divided by age and gender such as annual transition probability matrices (e.g. mortality or fertility tables). To assign a specific event to agents, the corresponding likelihood is compared with a random number u_i drawn from a uniform distribution [0.1] for each *i-th* individual. The event is assigned when u_i is lower than . The process is then repeated iteratively over time for each event following a sequential order.

The above description of the modelling framework reminds us how crucial econometric specifications and modular structure of the model are. First, finding optimal specifications that simulate with accuracy the likelihood of a number of events is not an easy task. To this end, the use of administrative data serves a great role. For example, T-DYMM's labour market module makes use of a rich unbalanced panel dataset that contains plenty of information on occupational statuses and contribution histories. On the other hand, the choice of what to simulate and when to simulate it within the structure of the model is an important one. Events are assigned following a sequential order that aims to replicate individual life path choices. Demographic and education modules usually come first than labour market modules, just as the pension and tax modules must come after the simulation of gross work-earned income. Mixing up the sequence of events so as to unrealistically depict life path choices can lead to biased projections.

Using the classifications put forward in O'Donoughe (2001), there are dynamic models that simulate behaviours of the entire population (*population models*) or a population group (*cohort models*). In both cases, individuals are aged using the same modelling framework. The only difference is that cohort models age a specific cohort following

[1] Hypothetical data are artificially created data for simulation purposes. See Hufkens *et al.* (2019) for an example.

each individual till the last period, when death occurs, while population models take all individuals to a given time *t* in the future. Another distinction that can be made is whether the model is *open* or *closed*. A closed model allows individuals to get married just with other individuals of the starting sample. On the contrary, open models artificially create spouses of sample individuals who experience the marriage event, making it more difficult to ensure population representativeness. In both cases, new-borns and migrants are externally generated by a cloning routine. Last, dynamic models are characterised by treating time as *discrete* or *continuous*. When time is discrete, behaviours are simulated at each period *t* (usually year) based on the predefined sequential order, and status transitions take place only at the beginning of the new period $t+1$ (e.g. from employed at time *t* to unemployed at time $t+1$, which means that the *i-th* individual will experience at least one year of unemployment). When time is continuous, instead, survival functions are used to simulate the time of events, and status transitions can occur at any period of time, allowing for a more precise computation of the effective time spent in specific statuses. Though more theoretically appealing, the treatment of time as a continuous variable poses a series of challenges starting with data availability and estimation problems (Li *et al.* 2014a). The functioning of continuous-time models was extensively treated in Willekens (2006), where an accurate explanation of the *quantile function approach* used in such modelling framework is provided.

In conclusion, the use of alignment procedures in dynamic microsimulation is worthy of mention. The importance of alignment lies in the limited predictive capability of dynamic models. The essential aim of such models is not to provide forecasts of demographic or macroeconomic variables; rather, these models incorporate demographic and macroeconomic projections and provide insights on aggregate and distributive traits of the simulation sample concerning the labour market, the tax-benefit system, the pension system etc.

Projecting the future characteristics of the simulation sample on the sole basis of the model's equations may lead to totals that hardly match with official projections or expected patterns. For these reasons, the use of alignment procedures is generally regarded as a fundamental step towards the construction of a coherent projection of the model's baseline scenario, although the debate on what variables should be aligned and how extensive alignments should be is far from settled (Baekgaard 2002; Li *et al.* 2014a).

The alignment consists of a calibration technique used to constrain the output of the model to externally derived aggregates (Scott 2001). The number of individuals who experience a specific event τ (e.g. giving birth, dying, being employed) can be aligned to projections (e.g. fertility rates, mortality rates, employment rates), often specified by age and gender. The selection of individuals may be carried out either randomly or by ranking them according to the likelihood to experience the aligned event – what is defined in the literature as *sorting-based alignment method* (Li and O'Donoughe 2014a). Note that reference projections can either be official ones made available by third-party institutions or autonomously computed by using macroeconomic models such as

Computable General Equilibrium (CGE) ones. In the latter case, the CGE model outputs serve as inputs for the alignment of dynamic models in a top-down setting (Peichl 2016).
A list of all dynamic models developed in Italy and main features is provided in Table 7.2.

7.1.3 Policy responses within a dynamic microsimulation framework

Given plausible assumptions on future macroeconomic aggregates, dynamic models have been primarily used to shed light on the long-term outcomes of current policy settings and alternative scenarios. Individual attributes are projected over time by making use of transition probabilities and sampling techniques. The focus of this paragraph is on how responses to policy changes should be modelled and incorporated within a dynamic modelling framework. Research efforts have been mainly focused on the inclusion of retirement decisions and labour supply responses, although other adjustments may be modelled, such as benefit take-up, consumption patterns and health self-care practices. In what follows, we will explore the possibility of enhancing the T-DYMM's simulation capabilities in light of the current state-of-the-art in this field. However, it needs to be reminded that the modelling of behavioural responses relies on the crucial assumption that simulated policy changes do not modify agent preferences, whether responses are identified using structural behavioural models or reduced form models.
At first, it is worth noting that the alignment procedure constitutes a first way of analysing alternative policy scenarios under certain circumstances. An example of such use has been sketched in O'Donoughe *et al.* (2010). As the authors argue, a potential solution for the inclusion of behavioural interactions is "to compare the average (pre-alignment) event value, such as the average transition rate or average earnings in the baseline scenario, with the average in the alternative scenario. One potential method is to increase alignment values by proportional difference". The attractiveness of such modelling strategy lies in its simplicity of application. At the same time, the approximation of agents' behavioural responses might be a too unrealistic assumption to rely on. As the key factor in microsimulation studies, agents' heterogeneity can be better captured with more properly microfounded approaches[2].
As noted by Li *et al.* (2014b), dynamic models could make use of "structural behavioural models, reduced form statistical model (*probabilistic regression analysis*) or as simple transition matrix to simulate change". Then, more specifically: "Behavioural models are grounded in economic theory, in the sense that changes in institutional or market characteristics [*such as exogenous changes in policy parameters or labour demand characteristics*] cause changes in the individual behaviours through an optimisation process" …

2 Keep in mind that the use of alignment procedures for a more coherent representation of the baseline scenario can coexist with alternative ways of including responses within a dynamic framework.

Table 7.2 Dynamic models in Italy

Name	Reference	Survey data	Administrative data	In operation	Cohort/ Population	Open/ Closed	Discrete/ Continuous	Main feature
CAPP_DYN	Mazzaferro and Morciano (2012); Tedeschi *et al.* (2013)	IT-SILC	No	No	Population	Closed	Discrete	Analysis of the evolution of private wealth and distributional effects
DYNAMITE	Ando and Nicoletti-Altimari (2004)	SHIW	No	No	Population	Closed	Discrete	Relationship between demographic structure and saving rate
IrpetDin	Maitino and Sciclone (2009)	IT-SILC	Yes (INPS)	Yes	Population	Closed	Discrete	Only example of regional model (Tuscany)
Italian Cohort	Baldini (2001)	SHIW	No	No	Cohort	Closed	Discrete	Redistributive effect of welfare state intervention
Michelangeli and Pietrunti's model	Michelangeli and Pietrunti (2014)	SHIW	No	Yes	Population	Closed	Discrete	Focus on households' indebtedness and debt-service ratio to monitor financial vulnerability
MIND	Bianchi *et al.* (2005)	SHIW	No	No	Population	Closed	Discrete	Emphasis on model alignment and validation issues
T-DYMM	Caretta *et al.* (2013)[1]; IESS (2016)[2]	IT-SILC	Yes (INPS)	Yes	Population	Closed	Discrete	Financial sustainability and adequacy of the social security system

Note: IT-SILC stands for 'European Union Statistics on Income and Living Conditions' for Italy; SHIW is the 'Survey on Household Income and Wealth' of the Bank of Italy; while INPS means 'National Social Security Institute'. See O'Donoughe (2001) and Li and O'Donoughe (2013) for an international review of dynamic models. In the table above, agent-based models are not included since they are traditionally distinguished from dynamic models in the literature (Bae *et al.* 2016). For an example in the Italian context, see Leombruni and Richiardi (2006).

[1] It refers to the first version of the model (T-DYMM 1.0).

[2] It refers to the second version of the model (T-DYMM 2.0).

"Reduced form models and transition matrices are often used to simulate mortality, fertility, family formation, labour market transitions etc. As these models usually do not depend on policy parameters, they are often restricted to simulating status quo, and are not suitable for reform analysis".

Despite their limited use for such modelling purposes, reduced form statistical models could be employed for capturing the magnitude and size of behavioural adjustments following a policy change. However, this is dependent on the inclusion of policy parameters (e.g. a covariate in the transition equation identifying the presence or absence of a specific policy setting, as well as its intensity) and the use of a dependent variable that effectively captures how individuals tend to react to policy changes.

The dynamic model of the Bank of Italy (DYNAMITE) (Ando and Nicoletti-Altimari 2004) gives a valid example of behavioural modelling using probabilistic regression analysis. The model was entirely built on SHIW data. The expected retirement age at the time of the survey was used to identify individual behavioural responses. Among the covariates, economic incentives such as the internal rate of return of several pension schemes and expected losses in the form of social security entitlements were included. A first specification was used to determine the expected retirement age once individuals enter the labour market for the first time. Then, by means of a second specification, which also takes account of *age variant characteristics* such as household composition changes, number of dependent grown children, partner retirement and temporary monetary shocks, the expected retirement age was recalculated annually within the age interval that goes from 50 to 57 years of age, the latter being the minimum required age.

Keeping the focus on the retirement decision, a structural behavioural model is implemented in SADNAP (van Sonsbeek 2011). The framework described here is the *option value model* put forward by Stock and Wise (1990).

Individuals choose the optimal age to retire maximising the expected lifetime utility made of labour and retirement income separately. For each year included in the discrete age interval taken as reference (e.g. 60-69), the expected value of both income components needs to be estimated. The choice of retirement age is made when individuals are 60 years old, supposing that they were all born on January 1. This means that labour income is unknown in the first and following years of the age interval when the retirement decision is planned (as it is for retirement income). The value function is therefore defined as follows:

$$V_t(R) = \sum_{s=t}^{r-1} \left(\frac{1}{1-p}\right)^{s-t} p(s|t)\, U_y((Y_s)^\gamma) + \sum_{s=r}^{S} \left(\frac{1}{1-p}\right)^{s-t} p(s|t)\, U_b(kB_s(R)^\gamma)$$

where t is equal to 1 (60 years old in our age interval); s ranges within the age interval, extremes included; p is the rate applied to discount future income flows; $p(s\,|\,t)$ stands

for the survival probabilities in the *s-th* year; γ is a parameter that captures individual risk aversion; $B_s(R)$ are the pension benefits received while retired; finally, S is the year in which the probability of dying is equal to 1[3].
A few models have attempted to take account of labour supply responses in dynamic microsimulation thanks to structural behavioural models. One of the first contribution came from the Swedish model MICROHUS (Klevmarken and Olovsson 1996). Behavioural adjustments to policy changes were evaluated in light of a major reform on the taxation of personal income and consumption, which took place in Sweden in 1991. While increases or decreases in the simulated annual working hours were computed using the Hausman-type model (Burtless and Hausman 1978; Hausman 1979), which would fall into the category of continuous models based on the classification proposed above, other behavioural adjustments such as moving to a new house or changes in household composition were included by means of probabilistic regressions. The isolation of labour market responses from all the remaining ones was conducted by comparing policy scenarios that differ for the kind of behavioural response considered. A first scenario including all adjustments accounted for was compared with a specular one except for labour supply responses, which were kept constant. The latter scenario was then compared with a further one that did not include any of the behavioural responses described above, therefore allowing for a complete separation of effects.
Other examples of adjustments on the labour supply side can be found in dynamics models such as SESIM (Ericson and Hussenius 1999) and LIAM (O'Donoughe *et al.* 2010), where a discrete framework rather than a continuous one was implemented. This way of approaching labour supply adjustments is even strengthened by the use of CGE models' outputs, where labour demand rigidities and constraints can be better tailored to meet the current functioning of a complex system such as the economic one.

7.2 The new features of T-DYMM

Following the taxonomy given in the above, T-DYMM falls into the category of dynamic microsimulation models.
Compared with other dynamic models, both locally and worldwide, T-DYMM has the great advantage of using a unique set of data. Its distinguishing feature is that it relies on the full individual's working history for the modelling of labour trajectories

---

3 For our purposes, S should be set equal to the individual's life expectancy defined by the current contributory pension system in Italy. This different specification with respect to the original model is driven by the way in which pension annuities are determined. The sum of contributions paid during the entire working life in addition to the interest accrued is divided by the life expectancy at the age of retirement in order to determine the amount of annual benefits granted.

and pension entitlements, as well as on the use of individual tax returns for a more precise simulation of the tax-benefit system. In addition to this, the improvements in the model's coverage and in its predictive capabilities make T-DYMM a crucial tool for the socio-economic modelling of the Italian welfare system. In a context where a few dynamic models are still active in Italy (see Table 7.2), T-DYMM is of great value not just for the microsimulation community in itself but also to better inform the policy makers and raise awareness on the medium- and long-term impact of existing phenomena.

7.2.1 Getting to T-DYMM 3.0: the previous versions of T-DYMM

The first version of the dynamic microsimulation model (DMSM) T-DYMM – hereafter T-DYMM 1.0, see MEF-FGB (2012) – significantly benefitted from the experience of the MIDAS-IT model (Dekkers 1999), a DMSM previously developed by ISAE (the Italian Institute for Studies and Economic Analyses)[4].

T-DYMM 1.0 is a dynamic ageing, discrete-time model: transitions in all updating processes are carried out year-by-year and are achieved by means of probabilistic methodologies (Monte Carlo technique). It is a closed model: new individuals enter the sample each year as new-borns, but migration flows are not simulated.

T-DYMM 1.0 inherits from MIDAS-IT the general structure and the simulation platform LIAM (O'Donoghue *et al.* 2009). Differently from MIDAS-IT, T-DYMM 1.0 contains in the fiscal module the main elements of EconLav (Coromaldi and Guerrera 2009), a static microsimulation model of the Italian tax-benefit system developed by ISFOL[5], with the support of the Italian Ministry of Economy and Finance and the Italian Ministry of Labour, for the analysis of the effects of tax and benefit system reforms.

The stylised structure of the model T-DYMM 1.0 consists of three main modules (demographic, labour market and pension modules) linked to each other by recursive feedbacks (i.e. in the same period, the causal relationship is unidirectional), integrated with a fourth external module (on taxation, running on Stata).

T-DYMM 2.0 retains largely the same architecture of its previous version, while innovating on a number of crucial aspects. First, T-DYMM 2.0 makes use of a new simulation platform, LIAM2[6], which represents a natural evolution of the previously employed LIAM and provides considerable improvements in terms of computation speed and data capacity. Concerning the structure of the model, T-DYMM 2.0 saw the development of two sub-modules on unemployment benefits and on private (complementary) pension schemes.

---

4 The model was developed in the context of AIM, a European-funded sixth framework project (see Dekkers *et al.* 2009).

5 The Italian research institute for vocational education and training employment and social policies. Since November 2016, ISFOL has been renamed INAPP (National Institute for Public Policy Analysis).

6 See: https://liam2.plan.be/.

The version of the model currently under development (T-DYMM 3.0) will be characterised by several innovations, stretching from an improvement of the representativeness of the starting sample, to the addition of new modules and the enrichment of existing ones, to the update and the expansion of the institutes and the legislation simulated within T-DYMM. The most important novelties are:

- A procedure to ensure better representativeness of results through the recalibration of sample units for T-DYMM's starting data set;
- The development of a *Migration Module*, where both immigration and emigration flows are modelled;
- A rethinking of the sequential structure of the *Labour market Module* that allows for multiple simultaneous choices by implementing the parameters of a multinomial logit analyses, whereas the previous versions of T-DYMM were solely based on binomial regression analyses;
- As far as the *Pension Module* is concerned, working pensioners are to be included in the simulations. In previous releases of T-DYMM, retirement and work were simplistically treated as incompatible, but the relevance of the phenomenon in real data, the far-reaching pensionable ages in years to come and the expected strengthening of the economic incentives to keep working, once the Notional Defined Contribution (NDC) computation rules settle in, suggest a less stylised representation of pensioners is due;
- The inclusion of the *Taxation Module* within the LIAM2 framework, which greatly improves the speed of the model. The availability of data on individual tax returns (see Paragraph 1.3 of this Report) will greatly benefit the update and development of the *Taxation module* and will allow for a more comprehensive analysis on income inequality;
- The development of a *Wealth Module*, aimed at modelling the whole spectrum of movable and immovable assets at the household's disposal (real and financial wealth, including a better design for the private pension sub-module). This module will further improve the simulation of means-tested benefits, where wealth information is also needed;
- The inclusion of a number of new institutes within the model, so as to make T-DYMM 3.0 suitable for broader analyses of the Italian welfare system. Most notably, we will update the welfare legislation by simulating the so-called Citizenship Income (*Reddito di Cittadinanza*), an enhanced minimum income scheme in effect since March 2019, and the so-called *Quota 100*, an alternative early-retirement scheme in effect since February 2019.

As additional developments, the working team will also explore the possibility of:

- Including behavioural responses in the choice of retirement age, possibly by implementing the modelling framework outlined in Stock and Wise (1990);

- Modelling the *death* and *disability events* as causal processes not randomly determined as in the previous versions of the model. By making use of the newly acquired information on health statuses contained in the latest version of AD-SILC and expanding the range of institutes that address the disability issues simulated in the model, a proper Disability model can be developed, thereby making T-DYMM 3.0 fit to assess the impact of policy changes in that area of the welfare system as well.

In the next paragraphs, we shall examine in more detail some of the innovations presented here. Before addressing the renovated modular structure of T-DYMM 3.0, let us examine the improvements in the starting datasets.

7.2.2 The new starting dataset of T-DYMM 3.0

The use of population-representative samples is the starting point of almost all models for a good depiction of the simulated reality. Individual heterogeneity is fully exploited and modelled so as to simulate socio-demographic interactions and economic behaviours. However, the level of information needed often goes beyond the representativeness provided by sample surveys. One can be interested in studying how a hypothetical reform on a small, specific subgroup of workers would affect them in terms of their future group-specific outputs, and how this would then affect overall inequality levels. To do so, the interested sample group must be a good representation of the corresponding one for the population as a whole. At the same time, microsimulation models can also be used to estimate aggregate and average values (e.g. overall pension expenditure), which also requires the appropriate use of weighting techniques. Given the above, T-DYMM 3.0 shall make use of a well-established technique for the reweighting of sample units (Creedy and Tuckwell 2004; Pacifico 2014) in order to execute our simulations on a dataset that realistically represents the many dimensions we are interested in. The recalibration is performed on the cross-sectional AD-SILC wave for the year 2017[7], which is the starting point of our model. Total frequencies of different sample subgroups are calibrated (at the individual level) to external aggregates made available by the Italian Department of Finance[8] and the Italian National Institute of Statistics (ISTAT)[9].

The weight recalibration is obtained by minimising the Lagrangian function that follows, with respect to , the recalibrated weights:

7 See Chapter 1 for a detailed description of the AD-SILC dataset.

8 See https://www1.finanze.gov.it/finanze3/pagina_dichiarazioni/dichiarazioni.php, where plenty of information on aggregate tax returns are provided for each tax year.

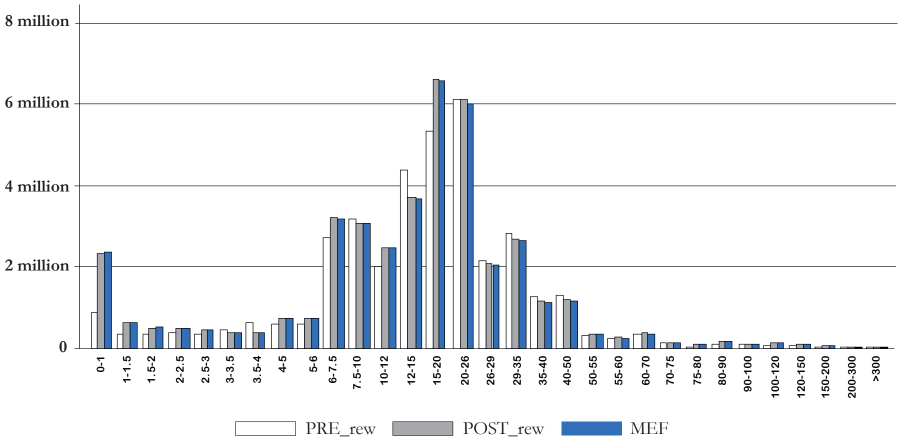
9 See http://dati.istat.it/, where data on age divided by classes, number of foreigners and highest educational attainment were taken as reference in the reweighting procedure.

$$L = \frac{1}{2}\sum_{j=1}^{k}\frac{\left(rw_j - ow_j\right)^2}{ow_j} + \sum_{k=1}^{m}\lambda_k\left[t_k - \sum_{j=1}^{k} rw_j x_j\right]$$

where *ow* is the original weight; $(rw_j - ow_j)/ow_j$ is the chi-squared distance function for the j-th individual; λ_k is the k-th Lagrange multiplier; t_k is the k-vector of external totals; x_j is the vector of the variables that are object of the reweighting.
The focus of the recalibration concerns different dimensions. So far, we have considered the following: number of taxpayers, employees, self-employed workers and retirees with positive gross income subject to the personal income tax (*Imposta sul Reddito delle Persone Fisiche*, IRPEF) by income group and geographical area; individuals and householders by gender; householders by number of family members and household typology; immigrant individuals and householders by gender and macro-area of birth; individuals by age class; individuals of at least 15 years of age by highest educational attainment.
Subsequent to the reweighting procedure on the AD-SILC 2017 dataset, individuals are duplicated in the sample as many times as their adjusted sample weights require. Then, once the dataset is expanded, householders are selected and sampled with repetition so as to extract a sample of 100,000 householders. Finally, the remaining members of the households are combined with their corresponding householders. The procedure described here is a common practice in dynamic microsimulation studies and allows the overtaking of the difficulties that modellers meet when using alignment methods, although alternative strategies that do not involve the expansion of the dataset have also been proposed (Dekkers and Cumpston 2012).
In what follows, the accuracy of the recalibration method is tested making use of the AD-SILC 2011 dataset, the base year of T-DYMM 2.0 (IESS 2016)[10]. The reference income distribution is gross income subject to Personal Income Tax (PIT) as of the 2011 tax year. The original IT-SILC weights were recalibrated for the variables listed above. As shown in Figure 7.1, the reweighted distribution of individuals with positive gross income subject to PIT adheres almost perfectly to external totals – which are defined below as the MEF distribution. The non-reweighted distribution underestimates the number of individuals falling into the first income group (0-1,000) by over a million, a typical example of how sample survey data fails to adequately represent the tail values of a distribution. The reweighting procedure also proves its usefulness in addressing the misreporting of middle-income individuals. This is especially true for two of the three most numerous income groups (12,000-15,000 and 15,000-20,000).

[10] AD-SILC 2017 could not be tested at this point, as it was not available at the time of the analysis.

Figure 7.1 Individuals with gross income subject to PIT by income group (values on the horizontal axis in thousands of euros)

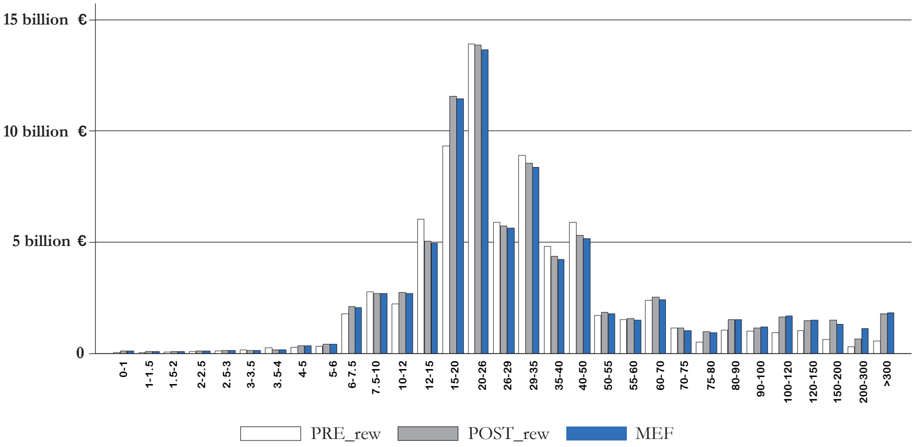


Source: own elaborations on IT-SILC 2011 and MEF data

As a result of these adjustments, reweighted total gross income subject to PIT better represents the administrative aggregate than the non-reweighted scenario for all the income groups involved (see Figure 7.2). The improvements are especially noticeable for the middle-income groups mentioned above and on the right tail of the income distribution. The pre-tax Gini index increases from 0.4043 to 0.4513 when the non-reweighted and the reweighted distributions are compared, while income inequality in Italy as calculated on individual tax returns ranges in the interval [0.45; 0.46] (Di Nicola *et al.* 2015; Di Caro 2018).

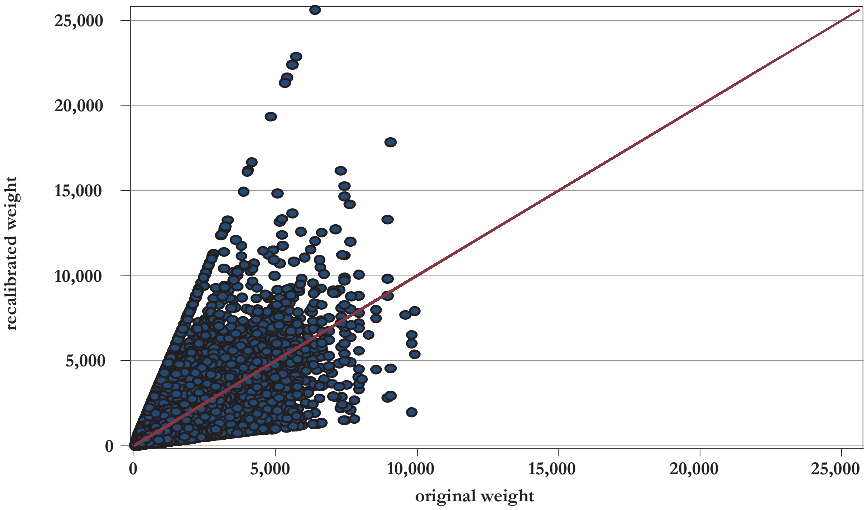
In Figure 7.3, the recalibrated weights are plotted against the original weights in order to give an intuitive idea about the distortions introduced by the reweighting procedure. What can be noted is that, with the exception of the outliers in the upper part of the graph, the new weights assigned to individuals are not profoundly different from the old ones even when the recalibration interested a substantial number of dimensions as in our case.

Figure 7.2 Total gross income subject to PIT by income group: values on the horizontal axis in thousands of euros



Source: own elaborations on IT-SILC 2011 and MEF data

Figure 7.3 Scatterplot of recalibrated weights against original weights for the AD-SILC 2011 dataset



Source: own elaborations on AD-SILC 2011 dataset

7.2.3 The new structure of T-DYMM 3.0

Figure 7.4 offers a schematisation of the modular structure of T-DYMM 3.0. The baseline year is 2017, corresponding to the latest vintage of IT-SILC at our disposal, and the simulation runs for 53 years until 2070, the last year of the macroeconomic projections of the Ageing Working Group (Economic Policy Committee of the European Commission)[11], employed in the alignment procedures.

The new version of the model shall include:

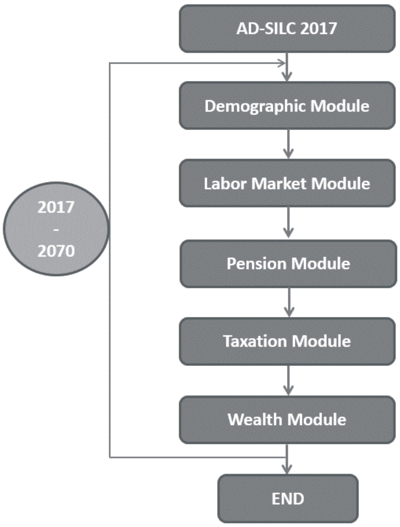
1. A **Demographic Module**, which encompasses intergenerational persistence, birth processes, migration flows, educational achievements and the marriage market (*de facto* relationships/separations, marriages/divorces). The main innovation with respect to T-DYMM 2.0 is the inclusion of the Migration block[12], while works on a disability sub-module are also underway;
2. A **Labour market Module**, which probabilistically simulates individual labour market dynamics, namely employment transitions (in and out of the labour market and among employment categories, sectors and contractual arrangements), labour incomes and possible unemployment spells (if requirements are met, unemployment benefits are attributed). In T-DYMM 3.0, upon in-depth analysis of AD-SILC, of INAPP surveys and of the newly available data from the Finance Department (see Chapter 1), the Labour market Module shall include sub-sections dedicated to working pensioners, migrant workers and non-standard workers, for whom we expect dedicated work patterns shall be designed;
3. A **Pension Module**, for the definition of eligibility requirements and retirement decisions and for the computation and indexation of pension benefits. The main additions to the Pension Module will consist in the update of the relative legislation and in the integration of a sub-module on working pensioners;
4. A **Taxation Module**, which computes net labour and pension incomes. As for to the Pension Module, the pertinent legislation will have to be updated for the Taxation Module as well. Additionally, the newly acquired information contained in the Tax Declarations (see Paragraph 1.3) will enable the simulation of new tax-free cash benefits (e.g. family and housing benefits);
5. A **Wealth Module**, which simulates real and financial wealth (including private pension schemes). This module will expand on the work done in IESS (2016) for private pensions and include other relevant components of wealth, basing the estimates on SHIW and SILC data on one side and on administrative data from the Finance Department on the other.

[11] For the scopes of the MOSPI project, the projections of the 2021 Ageing Report will be employed. See European Commission (2017).

[12] The Migration Module technically operates as a sub-module within the Demographic Module, but extends its reach into other Modules, as specific care will have to be paid to migrant workers and migrant pensioners respectively in the Labour Market Module and in the Pension Module.

While future reports will offer a detailed description of the functioning of each module, in what follows we shall focus on the main structural innovations of T-DYMM 3.0, which concern the introduction of a Migration Module and of a Wealth Module. A new sub-model on working pensioners is also set to be developed.

Figure 7.4 Structure of T-DYMM 3.0



7.2.4 The sub-module on working pensioners

Since 2001, old-age pension benefits and early-pension benefits computed on over 40 years of contributions can be cumulated with labour incomes. Since 2009, all early pensions can be cumulated regardless of the years of contributions on which the pension has been computed[13].

Earlier versions of T-DYMM have treated retirement as an absorbing state, i.e. retirement is not compatible with work. The sequencing structure of the model favoured a simplification in this sense, and the relative low number of working pensioners supported it. The number of working pensioners has appeared to be decreasing in the past few years: according to a 2020 report from ISTAT, the figures went from 492 thousand in 2011 to 406 thousand in 2018. That can be explained by the quick rise in pensionable ages that the 2011 Pension Reform (so-called 'Fornero Reform') has caused.

However, the incidence of the phenomenon could vary consistently in the future, since the economic incentives to postpone retirement are bound to strengthen as the Notional Defined Contribution (NDC) pension scheme settles in.

[13] Some limitations are still in place for disability pensions.

With regard to the observed evidence and taking account of the expected evolution of the population of retirees, T-DYMM 3.0 shall devote a sub-model to working pensioners. Dedicated working patterns will be designed and simulated in a sub-section of the Labour market Module according to the parameters estimated on the AD-SILC dataset and the appropriate contribution accrual and taxation rules will be applied. This development shall guarantee a more accurate representation of the population of retirees within the model, as well as allow the possibility of testing policy options specifically directed to elderly workers, a population projected to grow consistently in the coming decades.

7.2.5 The Migration Module

In a global context where some countries face the secular challenge of ageing populations, while others are in the midst of a significant demographic growth, migration pressures are bound to play a significant role in the economic, social and political spheres for years to come.

The modelling of migration has been a crucial step forward in the predictive capabilities of dynamic microsimulation models (O'Donoghue *et al.* 2010). Despite the many challenges that modelers still face, immigration and emigration flows can be simulated with a satisfactory level of accuracy. Setting aside the computational challenges of the process, some large empirical evidence is required to support crucial decisions in the simulation analysis.

First, information on age, gender, country of origin, birth and citizenship, educational achievement and household dimension are some of the essential aspects needed to characterise the migrant population in the data and understand how it should be cloned and included within the simulation model (Duleep and Dowhan 2008a; 2008b).

Secondly, one is faced with the demanding task of accounting for migration drivers. In truth, individuals move for a multitude of reasons. The search for better living conditions may well be considered as the main driver, but that often cannot be simplified as 'search for higher purchasing power'. Access to services and welfare in the country of destination can play a significant role in the decision to migrate, as can a number of cultural and contingent reasons (the outbreak of an armed conflict, climate change and its consequences).

The limited availability of data and the inherent complexity of the migration phenomenon leave wide margins of discretion to modelers when it comes to building the design of a dedicated module.

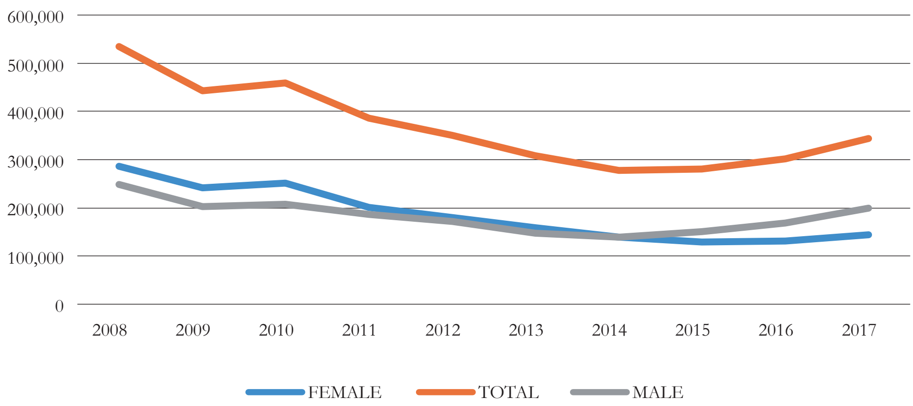
What follows offers an outlook on the migration phenomenon in Italy and outlines how we intend to develop a migration module within T-DYMM 3.0, stemming from the work of Dekkers (2015a).

What do data tell us?

The year 1971 saw the closing of a migration cycle begun over a century before: for the last time since its unity, Italy registered a negative balance between immigration and emigration flows. Since then, a growing trend has seen Italy transform from an emigration to an immigration country, with a peak of foreign registrations in 2007 following the access to the EU of Romania and Bulgaria (AISP 2017).

The outbreak of the economic crisis in 2008 abruptly interrupted this growing trend. Figures 7.5 and 7.6 report immigration flows to Italy by sex and macro-area of birth[14] over the 2008-2017 period. The number of people arriving has seen a substantial decrease throughout the time span considered (from 534,712 to 343,440), although the overall flow shows an increasing pattern in the last three years considered. The decrease in the number of arrivals has been more marked for women, as from 2014 onwards; male immigrants have represented a higher share of the new arrivals.

Figure 7.5 Immigration flows by sex



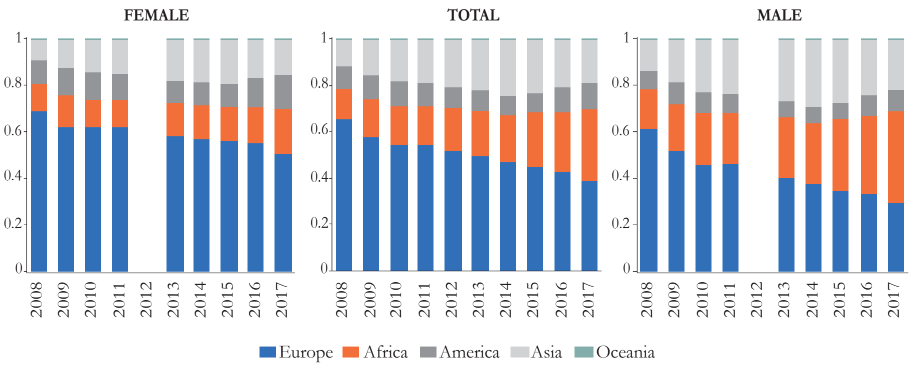
Source: own elaborations on EUROSTAT data

As Figure 7.6 points out, European immigrants are still the most consistent subgroup when considering male and female foreigners together, but it is interesting to note the increasing relevance of African newcomers compared with the decreasing frequency of the European ones, and this is particularly true for male immigrants. After almost thirty years where Eastern European immigrants were by far the most numerous

[14] In the current paragraph, we have made the conscious choice of focusing on foreign-born individuals rather than non-citizens: indeed, for the purpose of the migration module, we are interested in all individuals who enter the model without being born within it.

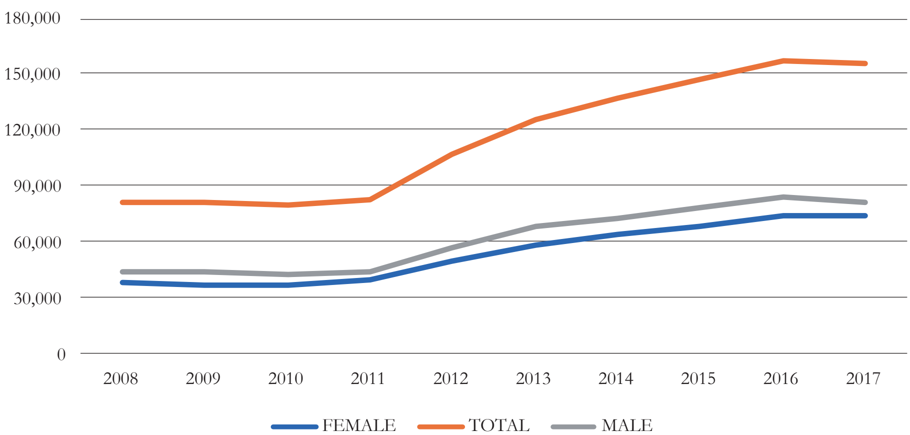
subgroup – the case of the first Albanian exodus to the Italian coasts in 1991 is emblematic – the arrival of people who were born in Africa is likely to continue growing in the near future. Amongst other factors, that is due to a structural surplus of local labour supply (Menonna and Blangiardo 2018). According to the *World Population Prospects 2019* of the United Nations, the population of sub-Saharan Africa is expected to double its current volume by the 2050, reaching 2.1 billion people[15].

Figure 7.6 Distribution of foreign-born immigrants by continent of birth and sex



Note: disaggregate data not available for the 2012 year.
Source: own elaborations on EUROSTAT data

Figure 7.7 Emigration flows by sex



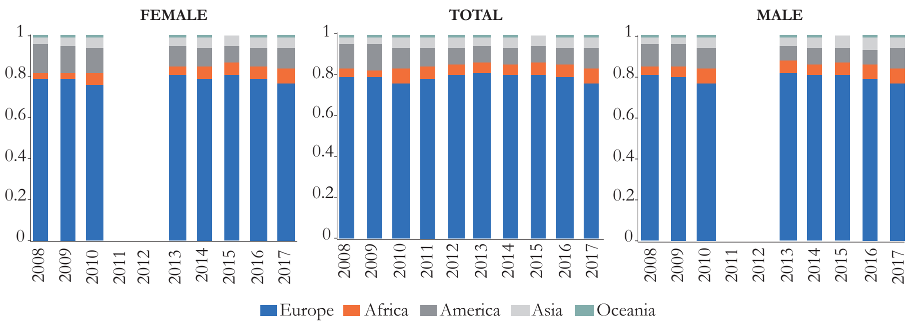
Source: own elaborations on EUROSTAT data

15 See: https://population.un.org/wpp/.

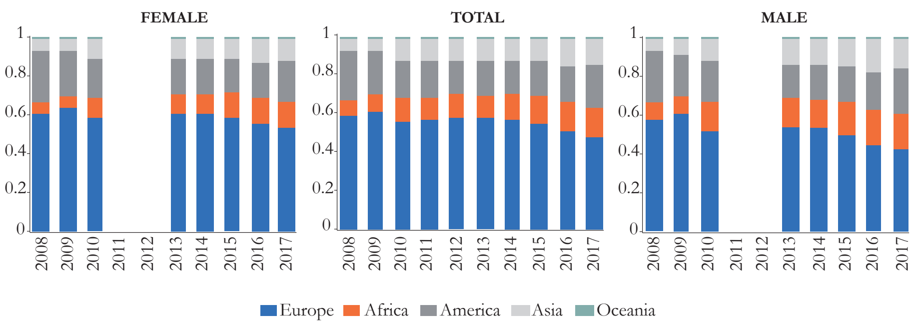
As far as emigration flows are concerned, the number of people leaving the country grew rapidly during the 2008-2017 period (from 80,947 to 155,110), with a growth rate peak that coincides with one of the toughest years faced by the Italian economy in the recent crisis (2012). Stable during the 2008-2011 period, the flow of emigrants grew almost constantly until 2016, then stabilised in 2017 (see Figure 7.7). Male and female flows show a similar pattern throughout the years.
Eight out of ten individuals emigrating from Italy are born in Europe – five of which are born in Italy – a ratio that is relatively constant through the whole period for both sexes. America, Africa and Asia share the remaining shares of emigrants (Figure 7.8.1).
Figure 7.8.2 shows that the foreign-born European component of emigrants has decreased over the years: inflows from European countries (Romania) have diminished in the past few years, and a good portion of those longer-term immigrants appear to have stabilised in the country.

Figure 7.8 Distribution of emigrants by macro-area of birth and sex

1. All emigrants



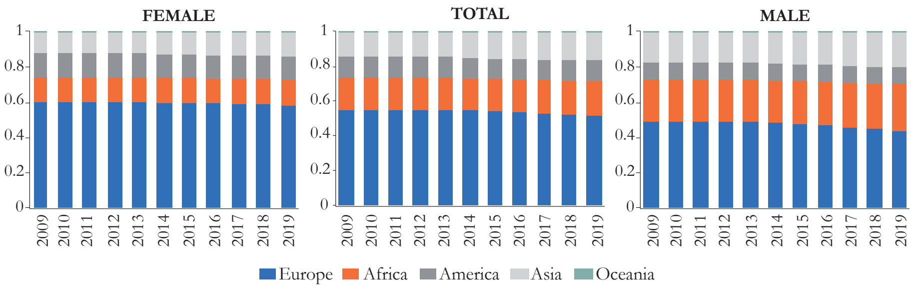
2. Foreign-born emigrants



* Data for Europe exclude Italian-born individuals.
Note: disaggregate data not available for the 2011 and 2012 years.
Source: own elaborations on EUROSTAT data

The evolution of migration flows over time has affected the characteristics of the foreign-born resident population in Italy. Figure 7.9 shows how between 2009 and 2019 the share of European foreigners decreased visibly, mostly in favour of Africans and Asians. One should notice how the variation is essentially due to the male component, whereas the composition of female foreigners has been quite stable through the past decade.

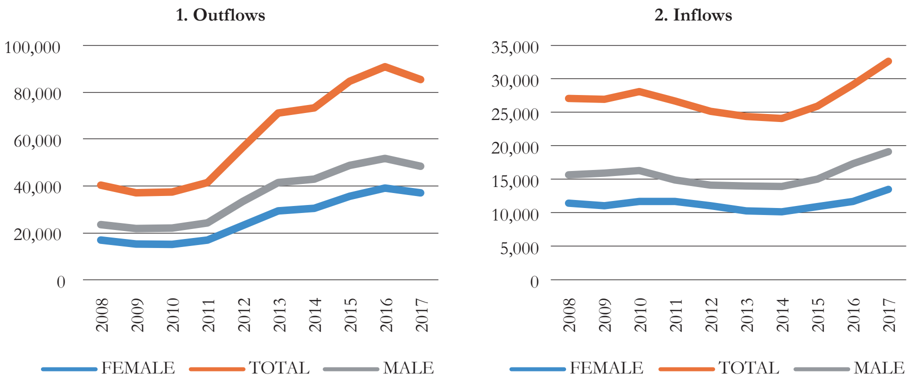
Figure 7.9 Stock of foreign-born residents by continent of birth and sex



* Data for Europe exclude Italian-born individuals
Source: own elaborations on EUROSTAT data

Last but not least, Figures 7.10.1 and 7.10.2 compare the outflow (individuals leaving the country) and the inflow (individuals returning to the country) of people born in Italy for each of the years considered. The first thing that one should notice is that Italy is characterised by having a severe negative balance that worsens over time. The gap slightly narrowed in 2017, both due to a decrease in expatriations and to an increase in repatriations compared to the previous year.

Figure 7.10 Migration flows of Italian-born individuals



Source: own elaborations on EUROSTAT data

Almost half of all emigrants from Italy were born in the country: the economic downturn has led to the emigration Italian-born and foreign-born residents alike, as the comparison between Figure 7.7 and Figure 7.10.1 demonstrates (the most interesting difference appears to be the relevantly lower predisposition of Italian-born females to migrate compared to foreign-born females). Finally, the reprise in migration inflows in the 2015-2017 period shown in Figure 7.5 appears in line with the growth in Italian-born inflows in the same period (Figure 7.10.2).

The modelling strategy of migration

The development of simulation strategy for the migration phenomenon within T-DYMM 3.0 starts from a detailed description of the level of representativeness that the 2017 IT-SILC dataset ensures when it comes to the stock of immigrants living in Italy. In Table 7.3, the weighted totals of immigrants by country of birth and sex are compared with Eurostat aggregate data to assess whether immigrants are well represented in the cross-sectional survey that constitutes the base year of T-DYMM 3.0. The values referring to the columns 'IT-SILC' are weighted using the original IT-SILC weights prior to any recalibration. The categories shown for countries and wider geographical areas are those listed in IT-SILC and cover all EU member states, some additional selected European countries, Canada, the USA, and eight macro areas. Table 7.3 shows how, for some countries of origin, SILC cannot guarantee representativeness. Even if one does not specify by gender, no individual from CY, HR, MT, IS, ME, MK, RS, TK, USA and CAN was surveyed in SILC 2017. According to Eurostat data, these individuals represent about 4% (240,755) of the stock of foreign-born residents in Italy in 2017 (6,020,614). Overall, the misrepresentation decreases as the size of the immigrant subgroup increases in relative terms (see the columns 'Diff.'). The three most numerous categories are other Europeans ('OEU'), Romanians ('RO') and North Africans ('NAF') respectively, and for all three the difference between weighted sample and population data does not exceed 10% in absolute terms.

Even if no existing model designs migrant classification at the level of detail shown in Table 7.3, modelers should be aware of the inherent discrepancies in representativeness that the use of sample survey data might impose to the modelling of immigration. As a result of our analysis, we have opted to include immigrants' sex and macro-area of birth among the dimensions considered in the recalibration algorithm described above (see Paragraph 7.2.2).

Table 7.3 Representativeness of IT-SILC 2017 stock of foreign-born residents

COUNTRY OF BIRTH*	MALE			FEMALE			TOTAL		
	IT-SILC	Eurostat	Diff.	IT-SILC	Eurostat	Diff.	IT-SILC	Eurostat	Diff.
AT	1,409	4,730	**-70%**	5,761	10,748	**-46%**	7,170	15,478	**-54%**
BE	32,248	19,604	**64%**	26,780	25,894	**3%**	59,028	45,498	**30%**
BG	12,511	20,556	**-39%**	24,626	39,235	**-37%**	37,137	59791	**-38%**
CY	0	186	**-100%**	0	246	**-100%**	0	432	**-100%**
CZ	1,976	1,778	**11%**	8,810	8,301	**6%**	10,786	10,079	**7%**
DE	61,837	91,179	**-32%**	120,475	119,068	**1%**	182,312	210,247	**-13%**
DK	4,872	1,052	**363%**	7,178	2,030	**254%**	12,050	3,082	**291%**
EE	0	245	**-100%**	469	1,229	**-62%**	469	1,474	**-68%**
EL	4,869	7,924	**-39%**	8,897	8,582	**4%**	13,766	16,506	**-17%**
ES	9,393	10,327	**-9%**	20,441	22,716	**-10%**	29,834	33,043	**-10%**
FI	0	577	**-100%**	3,398	2,140	**59%**	3,398	2,717	**25%**
FR	37,747	50,478	**-25%**	75,440	77,601	**-3%**	113,187	128,079	**-12%**
HR	0	9,588	**-100%**	0	13,568	**-100%**	0	23,156	**-100%**
HU	5,184	3,228	**61%**	18,992	9,798	**94%**	24,176	13,026	**86%**
IE	5,427	1,503	**261%**	0	2,100	**-100%**	5,427	3,603	**51%**
LT	0	1,330	**-100%**	5,592	4,605	**21%**	5,592	5,935	**-6%**
LU	765	2,089	**-63%**	0	2,593	**-100%**	765	4,682	**-84%**
LV	0	754	**-100%**	407	2,732	**-85%**	407	3,486	**-88%**
MT	0	528	**-100%**	0	1,363	**-100%**	0	1,891	**-100%**
NL	13,008	4,870	**167%**	6,666	7,910	**-16%**	19,674	12,780	**54%**
PL	18,546	26,509	**-30%**	62,370	87,621	**-29%**	80,916	114,130	**-29%**
PT	827	2,253	**-63%**	2,682	4,175	**-36%**	3,509	6,428	**-45%**
RO	394,325	412,664	**-4%**	533,991	612,365	**-13%**	928,316	1,025,029	**-9%**
SE	0	660	**-100%**	5,280	1,065	**396%**	5,280	1,725	**206%**
SI	2,489	1,804	**38%**	7,391	2,800	**164%**	9,880	4,604	**115%**

(continued)

COUNTRY OF BIRTH*	MALE			FEMALE			TOTAL		
	IT-SILC	Eurostat	Diff.	IT-SILC	Eurostat	Diff.	IT-SILC	Eurostat	Diff.
SK	1,822	2,571	**-29%**	4,168	8,275	**-50%**	5,990	10,846	**-45%**
UK	14,672	26,298	**-44%**	56,400	36,910	**53%**	71,072	63,208	**12%**
CH	96,156	88,901	**8%**	106,381	103,302	**3%**	202,537	192,203	**5%**
IS	0	69	**-100%**	0	136	**-100%**	0	205	**-100%**
ME	0	1,263	**-100%**	0	1,550	**-100%**	0	2,813	**-100%**
MK	0	38,846	**-100%**	0	32,046	**-100%**	0	70,892	**-100%**
NO	2,147	660	**225%**	0	1,065	**-100%**	2,147	1,725	**24%**
RS	0	24,196	**-100%**	0	20,441	**-100%**	0	44,637	**-100%**
TK	0	11,201	**-100%**	0	8,484	**-100%**	0	19,685	**-100%**
OEU	520,511	452,800	**15%**	734,239	745,685	**-2%**	1,254,750	1,198,485	**5%**
NAF	406,569	417,043	**-3%**	243,507	296,826	**-18%**	650,076	713,869	**-9%**
WAF	115,890	221,963	**-48%**	88,888	100,935	**-12%**	204,778	322,898	**-37%**
OAF	23,609	**	******	40,533	**	******	64,142	**	******
CAN	0	11,097	**-100%**	0	13,752	**-100%**	0	24,849	**-100%**
CSA	223,651	206,313	**8%**	393,405	329,100	**20%**	617,056	535,413	**15%**
USA	0	22,575	**-100%**	0	29,620	**-100%**	0	52,195	**-100%**
NME	15,994	21,473	**-26%**	21,360	23,978	**-11%**	37,354	45,451	**-18%**
OAS	538,617	527,639	**2%**	443,505	428,009	**4%**	982,122	955,648	**3%**
OCE	6,664	9,064	**-26%**	8,855	11,518	**-23%**	15,519	20,582	**-25%**
OTH	2,025	**	******	0	**	******	2,025	**	******
TOTAL	**2,575,760**	**2,759,860**	-6.7%	**3,086,887**	**3,260,754**	-5.3%	**5,662,647**	**6,020,614**	-5.9%

Note: AT: Austria; BE: Belgium; BG: Bulgaria; CY: Cyprus; CZ: Czech Republic; DE: Germany; DK: Denmark; EE: Estonia; EL: Greece; ES: Spain; FI: Finland; FR: France; HR: Croatia; HU: Hungary; IE: Ireland; LT: Lithuania; LU: Luxembourg; LV: Latvia; MT: Malta; NL: Netherlands; PL: Poland; PT: Portugal; RO: Romania; SE: Sweden; SI: Slovenia; SK: Slovak Republic; UK: United Kingdom; CH: Switzerland; IS: Iceland; ME: Montenegro; MK: The Former Yugoslav Republic of Macedonia (FYROM); NO: Norway; RS: Republic of Serbia; TK: Turkey; OEU: Other European Countries; NAF: North Africa; WAF: West Africa; OAF: Other Africa; CAN: Canada; CSA: Central and South Africa; USA: United States; NME: Near and Middle East; OAS: Other Asia; OCE: Australia and Oceania; OTH: Other.

* No individuals (either males or females) were interviewed for countries of origin: CY; HR; MT; IS; ME; MK; RS; TK; USA; CAN.

** Data not available.

As a common practice in dynamic microsimulation, T-DYMM 3.0 makes use of alignment procedures to project some aggregate results (see Paragraph 7.1.2). The main sources for alignment are the Ageing Working Group Projections (hereafter AWG data), which rely on Eurostat periodical updates for what concerns demographic projections. As of today, only net migration flows projections are available per gender and age. The specification of inflows and outflows as well as the partition by area of origin are not available and will have to be developed 'in house'. A certain level of aggregation by macro-area of origin will be necessary, and it shall be based on the evidence described above and on the interpretation of such data provided by the relevant literature and the institutions that study these phenomena (ISTAT).

For what concerns immigration flows, we have seen in Figure 7.6 that the ratio of African immigrants out of all the newcomers over the 2008-2017 period grew rapidly, the incidence of European immigrants decreased and the remaining categories appear to have a somewhat regular distribution over time. This trend pattern combined with the expected increase in the African population worldwide is likely to gradually transform the face of immigration over the long run.

Concerning emigration, figures have been shown to worryingly rise in the past decade. The trend appears to have stabilised and somewhat reversed for Italian-born individuals, not so much for the foreign-born: Italy has become a destination country but has also established itself as a transition country for migrants.

Different alternatives on how to project the frequency distribution of immigration and emigration flows are currently being evaluated, but a certain level of simplification is bound to be applied.

Once alignment procedures are settled, a further step in the modelling of migration consists in the identification of target households. According to the procedure established in Dekkers (2015a), the migration process consists in the cloning (immigration) or elimination (emigration) from the simulation sample of target households, i.e. the households that we wish to include in the process because they display features that we deem characteristic of the target population.

Immigrants are identifiable in our dataset AD-SILC 3.0 through IT-SILC variables like 'db210' (country of birth), 'pb220a'/'citesa' (first citizenship) and all of the variables presented in Annex 1.1 of Section 1 of the present report.

IT-SILC also contains useful information for the identification of emigrants. Variables 'rb031'/'aita' (the year of first immigration to Italy) can be usefully confronted with 'aallo' (the year after which the household has lived in Italy without leaving for a period of one year or more, only available 2016-2017) to determine which migrant households are more likely to move again and explore their characteristics. For Italian-born emigrants, useful information could also be extracted from EU-SILC datasets for all countries (except Italy), where emigrant Italian-born individuals are surveyed, hence the working group is exploring the possibility of accessing those datasets as well.

The set of information available in AD-SILC 3.0 and in possible external sources will allow us to identify a number of homogeneous sub-groups within the sample of migrants and thus treat these different groups separately.
The next step requires the selection of the households that act as donors in the cloning procedure for newcomers, as well as of the households that emigrate.
The selection process may be directed by logit estimations or by the simple extraction of a random number u_i drawn from a uniform distribution [0,1] for each *i-th* household belonging to the *j-th* immigrant cloning group. The random number is then used to rank households and it is positively correlated with the probability of being chosen as donor household. Note that different ranking methods might be preferred to a random imputation. For example, Dekkers (2015a) ranks households according to the number of members.
Once a household is selected to emigrate at time *t*, one needs to account for the possibility for that household to re-join the sample, i.e. to move back to the country of origin. The emigrated household would have to age and keep existing while outside of the simulated Italian context, while the probability to move back estimated as a (non-linear) function of the time spent abroad. We are also considering whether to reward returning households by assigning them a higher overall income and to what extent, i.e. suppose an economic incentive is driving the choice to move back. Another important issue that relates to the modelling of emigration is how to deal with pension entitlements accumulated in foreign countries. It is an arduous problem and we are not aware of any common solutions for active microsimulation models: pension systems vary widely and the way to deal with the concurrence of pension entitlements accrued in different countries is quite dependent on the countries involved.
Once migrating household have been selected, the last step in the simulation process consists in the use of the Pageant algorithm by Chénard (2000). The method is employed as a way to ensure coherence between alignment procedures and the distribution of immigrant and emigrant individuals by sex and age class. Dekkers (2015a) offers an extensive explanation with application to LIAM2.

7.2.6 The wealth module

One of the novelties in T-DYMM 3.0 is represented by the introduction of a wealth module. Even though this topic is commonly recognised as relevant in the development of dynamic microsimulation modelling aimed at assessing the long-run distributional effects of social and economic reforms, it has hardly been tackled in the relative literature and just a few models are endowed with a module able to account for the evolution of private wealth over time. One of the exceptions is the paper by Tedeschi *et al.* (2013), in which the authors introduce a wealth module in the Italian model CAPP-DYN. We consistently draw from their work in the construction of this new module in T-DYMM.

The inclusion of private wealth in the model is crucial in assessing the overall well-being of the simulation units before and after retirement, and it is indispensable to allow for proper estimation of means-tested allowances[16]. In particular, *ex ante*, the role of wealth determines saving/investing as well as retirement decisions. Individuals who are part of wealthier households, for instance, may decide to anticipate retirement since they can count on the 'safety net' represented by the household accumulated wealth. In the Italian context, the several reforms to the pension system adopted since the '90s have sensibly reduced returns on social contributions compared to the past and younger generations are on average set to enjoy less generous public pensions than their parents did. In order to achieve consumption smoothing in the future, individuals will have to rely on private resources more than it was necessary in the past and may wish to increase their savings during working life in order to channel them into 'retirement-oriented investments' such as pension funds, real estate and other form of financial assets.

While the generality of T-DYMM modules takes the individual as unit of analysis, the Wealth Module considers the household.

The sequence of processes that incorporate the mechanisms of formation, transmission and spending down of household wealth is summarised in Figure 7.11. The first process to be modelled is the intergenerational transmission of wealth (1), followed by the update of net wealth (2), the decision to buy or sell dwellings (3) and finally the decision on consumption behaviour (4), which in turn determines how much is saved (5). The set of decisions presented in the scheme depends on a series of econometric estimations that we are going to briefly explain and discuss in the rest of this paragraph.

As mentioned above, the first simulated events are the intergenerational transfers of wealth between parents and children outside the family of origin (*inter vivos* and *mortis causa*). By using SHIW data (see Paragraph 1.3), the *inter vivos* transfers are modelled by means of a probabilistic approach based on a Heckman two-step procedure (Heckman 1979), in order to account for selection bias. The two sides of the transfer (potential donor and potential recipients) are modelled separately and then aligned. *Mortis causa* transfers or bequests instead have a mechanical functioning. Once the number of inheriting families is defined, bequests are deterministically simulated with the stock of wealth being proportionally distributed among them in the form of financial wealth.

In the following step (2), the module updates the wealth stock by assigning a specific return to each asset in order to determine current wealth value (as a random walk with drift based on the mean of the asset return itself).

Because of the lack of data and the simplification needs of microsimulation, the assumption is made that real wealth is only constituted by the value of the house

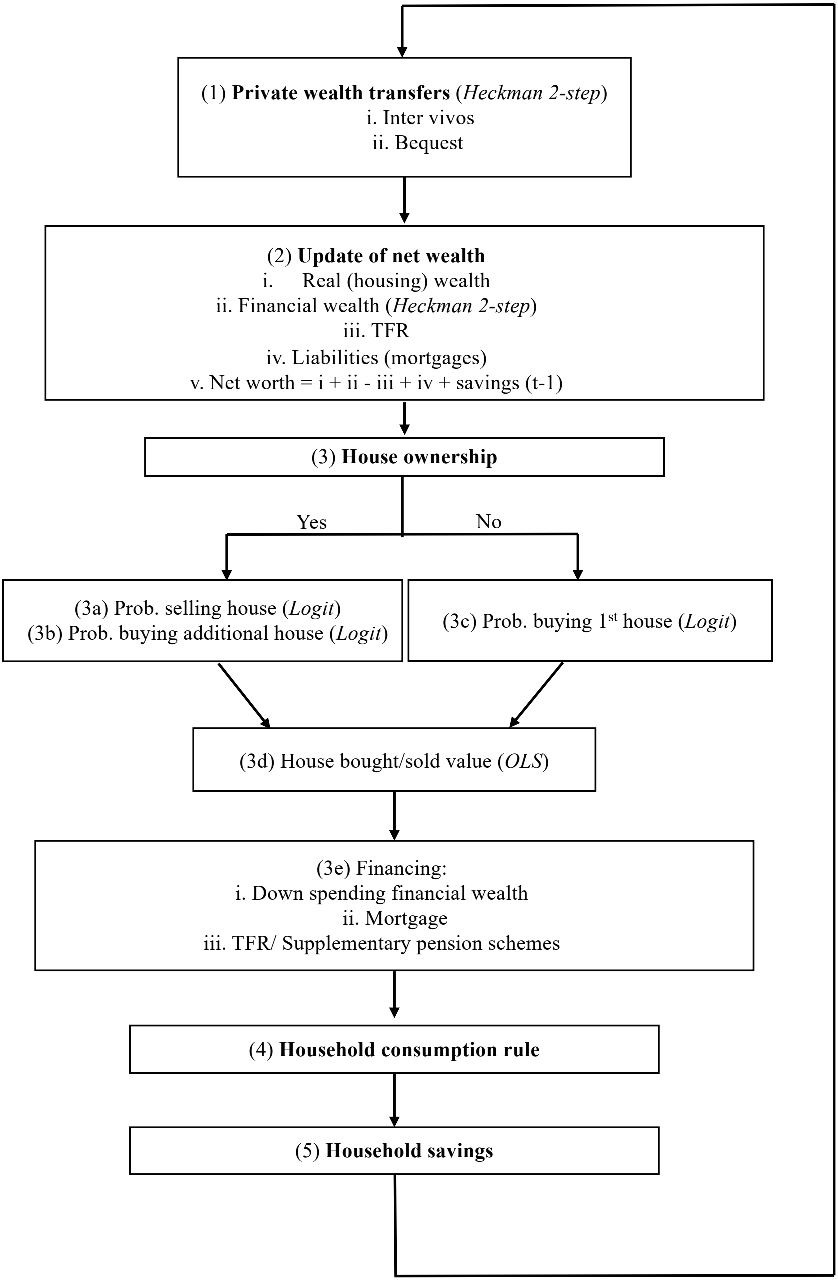
[16] For instance, the so-called 'Citizenship Income'specifically refers to threshold amounts in real (30,000 euros, not accounting for the first house) and financial wealth (6,000 euros to be adjusted to the dimension of the household and the possible presence of disabled individuals).

owned. Returns on real wealth are aligned to external data sources (such as the AWG projections). An open point at this stage of the work regards the opportunity of taking into account the idiosyncratic and systemic components of risk in the rate of return. A possible way to account for the former consists of drawing returns from an asset-return specific distribution (as in Tedeschi *et al.* 2013). This holds for the return on real wealth as well as for the other rates of return that are covered in the module.
Financial wealth is divided into risky and non-risky assets. According to the SHIW classification, risky assets include stocks, mutual funds, private bonds, foreign government bonds, shares of limited liability companies. Non-risky assets are composed of bank and postal deposits CDs, PCTs, BFPs and government securities; amongst non-risky assets, we also add real (tangible) goods other than real estate (firms, valuable objects). We estimate financial wealth allocation between these two components with an econometric model accounting for persistence in the attitude toward risk and for the role of other observables with a selection model à la Heckman.
From 2016, the SHIW questionnaire includes three questions related to the level of financial literacy of the respondent (the head of the household). We are exploring the possibility of integrating this cross-sectional information as an explanatory variable to attempt capturing the relationship between financial knowledge and investment decisions[17]. This variable, which can be influenced by policy, can be calibrated/aligned over the simulation time span in order to affect out-of-sample prediction of investment variables.
We assume the non-risky share of financial wealth accrues a null real return whereas returns on financial risky assets follow external data sources (for instance the S&P 350 index, as in T-DYMM 2.0).
One extra component of financial wealth is the *Trattamento di Fine Rapporto* (TFR)[18] already paid off by the employer to the employee in case of employment termination (dismissal, resignation or retirement). This component is accrued at the individual level. Household savings pertaining to the previous year are another component of financial wealth. Finally, the outstanding debt is obtained by subtracting the capital component of the mortgage instalment paid in the previous period from its lagged value[19]. Mortgages are the only form of borrowing we allow in the module.
Once wealth and debt are computed, net wealth is then given by the sum of real and financial wealth minus the outstanding debt.

---

[17] The economic literature on financial literacy and its relationship with investment decisions has been growing in the last years. See Lusardi and Mitchell (2014) for a review.
[18] The *Trattamento di Fine Rapporto* (end-of-service allowance) is a mandatory severance payment for public and private employees. See Paragraph 7.2.7 for a description of how the TFR is to be simulated within T-DYMM 3.0.
[19] Mortgage instalments are simulated under simplifying assumptions on mortgage duration.

Figure 7.11 Wealth module scheme



The following process (3) consists of the decision whether to acquire or sell a house. It is modelled on a set of discrete choice models (logit) estimated on the pooling of 2004-2016 waves of SHIW. The totals are then aligned to match an official external source (OMI dataset, see Paragraph 1.3.1).
The selection for the 'sale' event occurs through a logit based on a pooling of SHIW cross sections, afterwards the value of home equity sold is heuristically assumed to be the current value of real wealth divided by the number of houses owned. This event implies an increase in the financial wealth of the selling household and it may eventually imply the arising of a new mortgage for the acquiring household (in case the financial wealth does not cover the expense).
The procedure applied to select the households that decide to purchase a new house is similar to the one used to determine the 'sale' event and it is based on a logit estimation. When a household is selected to buy property, the value of the purchased dwelling is estimated using an OLS on a pooling of SHIW cross sections (or, in alternative, using the administrative data from OMI) with the ratio of house value to household net wealth as the dependent variable.
Regarding the financing of house acquisition, the module foresees three possible cases: Purchase covered by financial wealth; in this case, financial wealth decreases by the price of the house bought, real wealth increases by the same amount and there is no creation of additional debt in the system.
If the price of the house exceeds financial wealth, the household's financial availability may be increased by creating new debt in the form of a mortgage.
Purchase covered by (partial) retrieval of wealth accrued in TFR or private pension funds (for individuals who are not retired). If at least one of the household members has an accrued TFR or an individual pension scheme and the purchase concerns the first house, a redemption of it as a set-off of debt contracted is allowed.
The last two steps of the module are strictly related. The application of a consumption rule determines the amount of consumed income at the household level (4). In a first implementation of the model, the consumption rule will be a function of current household disposable incomes and socio-demographic characteristics (without a proper accounting for the role of expected lifetime resources). Within the simulation, consumption will be prevented to exceed the sum of all disposable household income and of 'liquid' non-risky financial wealth net of any mortgage instalment.
In the last step (5), yearly household savings are obtained as the difference between disposable income (net of mortgage instalment) and consumption.

7.2.7 The sub-modules on TFR and private pensions

In the present paragraph, we shall outline the features of two sub-modules to the Wealth Module, the first on the so-called 'end-of-service allowance' (*Trattamento di Fine Rapporto*, TFR) the second on supplementary (private) pensions. Both are relevant additional components of wealth, especially for their role in smoothing consumption over time. One key premise is that these sub-modules, differently from the generality of the Wealth Module, work at the individual and not at the household level.

Trattamento di Fine Rapporto (TFR)

The TFR is a sort of mandatory severance payment for public and private employees. It acts as a deferred share of wage: TFR contributions are withheld and managed directly by the employer (the accrual rate is constant and mandated by law), who has to pay the accumulated amount to the employee in the event of dismissal or retirement. A crucial change in the legislation concerning TFR took place in 2005, when the 'silent consent' formula was introduced (with the new regulations made executive from 2007 onwards): if a worker does not explicitly disagree, his/her TFR flows (not the stock already accrued by firms) are transferred from firms to pension funds. According to the last available data (2018) from the Commission for the Surveillance of Pension Funds (*Commissione di Vigilanza sui fondi Pensione*, COVIP)[20], only about 23% of the overall amount of accrued TFR has been transferred to pension funds.

The convenience of treating TFR and private pensions as adjacent topics is straightforward. Within T-DYMM 3.0, the choice of the employee between the two possible destinations of the TFR will be modelled based on econometric estimates carried out on the new AD-SILC dataset, also making use of the information contained in the SHIW data. The share of workers that transfer their TFR to private pension plans will be aligned to external data from COVIP (a number of different scenarios can be implemented).

The amount of TFR already accrued at the beginning of the simulation will be attributed crossing the available information from AD-SILC 3.0 on the years of contributions and other available variables that describe the evolution of salaries in the sample and over time.

Within the model, the TFR will appear as a deduction on the wage for public and private employees, amounting to 6.91% of the wage gross amount. In case the TFR is kept in the firm, it is re-evaluated every year by 1.5% plus 0.75% of the inflation rate. In case the employee decides to locate his/her TFR to a collective pension fund (see below)[21], it is re-evaluated every year by an interest rate that depends on the type of investment made by the pension fund. As was done for T-DYMM 2.0, we will study and draw the portfolio composition of the pension funds from COVIP data and elaborate projections on rates of return depending on the riskiness of the investments (for instance by discerning between non-risky and risky assets to be consistent with the distinction made in the wealth module).

---

20 For more information about COVIP data, see: http://www.covip.it/.

21 By assumption, we exclude the case of employees transferring TFR to a private pension fund.

Following the pertinent legislation, whenever a work relationship is interrupted, either for resignation or for retirement, the TFR is paid off to the employee as a lump sum. Within T-DYMM, as seen in Paragraph 7.2.6, employees will also have the possibility of accessing a partial anticipation of the TFR for the purchase of the first house[22], for a maximum amount of 70%.

Private pensions

For what concerns private (or supplementary) pension schemes, we shall first offer a brief overview on the topic and on the Italian framework.

Supplementary pillars operate on a voluntary basis; they are fully funded and mainly compute benefits according to Defined Contributions rules[23]. Following the 1993 Reform (Italian Legislative Decree no. 124/1993) and subsequent revisions, the supplementary pillars are organised into three different types of pension institutions: closed (collective occupational) funds (CPFs), open funds (OPFs), and personal pension plans (*Piani Pensionistici Individuali*, PIPs). Even though the legislation has favoured the development of supplementary pensions since the '90s[24], according to the most recent data, the take-up rate in private supplementary schemes is still rather limited. In 2018, the total number of individuals enrolled in supplementary schemes amounted to about 7.9 million, namely 30.2% of all workers.

In what follows, we shall offer a description of how the sub-module on private pensions is to be developed in T-DYMM 3.0. Compared to the simplified version of a sub-module on private pensions included in T-DYMM 2.0, which only made use of the information available in SILC, T-DYMM 3.0 shall also take advantage of the useful information contained in the SHIW survey[25].

This is the information available in SHIW that shall provide useful to our purposes:

- A question on the duration of the contribution in supplementary schemes (the year from which the contribution to the complementary pension has started);
- A question on the percentage of the salary allocated to the supplementary schemes. The same question is also posed with respect to the employer's contribution to the scheme;
- A question regarding the value of the supplementary pension accrued at the end of the reference year;
- A question regarding the investment choices of the pension funds.

---

[22] This anticipation is possible also for extraordinary medical reasons; however, these occasions are not foreseen in the model.

[23] Before the 1993 reform of private pensions (Italian Legislative Decree no. 124/1993), only pension funds sponsored by banks and insurance companies in favor of their employees, the so called 'pre-existing funds', were operating. Only pre-existing funds may provide benefits calculated according to Defined Benefit formulas.

[24] All contributions to private pension funds are deductible up to 5,164.57 euros per year.

[25] See Paragraph 1.3 for a brief description of SHIW and of its linkage with AD-SILC 3.0.

In the model, we will simplify the complex system of private pension schemes by featuring a limited set of alternatives to the public scheme that will constitute our supplementary pension scheme, encompassing all the three aforementioned institutions available in the Italian system.
The choice to participate in a supplementary pension scheme will be modelled based on econometric estimates carried out on the SHIW dataset. SHIW will also guide us in estimating the amount devoted to pension funds for a given year (including whether it came from TFR or not) as a share of the yearly income and in assigning the amounts of contributions accrued before the start of the simulation for the sample in the base year. Both probabilistic estimations will constitute an improvement of the methodology of T-DYMM 2.0, where both dimensions where deterministically imputed.
Similarly to what outlined for collective pension funds (see above), the rate of return of supplementary pension funds will be attributed according to a set of assumptions on future portfolio compositions based on historical COVIP data and on the development of securities by risk-class.

References

Aaberge R., Colombino U. (2014), Labour Supply Models, in O'Donoghue C. (ed.), *Handbook of Microsimulation Modelling*, Bingley, Emerald, pp.167-221

Aaberge R., Colombino U. (2018), Structural Labour Supply Models and Microsimulation, *International Journal of Microsimulation*, 11, n.1, pp.162-197

Albarea A., Bernasconi M., Di Novi C., Marenzi A., Rizzi D., Zantomio F. (2015), Accounting for tax evasion profiles and tax expenditures in microsimulation modelling. The BETAMOD model for personal income taxes in Italy, *International Journal of Microsimulation*, 8, n.3, pp.99-136

Ando A., Nicoletti-Altimari S. (2004), *A micro simulation model of demographic development and households' economic behavior in Italy*, Temi di discussione n.533, Roma, Banca d'Italia

AISP (Italian Association for Population Studies) (2017), *Rapporto sulla Popolazione – Le molte facce della presenza straniera in Italia*, Bologna, Il Mulino

Azzolini D., Bazzoli M., De Poli S., Fiorio C., Poy S. (2017), Developing and Validating Regional Microsimulation Models. TREMOD: The Tax-Benefit Model of the Italian Province of Trento, *Economia Pubblica*, n.1, pp.5-33

Bae J.W., Paik E., Kim K., Singh K., Sajjad M. (2016), Combining Microsimulation and Agent-based Model for Micro-level Population Dynamics, *Procedia Computer Science*, 80, special issue, pp.507-517

Baekgaard H. (2002), *Micro-macro linkage and the alignment of transition processes: some issues, techniques and examples*, National Centre for Social and Economic Modelling Technical Paper n.25, Canberra, NATSEM

Baldini M. (2001), Inequality and redistribution over the life-cycle in Italy. An analysis with a dynamic cohort microsimulation model, *Brazilian Electronic Journal of Economics*, 4, n.2, pp.1-5

Baldini M., Toso S. (2009), *Diseguaglianza, povertà e politiche pubbliche*, Bologna, Il Mulino

Baldini M., Giarda E., Olivieri A. (2015a), *A tax-benefit microsimulation model for italy: a partial evaluation of fiscal consolidation in the period 2011-2014*, Prometeia Nota di lavoro n.1, Bologna, Prometia

Baldini M., Pacifico D., Termini F. (2015b) *Imputation of missing expenditure information in standard household income surveys*, DEMB Working Paper Series n.49, Università di Modena e Reggio Emilia

Betti G., Donatiello G., Verma V. (2011), The Siena microsimulation model (SM2) for net-gross conversion of EU-SILC income variables, *International Journal of Microsimulation*, 4, n.1, pp.35-53

Bianchi C., Romanelli M., Vagliasindi P.A. (2005), Validating a Dynamic Microsimulation Model of the Italian Households, in Leskow J., Punzo L.F.,

Boscolo S. (2019), Quantifying the Redistributive Effect of the Erosion of the Italian Personal Income Tax Base: A Microsimulation Exercise, *Economia Pubblica*, n.2, pp.39-80

Bourguignon F., Spadaro A. (2006), Microsimulation as a tool for evaluating redistribution policies, *The Journal of Economic Inequality*, 4, n.1, pp.77-106

Buddelmeyer H., Hérault N., Kalb G., van de Zijll de Jong M. (2012), Linking a Microsimulation Model to a Dynamic CGE Model: Climate Change Mitigation Policies and Income Distribution in Australia, *International Journal of Microsimulation*, 5, n.2, pp.40-58

Burtless G., Hauman J.A. (1978), The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment, *Journal of Political Economy*, 86, n.6, pp.1103-1130

Buslei H., Bach S., Simmler M. (2014), Firm Level Models, in O'Donoghue C. (ed.), *Handbook of Microsimulation Modelling*, Bingley, Emerald Publishing, pp.479-503

Caretta A., Flisi S., Frale C., Raitano M., Tedeschi S. (2013), *T-DYMM: the Treasury Dynamic Microsimulation Model of the Italian Pension System*, MEF Working Paper n.11, Roma, Ministero dell'Economia e delle Finanze

Ceriani L., Fiorio C.V., Gigliarano C. (2013), The importance of choosing the dataset for tax-benefit analysis, *International Journal of Microsimulation*, 6, n.1, pp.86-121

Chénard D. (2000), Individual alignment and group processing: an application to migration processes in DYNACAN, in Mitton L., Sutherland H., Weeks M. (eds.), *Microsimulation Modelling for Policy Analysis*, Cambridge, Cambridge University Press.

Colombo G. (2010), Linking CGE and Microsimulation Models. A Comparison of Different Approaches, *International Journal of Microsimulation*, 3, n.1, pp.72-91

Colombino U. (2015), Five Crossroads on the Way to Basic Income. An Italian Tour, *Italian Economic Journal*, 1, n.3, pp.353-389

Coromaldi M., Guerrera D. (2009), Modello di Microsimulazione EconLav: la costruzione del data-set di input, MEF Working Papers n.4, Roma, Ministero dell'Economia e delle Finanze

Cozzolino M., Di Marco M. (2015), Micromodelling Italian Taxes and Social Policies. *Rivista di Statistica Ufficiale*, 2, pp.17-26

Creedy J., Tuckwell I. (2004), Reweighting Household Surveys for Tax Microsimulation Modelling. An Application to the New Zealand Household Economic Survey, *Australian Journal of Labour Economics*, 7, n.1, pp.71-88

Creedy J., Kalb G. (2005), Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation, *Journal of Economic Surveys*, 19, n.5, pp.697-734

Curci N., Savegnago M., Cioffi M. (2017), *BIMic. The Bank of Italy microsimulation model for the Italian tax and benefit system*, Questioni di economia e finanza n.394, Roma, Banca d'Italia

Dekkers G. (1999), *The future development of living standards of the retirees in Belgium: an application of the static microsimulation model station*, MPRA Paper n.36005, University Library of Munich

Dekkers G. (2015a), *On the modelling of immigration and emigration using LIAM2*, NOTE-LIAM2-11155, Federaal Planbureau <https://bit.ly/38lUYXr>

Dekkers G. (2015b) The simulation properties of microsimulation models with static and dynamic ageing – a brief guide into choosing one type of model over the other, *International Journal of Microsimulation*, 8, n.1, pp.97-109

Dekkers G., Buslei H., Cozzolino M., Desmet R., Geyer J., Hofmann D., Raitano M., Steiner V., Tanda P., Tedeschi S., Verschueren F. (2009), *What Are the Consequences of AWG Projections for the Adequacy of Social Security Pensions?*, ENEPRI Research Report n.65, AIM Work Package 4

Dekkers G., Cumpston R. (2012), On weights in dynamic-ageing microsimulation models, *International Journal of Microsimulation,* 5, n.2, pp.59-65

Di Caro P. (2018), *Redistribution in real-world PIT. Evidence from Italian records*, DF Working Paper n.2, Roma, MEF

Di Nicola F., Mongelli G., Pellegrino S. (2015), The static microsimulation model of the Italian Department of Finance. Structure and first results regarding income and housing taxation, *Economia pubblica*, n.2, pp.125-157

Duleep H.O., Dowhan D.J. (2008a), Research on Immigrant Earnings, *Social Security Bulletin*, 68, n.1, pp.31-50

Duleep H.O., Dowhan D.J. (2008b), Incorporating Immigrant Flows into Microsimulation Models, *Social Security Bulletin*, 68, n.1, pp.67-76

D'Amuri F., Fiorio C.V. (2006), Grossing-up and Validation Issues in an Italian Tax-Benefit Microsimulation Model, Econpubblica Working Paper n.117

Ericson P., Hussenius J. (1999), *Distributional Effects of Public Student Grants in Sweden – a Presentation and an Application of the Dynamic Microsimulation Model SESIM*, APPAM Seminar "Public Policy Analysis and Management: Global and Comparative Perspectives", Washington DC, November 4-6

European Commission (2017), *The 2018 Ageing Report. Underlying Assumptions and Projection Methodologies*, Institutional Paper n.065, Luxembourg, Publications Office of the European Union

Gastaldi F., Pazienza M.G., Pollastri C. (2017), *I Modelli di Microsimulazione dell'UPB: Famiglie e Imprese,* Mimeo online printing

Hausman J.A. (1979), The econometrics of labor supply on convex budget sets, *Economic Letters*, 3, n.2, pp.171-174

Heckman J.J. (1979), Sample selection bias as a specification error', *Econometrica*, 47, n.1, pp.153-161

Hufkens T., Goedemé T., Gasior K., Leventi C., Manios K., Rastrigina O., Recchia P., Sutherland H., Van Mechelen N., Verbist G. (2019), *The Hypothetical Household Tool (HhoT) in EUROMOD: a new instrument for comparative research on tax-benefit policies in Europe*, JRC Working Papers on Taxation and Structural Reforms n.05, European Commission, Seville, Joint Research Centre

IESS (2016), *Final Report*, see <https://bit.ly/2BwUMbU>

Immervoll H., O'Donoghue C. (2001), *Imputation of gross amounts from net incomes in household surveys: an application using EUROMOD*, EUROMOD Working Paper n.EM1, Essex, EUROMOD

Klevmarken A., Olovsson P. (1996), Direct and behavioural effects of income tax changes: simulations with the swedish model MICROHUS, in Harding A. (ed.), *Microsimulation and Public Policy*, Amsterdam, Elsevier Science Publishers, pp.203-229

Leombruni R., Richiardi M. (2006), LABORsim. An Agent-Based Microsimulation of Labour Supply–An Application to Italy, *Computational Economics*, 27, pp.63-88

Li J., O'Donoughe C. (2013), A survey of dynamic microsimulation models: uses, model structure and methodology, *International Journal of Microsimulation*, 6, n.2 pp.3-55

Li J., O'Donoughe C. (2014a), Evaluating Binary Alignment Methods in Microsimulation Models, *Journal of Artificial Societies and Social Simulation*, 17, n.1, pp.1-15

Li J., O'Donoughe C., Dekkers G. (2014b), Dynamic models, in O'Donoghue C. (ed.), *Handbook of Microsimulation Modelling*, Bingley, Emerald Group Publishing, pp.305-343

Li J., O'Donoughe C., Loughrey J., Harding A. (2014c), Static Models, in O'Donoghue C. (ed.), *Handbook of Microsimulation Modelling*, Bingley, Emerald Group Publishing, pp.47-75

Lusardi A., Mitchell O.S. (2014), The economic importance of financial literacy: theory and evidence, *Journal of Economic Literature*, 52, n.1, pp.5-44

Maitino M.L., Sciclone N. (2009), *Assessing the Implications of Population Ageing on Tuscan Wellbeing. A Microsimulation Approach*, unpublished available https://bit.ly/3eU6w6B
Maitino M.L., Ravagli L., Sciclone N. (2017), MicroReg: a Traditional Tax-Benefit Microsimulation Model Extended to Indirect Taxes and In-Kind Transfers, *International Journal of Microsimulation*, 10, n.1, pp.5-38
Martini A., Trivellato U. (1997), The Role of Survey Data in Microsimulation Models for Social Policy Analysis, *Labour*, 11, n.1, pp.83-112
Mazzaferro C., Morciano M. (2012), *CAPP_DYN. A Dynamic Microsimulation Model for the Italian Social Security System*, CAPPaper n.48, Università di Modena e Reggio Emilia
MEF-FGB (2012), Innovative datasets and models for improving welfare policies, Final Report of the INDIW project, Roma, Ministero dell'Economia e delle Finanze - Dipartimento del Tesoro, Fondazione Giacomo Brodolini
Menonna A., Blangiardo G.C. (2018), Flussi migratori dall'Africa: scenari per il futuro, neodemos.info <https://bit.ly/3ipnagQ>
Merz J. (1991), Microsimulation. A survey of principles, developments and applications, *International Journal of Forecasting*, 7, n.1, pp.77-104
Michelangeli V., Pietrunti M. (2014), A Microsimulation Model to evaluate Italian Households' Financial Vulnerability, *International Journal of Microsimulation*, 7, n.3, pp.53-79
O'Donoughe C. (2001), Dynamic microsimulation. A methodological survey, *Brazilian Eletronic Journal of Economics*, 4, n.2, pp.1-77
O'Donoughe C., Lennon J., Hynes S. (2009), The Life-Cycle Income Analysis Model (LIAM). A Study of a Flexible Dynamic Microsimulation Modelling Computing Framework, *International Journal of Microsimulation*, 2, n.1, pp.16-31
O'Donoughe C., Redway H., Lennon J. (2010), Simulating Migration in the PENSIM2 Dynamic Microsimulation Model, *International Journal of Microsimualtion*, 3, n.2, pp.65-79
O'Donoughe C., Morrissey K., Lennon J. (2014), Spatial Microsimulation Modelling: a Review of Applications and Methodological Choices, *International Journal of Microsimulation,* 7, n.1, pp.26-75
Orcutt G.H. (1957), A new type of socio-economic system, *Review of Economics and Statistics*, 39, n.2, pp.116-123
Orcutt G.H. (1967), Microeconomic Analysis for Prediction of National Accounts, in Wold H., Orcutt G.H., Robinson E.A., Suits D., de Wolff P. (eds.), *Forecasting on a Scientific Basis – Proceedings of an International Summer Institute*, Lisbon, Gulbenkian Foundation, pp.67-127
Pacifico D. (2009), A behavioural microsimulation model with discrete laour supply for Italian couples, CAPPaper n.65, Università di Modena e Reggio Emilia
Pacifico D. (2014), *Reweight: a Stata Module to reweight Survey Data to External Totals*, Working Paper n.5, Roma, MEF

Peichl A. (2016), Linking Microsimulation and CGE Models, *Internationl Journal of Microsimulation*, 9, n.1, pp.167-174

Pellegrino S., Piacenza M., Turati G. (2011), Developing a Static Microsimulation Model for the Analysis of Housing Taxation in Italy, *International Journal of Microsimulation*, 4, n.2, pp.73-85

Rutter C.M., Zaslavsky A.M., Feuer E.J. (2010) Dynamic Microsimulation Models for Health Outcomes. A Review, *Medical Decision Making*, 31, n.1, pp.10-18

Scott A. (2001), *A computing strategy for SAGE: 1. Model options and constraints*, Technical Note n.2, London, ESRC-Sage Research Group

Stock J.H., Wise D.A. (1990), Pensions, the option value of work, and retirement, *Econometrica*, 58, n.5, pp.1151-1180

Sutherland H., Figari F. (2013), EUROMOD: the European Union tax-benefit microsimulation model, *International Journal of Microsimulation*, 6, n.1, pp.4-26

Sutherland H. (2018), Quality Assessment of Microsimulation Models. The Case of EUROMOD, *International Journal of Microsimulation*, 11, n.1, pp.198-223

Tanton R. (2018), Spatial Microsimulation. Developments and Potential Future Directions, *International Journal of Microsimulation*, 11, n.1, pp.143-161

Tedeschi S., Pisano E., Mazzaferro C., Morciano M. (2013), Modelling Private Wealth Accumulation and Spend-down in the Italian Microsimulation Model CAPP_DYN: A Life-Cycle Approach, *International Journal of Microsimulation*, 6, n.2, pp.76-122

Van Sonsbeek J. (2011), Micro simulations on the effects of ageing-related policy measures. The Social affairs Department of the Netherlands Ageing and Pensions Model, *International Journal of Microsimulation*, 4, n.1, pp.72-99

Willekens F. (2006), *Description of the micro-simulation model (Continuous-time micro-simulation)*, Deliverable D8 (first part), The Hague, Netherlands Interdisciplinary Demographic Institute

Concluding remarks

The present report has provided a comprehensive overview of the activities undertaken to upgrade the ADSILC database and a thorough analysis of the labour market features most prominent for the MOSPI project. We have also laid down the main advances to be achieved in terms of the evolution of the T-DYMM model.

In the next stages, we shall turn our focus on implementing the planned revisions and enhancements to the model, which will enable us to deliver the policy insights at the heart of this project. In particular, in line with the presented proposal, we shall: a) upgrade the capability of the model to simulate social protection subsides by updating its current blocks with new existing institutes and, at the same time, enabling it to encompass additional future policy applications; b) include a sub-module on working pensioners and, more in general, enable the model to encompass any flexible working arrangement which could be envisaged for senior workers; c) introduce a model block for real estate ownership and financial wealth that takes into account supplementary pensions schemes and portfolio decisions; d) expand the taxation module, also thanks to the newly available administrative fiscal data; e) introduce a migration module, taking into account the impacts on population projections and on income dispersion. Overall, the new model features will allow us to undertake a more informed evaluation of the adequacy of the welfare system in a forward-looking perspective, with a special focus on disadvantaged individuals and non-standard workers and in the light of a wide range of demographic and socio-economic scenarios. Possible policy actions, as promised, will be evaluated in depth.

Over the next months we shall proceed with the final activities related to data acquisition (e.g. introducing in AD-SILC the fiscal administrative information) and then move to the estimation of the model equations; the latter activity will be informed by the know-how gained while analysing the new dataset and illustrated in the current report. While working on the new tasks, an essential contribution and stimulus will be provided by the monitoring and evaluation process undertaken by Fondazione Giacomo Brodolini. A number of foreseen activities will be specifically helpful in focusing the efforts undertaken by the modelling team; we deem the forthcoming feedback in terms of transferability of the MOSPI project (with respect to other European MS) particularly relevant. The enhancement of the model will strongly benefit from the peer review workshop – to be planned – in which EU experts will assess the state of progress of the project and allow us to fine-tune our line of action. On top of officially planned activities, we shall keep exchanging views with national and international experts. Data providers, policy institutions and academic experts have been instrumental in the development of T-DYMM in the past; we look forward to strengthening our collaboration to ensure its wider success in the future.

Finished printing in August 2020
from Rubbettino print
on behalf of Rubbettino Editore Srl
88049 Soveria Mannelli (Catanzaro)
www.rubbettinoprint.it