Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA
IN SCIENZE STATISTICHE

Ciclo 34

Settore Concorsuale: 13/A3 - SCIENZA DELLE FINANZE
Settore Scientifico Disciplinare: SECS-P/03 - SCIENZA DELLE FINANZE

INEQUALITY OF INCOME AND POVERTY PERSISTENCE IN THE MUNICIPALITY OF BOLOGNA: A CROSS-SECTIONAL AND LONGITUDINAL ANALYSIS THROUGH ADMINISTRATIVE DATA

Presentata da: Marika Bazzocchi

Coordinatore Dottorato
Monica Chiogna

Supervisore
Carlo Mazzaferro

Esame finale anno 2022

Abstract

This work is part of a project promoted by Emilia-Romagna that aims at encouraging research activities in order to support the innovation strategies of the regional economic system through the exploitation of new data sources. To gain this scope, a database containing administrative data is provided by the Municipality of Bologna. This is achieved by linking data from the Register Office of the Municipality and fiscal data coming from the tax returns submitted to the Revenue Agency and released by the Ministry of Economy and Finance for the period 2002-2017. The main purpose of the project is the analysis of the medium term financial and distributional trends of income of the citizens residing in the Municipality of Bologna. Exploiting this innovative source of data allow us to analyse the dynamics of income at municipal level, overcoming the lack of information in official survey-based statistic. We investigate these trends by building inequality indicators and by examining the persistence of in-work poverty. Our results represent an important informative element to improve the effectiveness and equity of welfare policies at the local level, and to guide the distribution of economic and social support and urban redevelopment interventions in different areas of the Municipality.

Acknowledgements

Contents

Chapter 1

Introduction

In the last decades, income inequality has been rising in many countries all over the world, becoming one of the most discussed issue in the economic policy agenda. Much of the focus of this debate has dealt with the highest earners, but there is also growing concern about the economic situation of the low-income earners. Evidence suggests that the wide gap between rich and poor contribute to slow the economic growth and also affects other social issues, i.e. lower the social mobility, increase the rate of crime and bad health (OECD, 2015). Government action can reduce disparities and encourage a more inclusive growth. In seeking to promote this multifaceted task, policies approaches must encompass a range of different areas, including education, labour market and taxation. Drawing on the research about income inequalities and its trends over time, governments can shape policies to ensure a more even distribution of income and opportunities.

Since the publication of the pioneering work of Piketty (2003) and Piketty and Saez (2003), researchers started to use tax data for the analysis of income inequality. As a consequence, a considerable effort by the National Statistical Offices has allowed to make these

archives available and spread the systematic use of administrative sources. The development of information technology in Public Administrations and the increased awareness among policy makers on the relevance of statistics to governance have been two factors that have favoured this process (Calzaroni, 2011).

Along this line of research, this work exploits newly available administrative data on individual income tax returns linked to data coming from the Registry Office of the Municipality of Bologna for the period 2002-2017.

Although the use of this new sources is often followed by serious challenges, the exploitation of administrative data is now widespread among researchers and statisticians, replacing or integrating census and sample surveys. Indeed, these two types of information are primarily designed for different purposes: administrative data are collected by government institutions or agencies for tax, benefit or public administration purposes, whereas survey data are collected specifically for statistical and research purposes. A complete review of advantages and potential drawbacks of administrative data is therefore essential. Following the works of Penneck (2007) and Johnson and Moore (2005), we highlight in detail both the strengths and weaknesses of administrative records compared to those of surveys. Advantages of administrative data include a complete coverage of a population, low data collection costs, reduced respondent burdens, and better data quality. On the other hand, the stability of a program over time, privacy concerns about non-administrative use of data and conceptual issues relative to the population and items collected are some of the drawbacks.

The main feature of this work resides in the use of this new source

of data, overcoming the lack of statistical information at municipal level and at finer level of territorial reference. Therefore, it supplies estimates that are difficult to recover with other available instruments. Indeed, sample surveys typically used for the analysis of income distribution do not allow for sufficiently fine partitions of the territory. In this work, on the contrary, thanks to our innovative dataset, we can analyse income dynamics and build inequality indicators for the Municipality of Bologna and its neighbourhoods. Moreover, this work examines the distribution of poverty among working individuals (Ahrendt et al., 2017; Raitano et al., 2019). Indeed, research literature has shown that being in the labour market does not exclude one from the risk of falling into poverty. Moved by this interest and the availability of longitudinal income data, we investigate the dynamics of in-work poverty and construct poverty measures. In this regard, following the approach based on the seminal works of Allison (1982) and Bane and Ellwood (1986), and also some ensuing works (Stevens, 1999, Devicienti, 2002; Biewen, 2006), we address poverty for the period 2007-2017 by introducing two new measures that account for the duration of poverty spells, analysing recurrence and transitions into and out of poverty, as well as estimating exit and re-entry risk rates in the working population aged 16-64. The contribution of this research project is threefold. First, analysing income distribution at a local level and seeing how it has changed over the years has an important value, since the national average data can hide very different situations at the territorial level. This is particularly true for Italy where, for many economic indicators, the territorial gaps are much wider than those recorded in comparison with other countries (Ciani and Torrini, 2019).

The second contribution of this work is the possibility to investigate the performance of the top income shares. Until very recently (Piketty and Saez, 2006; Alvaredo and Pisano, 2010; Alvaredo et al., 2013), nothing was known about this aspect in the distribution of income, primarily due to a lack of data. The examination of the top income shares is important because it enriches the analysis of inequality, allowing us to observe the trends relating to the high tail of the distribution with a much greater precision than that ensured by the sample surveys, which, due to their limited number of observations, run the risk of not adequately representing high-income individuals in the sample.

Finally, the study of the longitudinal dimension of the poverty problem at a local level represents an important novelty with respect to the existing research, which has mainly focused on the spread and intensity of poverty at a given point in time. Indeed, the analysis of the dynamics of in-work poverty at the local level is of great interest in order to be able to characterize the complete low income pattern of individuals along time. It also allows to distinguish the characteristics of workers that suffered long period of poverty in contrast with those that experience poverty for a relatively short time, which is of particular importance for the decision maker in order to address adequate policy measures. To this regard, another novelty of this study is the adoption of a local threshold, based on income declared by individuals in the tax report and referred to the level of well-being of the population residing within the municipal boundaries. Recent literature (Biggeri and Pratesi, 2017; Bishop et al., 2017) has in fact highlighted the limits of adopting a single national threshold to estimate poverty at the sub-national level.

The thesis is organized as follows. In Chapter 2 administrative data are introduced, their strengths together with the potential drawbacks of using them for statistical purposes, with respect to traditional survey data, are proposed and their empirical implications explored. We also present the administrative data set used for the ensuing analyses together with a comparison with data from the Italian Statistics on Income and Living Conditions survey. Chapter 3 presents a detailed description of the income trends in the Municipality of Bologna together with a list of indices of inequality, their characteristics, and their variation over the 2002-2017 period. It also analyses tax levy trends and focuses on the role played by the personal income tax in the redistribution of income at a local level. Chapter 4 gives a picture of both the cross-sectional and longitudinal poverty for the working population of the Municipality of Bologna over the 2007-2017 period using a local poverty threshold. Eventually, we report our conclusions and directions for further research in Chapter 5.

Chapter 2

Using administrative data for statistical purposes

2.1 Introduction

The demand for statistics on all aspects of our lives, our society, and our economy is constantly growing. At the same time, the increased ability to handle and manipulate large sets of data together helps in answering to the desire of understanding social and economic systems better, and make predictions helpful to improve the welfare in the society we live in.

In this context, statistical agencies are exploring new ways to produce official statistics in near real time by exploiting new kind of data sources: the big data. These data have some advantages over the traditional one derived from probability samples: they have a much larger sample size, they are less costly, and they produce statistics at the same time or shortly after the required information is extracted (Beaumont, 2020). Big data include administrative data, transaction data, social media data, internet of things and scrape data from websites, sensor data and satellite images (Rao, 2020).

In this scenario of increased availability of large dataset, the ability

7

to manipulate them has become possible thanks to the rapid development of information technology in the Public Administrations, which has made available a large amount of structured information both on institutions and individuals.

Therefore, while until the previous decade data for official statistics were produced almost exclusively using sample surveys and census, nowadays statistical agencies are exploring new ways to integrate survey and administrative data to create information systems capable of supporting a variety of research and policy purposes. The combination of these two sources is a promising and innovative strategy which affects the quality and quantity of research and increases the potential of the data.

However, this usage often comes with serious challenges, given that these two types of data are primarily designed for different purposes. Administrative data are defined as datasets collected by government institutions or agencies for tax, benefit, or public administration purposes. On the other hand, survey data are collected specifically for statistical and research purposes (Johnson and Moore, 2005).

This Chapter introduce the administrative data and examine their strengths with respect to traditional survey data, together with the potential drawbacks of using them for statistical purposes. A complete coverage of a population, low data collection costs, reduced respondent burdens, and better data quality are listed between the advantages. On the other hand, the stability of a program over time, privacy concerns about non-administrative use of data, conceptual issues relative to the population and items collected, and costs of transforming the data into a form useful for research purposes are some of the drawbacks.

After going through the dissimilarities associated with the use of these new data sources with respect to the traditional ones, in the remainder of the Chapter the dataset used for this research project is presented. This data source links data on personal income tax returns submitted to the Revenue Agency with those coming from the Registry Office of the Municipality of Bologna for the period 2002-2017.

The underlying idea is to use the dataset to display the informative potentials of a tax administrative source and the possibility to use it for exploring the medium-term trend in the distribution of personal income in a local area. Besides, the main aspects of the administrative source used will be presented, both the characteristics of fiscal data and those of the registry office data.

The Chapter is organised as follows. In Section 2 and 3 the big data and specifically administrative data are respectively presented. In Section 4, a detailed description of the differences between administrative and survey data are proposed and their empirical implications explored. Among other things, the matter of when it is preferable to use administrative data and when survey data is addressed. Section 5 presents the administrative dataset used for the ensuing analyses. The complex features of the total income for the personal income tax are exposed, together with clarifications related to the use of this variable for statistical purposes. A description of the demographic variables available in the dataset presented also follows. Section 7 concludes with a relative straightforward description of the characteristics which assure the quality of the dataset, along with a comparison between our dataset and data from the Italian Survey on Household Income and Living Conditions (abbreviated

IT-SILC).

2.2   Big Data

Sample survey has long been the leading tool used to obtain reliable results of finite population parameters. If the dataset comes from a probability sample, parameter estimation is straightforward and an extensive literature is available. In 1934, Neyman published an article in the *Journal of the Royal Statistical Society* which laid the foundations for this topic (Neyman, 1934). He introduced the idea of efficiency of design-unbiased estimator and optimal sample size allocation in his theory of stratified simple random sampling. His approach was almost universally accepted by statisticians, and it was a source of inspiration for various important further extensions. The early landmark contributions post-Neyman theoretical developments under the design-based approach, carried out by Mahalanobis and survey statisticians at the U. S. Census Bureau under the leadership of Morris Hansen, were mostly motivated by practical and efficiency considerations. The first one extended the theoretical setup to sub-sampling of sampled clusters. The second worked on the development of the basic theory of stratified two-stage cluster sampling with one cluster within each stratum, drawn with probability proportional to size. They also work on the introduction of rotation sampling with partial replacement to handle response burden in surveys repeated over time. Rao (2020) examined several developments of sample survey theory over the 20th century, such as the development of methods for combining two or more probability samples to increase the efficiency of estimators for a given cost, together with the more recent techniques of data collection to control

costs, maintain good coverage and response rates, and adjust for non-response bias and measurement errors[1].

Despite the additional efforts made in setting up sample surveys, some of them, conducted by national statistical agencies, still have very low response rates, high costs of data collection, and burden on respondents are increasing. So, it becomes dangerous to trust uniquely on data collection and estimation methods to correct potential non-response bias (Rao, 2020).

At the same time, due to technological innovations, large amount of less expensive data, the big data, are becoming increasingly accessible. Big data include administrative data, transaction data, social media data, internet of things and scrape data from websites, sensor data and satellite images. These data have the advantages over those derived from probability samples of having a much larger sample size and being less costly. They also provide estimates in near real time, i.e. they allows to produce statistics at the same time or shortly after the information is extracted (Rao, 2020). For this reason, statistical agencies publishing official statistical information are exploring new ways to integrate data from different sources by exploiting these new kind of data source.

However, even if some researchers have come to believe that probability survey could gradually disappears (see Couper, 2000; Couper, 2013; Miller, 2017), naïve use of data from non-probability sources can also lead to serious sample selection bias. Meng (2018) points out that, without adjustment, sample selection bias becomes dominant as the non-probability sample size increases, which significantly reduces the effective sample size, leading to the "big data para-

---

[1]For further details, see Bethlem, 2009, Rao and Fuller, 2017, and Kalton, 2019,

dox: the bigger the data, the surer we fool ourselves" (Meng, 2018). In his article, which focuses on population inference from the big data, the author proved with practical examples that, without taking data quality into account, population inference with big data gained lower quality than the ones obtained from a smaller sample size survey. Therefore, large non-probability samples alone cannot ensure the production of estimates with an acceptable level of quality. As noted by Elliott and Valliant (2017), the pre-election poll conducted by the *Literary Digest* magazine to forecast the outcome of the 1936 US presidential election is one of the early example of failure of big non-random sample. Despite a huge sample size of 2.3 million responses, the poll incorrectly predicted a sizable victory of Alf Landon over Franklin Roosevelt. This was due to the fact that this pool under-represented Roosevelt's voters, since it mainly consisted of car and telephone owners and magazine's own members. Couper (2000) and Elliott and Valliant (2017) have illustrated other more recent examples of polls that failed to correctly predict election outcomes, which underlined the potential problems that affect non-probability samples.

On the other hand, selecting a probability sample does not necessarily guarantee a good basis for inference to a population. In many types of survey, response rates have declined dramatically since there have been increasing difficulties to contact potential respondents and to persuade them to participate. (Elliott and Valliant, 2017). According to *Pew Research*, the response rate for a typical telephone survey fell from 36 per cent in 1997 to 9 per cent in 2012 (Kohut et al., 2012). Low response rates have risen doubts of how well a sample initially selected randomly can be called a probability sam-

ple from the desired population. These challenges have led many to wonder if the surveys are still the appropriate method to provide fair and unbiased estimates without using more expensive methods of data collection which usually have higher response rate.

In a scenario of profound social and economic changes, the information required by official statistics is also transforming (Calzaroni, 2011). Besides, the lower availability of resources for official statistics has led national statistical agencies to a progressive and systematic use of administrative sources. This implied an increasing work to determine terms and conditions of these archives, which has been possible on account of the rapid development of information technology in Public Administrations. Indeed, a large amount of well-organised information on companies, institutions and individuals have been made available. Calzaroni (2011) underlined that the increasing availability of these sources has led to the necessity of defining the guidelines for the collection and analysis, together with the formalization of the statistical methodologies to treat administrative data in order to obtain more reliable estimates.

In the face of the numerous advantages, the use of administrative data also has costs. To the costs deriving from the development of specific methodologies and organizational changes necessary for the management, it must be added those imposed by the need for coordination between the national statistical institutes and the company holding the national and territorial administrative sources of data.

In spite of that, traditional sample surveys have become even more important than in the past when they are integrated to big data source. Nowadays, in fact, it is deemed to be better integrating the sources of information rather than putting them in competition,

when the purpose is to obtain the best results in terms of quality and quantity of statistical information.

2.3 Administrative Data

Administrative data are defined as datasets collected by government institutions or agencies for tax, benefit, or public administration purposes (United Nations Economic Commission for Europe, 2011).

Types of administrative data

These data have several forms and compositions. First, they may be distinguished those administered nationally from those administered by provinces and municipalities. Besides, they change in terms of their purpose, which constitute a leading crucial factor of their coverage, quality and statistical usefulness.

One of the most extended definition of administrative source was issued by Brackstone (1987), who suggested to distinguish the administrative records in six categories.

1. *Records from the flow of goods and people across borders.* These include records of imports, exports, immigration, and emigration, which depends on laws usually established by national administrations;

2. *Records from the registration of legal events.* Examples are data from births, deaths, marriages, divorces, businesses actions, and other forms which requires to be formal registered by institutions;

3. *Records needed to administer benefit or obligations.* Taxation,

unemployment and health insurance, pensions, and family allowances belong to this category;

4. *Records needed to administer public institutions.* Examples of this data are records related to schools, universities, health institutions, courts or prisons;

5. *Records from the regulation of industry.* These include records from transportation, banking, broadcasting, telecommunications, and records from the management of supply and price of some commodities.

6. *Records from the provision of utilities.* Electricity, phone, and water services are some examples of data included in this category. Generally, these data are administrated at the provincial or municipal levels.

Brackstone (1987) suggested that administrative records also change in terms of the processes by which they are gathered. Depending on whether the data is administered nationally or locally, formats and hardware are different. Administrative records also differ in terms of accessibility. Some datasets are affected by laws and restrictions that govern their access and limit their use from a statistical point of view. Similarly, in some case the quantity of variables available to researchers is limited, making difficult to assess the quality and, therefore, the use of the administrative data for statistical purposes.

Use of administrative data

As a research tool, administrative records have many applications, including the direct tabulation and the indirect estimation of models or other statistics, as well as the construction of survey frames and

the validation of survey results. Following Brackstone (1987), we sum them up as follows.

1. *Direct production of data from an administrative source.*
   In this case, the reference population should match the target population, and the variables observed by the administrative process should correspond to the statistical variables. When these conditions do not hold, it is necessary to define a process to adjust the original information into data statistics.
   Examples are the annual summary of income distribution for each country based on personal income tax, the publication of monthly count of unemployment insurance claimants and beneficiaries by province, age, sex, length and type of benefit.

2. *Direct production of information through the integration of multiple administrative sources.*
   This process involves the integration of different administrative databases. In this regard, the definition of the procedures necessary for the creation of statistical registers, i.e. exhaustive and updated lists of units belonging to population of statistical interest, is crucial.
   In general, this category includes any use of administrative data which involved the linkage of different administrative or statistical sources. For example, the combination of the death register with records from individuals exposed to a particular hazard in order to estimate different mortality rates, or the combination of data from tax returns and unemployment insurance files in order to analyse labor market attachment and adjustment.

3. *Indirect use of data from administrative sources to support sta-*

tistical surveys.

In this category, it is included any use of administrative data to create, supplement or update frames to be used in censuses and surveys, as well as the use of these records for checking, validating or evaluating data coming from survey.

Examples are the use of administrative data to identify changes in the statistical registers, and the use of immigration/taxation records to evaluate census questions on immigration/income.

Since the use of administrative data for statistical purposes may have several advantages as well as potential drawbacks, it is useful to provide some reasons for reflection.

2.4 Differences between administrative and survey data

In the production of official statistics, the available data may come from statistical surveys and administrative sources. Surveys differ from administrative data since they are collected specifically for statistical purposes. They are generally more flexible than administrative sources as they are conducted to answer specific classes of research or public policy questions. Some advantages of survey data over the administrative ones include the targeting of a specific population and variables of interest, the interaction with respondents, and the guarantee that the data will be used exclusively for statistical purposes. By contrasts, potential drawbacks of survey data include difficulties in the construction of a suitable population frame, high costs to build a large sample size, unit or item non-response, and measurement errors.

Based on the work of Johnson and Moore (2005) and Penneck (2007), strengths and weaknesses of administrative records compared to those of surveys are summarized in the following three sub-Sections.

Population coverage

The target population defined by administrative systems is based on the legislation requirements needed for the registration of a specific program. For this reason, a major advantage is that statistics usually cover the whole target population. This provides more reliable results for sub-populations, i.e. small geographic areas or groups with specific characteristics. Therefore, it is especially useful for those interested in local area information.

By contrast, the population of sample surveys is usually defined by the analytical purpose and the budget available. Coverage, definitions, methodology and timing are designed to meet analytical needs. The sampling frame for most of the surveys is derived from existing sources, such as geographically based population data, address listings, telephone directories, or administrative sources. However, sample sizes can be small, and the wrong sampling frame may bias the results obtained. Besides, the units of observation are determined by the data required for the statistical purpose. The definition of the units can be also influenced by the type of population designed for the experiment. For all these reasons, surveys have limited use for analysing small populations or local areas.

Besides, the potential cost of obtaining the information, interviewing a sample of the desired size, or reaching a rare population can be prohibitive. More generally, the set-up costs for a statistical survey are almost as high as the ones of using administrative data to collect

the same information, but the running costs are usually significantly higher. Even if computers take over a large part of the processing burden, the whole approach for the realization of a sample surveys is still labour intensive and it probably never be fully automated, referring to the response chasing stage.

On the other hand, a drawback of the administrative population is that, in some case, individuals may need to undertake some actions to become part of the administrative system, and it is important to consider what encourages them to register to a program. There may be factors to avoid registration, especially if their situation places them close to a threshold that requires compulsory participation, or gets associated with financial costs, such as setting a minimum wage on which contributions are calculated. The resulting statistics on this type of population are highly relevant to the management of the system, but the use of them for broader purposes immediately rise quality issues that may lead to misleading conclusions.

Contents

The content issues that researchers need to consider while working with administrative and survey data are several. One of them is the purpose for which administrative data are collected, which may have a significant impact on their usefulness for statistical analysis regarding the amount of available data, data definitions, consistency between different time periods, and data quality. The usefulness of administrative registry systems may be limited because only the variables needed to manage the related program are collected in the database. These variables can only be a small part of the data reported in an administrative form, implying the necessity to link

the administrative records with other sources to fulfill the statistical purpose. This fact underlines the noteworthy value of big data as an additional source of information and consequently, as already pointed out, the importance of integrating different data sources available.

The previous issue is extended also to the definition of concepts and variables. For administrative records, the definitions are based on the legislation, and these can differ to those needed for social or economic analysis. This problem can be solved first trying to understand the differences and quantify the impact of such dissimilarities; later, systematic differences can be adjusted with additional data or editing the existing.

An advantage in the use of administrative data is that they may allow to produce statistics more frequently than sample surveys without extra response burden and with the addition of minimal extra cost. In fact, statistical surveys generally require more time to plan and design, reach the target population, and adjust the related issues, especially in the case of annual or ad-hoc data collections.

The main problem connected to the use of administrative source is about the timeliness. Administrative statistics will be regularly updated by the administrative system, but their timeliness will be driven by the needs of that system. Administrative sources that are not based on a particular time period are more flexible, since they are available once the data are registered. An example is administrative sources based on record events, e.g. births or deaths registration. On the other hand, administrative data, such as tax returns, are updated only once a year, so it is difficult to make them available for more frequently statistics.

The close linkage between the administrative population and the administrative system can also rise problems of continuity. All the features of the administrative population are defined by the needs of the administrative system: definition, units, classification, coverage, frequency, and timing. Changes in policies and rules may fluctuate the population taken in study from year to year, having serious implications for the resulting statistical measures.

The units of observations available from administrative data often focus on identifying specific entities. An example is the case of income tax return where married couples may fill out it jointly, but they are also allowed to compile the declaration separately in case of tax advantages. Differences in economic units reported in the administrative records may limit their usefulness for statistical purposes.

The problem of missing data is common to both sources. In administrative sources, some variables are not a primary importance, so they are collected only for certain categories of units, or they are not collected at all. On the other hand, surveys are dealing with full or partial non-response or the necessity to remove some values during the editing process. In fact, interviewers usually try to convince respondents to voluntarily participate at the survey but, even for respondents who agree to attend and respond to all survey questions, measurement errors are still a concern. Respondents can refuse to answer some questions due to fears in revealing data to protect privacy, or they may have difficulty to remember past events. Other typical problems about measurement errors include rounding of amounts and misunderstanding of questions. For administrative data, non-response is not a primary problem. It is however true that

in certain cases individuals tend to misreport the true value. For example, this may happen in tax income reports in order to reduce the required amount of taxes.

Privacy

As already pointed out, data privacy is a very important issue. Any use of administrative data for research purposes should consider the laws that protect the privacy of the data. In most cases, the use of these data is limited within the administrative agency or under an agreement that allow employees of various institutions to exchange the information.

Besides, due to privacy and confidentiality laws, the available data, both in case of administrative and statistical survey, do not contain unique identifying variables. For example, the identification code which directly identifies the individual is always missing. The availability of this piece of information could lead to wider statistical use and to the possibility to link data from different sources, which is essential for the quality assurance process.

Another discussed question is about the public opinion related to the sharing of personal data. Data sharing is often frowned upon by the public but, on the other hand, this can greatly improve government efficiency in the use of public resources. A key point is therefore the communicative strategy through which data sharing is presented by local and central authorities. A way to deal with public hostility could be to publish costs and benefits analyses of using different sources. The fact that no questionnaires are sent, data are not stored on paper or electronically by the interviewers, so fewer people have access to sensitive data, could favour the use of admin-

istrative sources over surveys. On the other hand, the employment of administrative data reduces the direct contact with the public, and at the same time public awareness about its usefulness.

In conclusion, if compared to each other both survey and administrative data present advantages and drawbacks. The key point is the possibility to have access to administrative data by the statistical agencies and the level of control over quality that researches can have over the data.

A further challenge is the possibility to link administrative records from different system programs together and/or with survey or census data to enhance the research potential of the data. The combination of highly reliable administrative records with detailed survey data opens new scientific and policy-related research opportunities. However, despite data linkage has been improved and most of the technical problems have been solved, it remains challenging and time-consuming. The main drawbacks of linking administrative data with survey data for research purposes derives from legal constraints related to privacy concerns, which have led to restrictions on data usage (Künn, 2015).

With the growing attempt to reduce costs and the increasing interest of data at a local level, the wish is that governments enhance the use of administrative data for statistical purpose. They should put effort in the improvement of the collaboration between administrative and statistical agencies and the implementation of automatic procedures that edit and clean the data, in order to make them easily available to statisticians.

2.5   Data and methodological issues

In this Section, we present the integrated source of administrative data used in this work. It links data on the Italian personal income tax returns coming from the Ministry of Economy and Finance (MEF) with data coming from the Registry Office of the Municipality of Bologna through the individual fiscal code for the period 2002-2017.

This database, although experimental, lends itself to becoming a useful source of information for the elaboration of statistics, and an important policy instrument to support local administrations. Indeed, the underlying idea is to use this dataset to display the informative potentials of a tax administrative source, and the possibility to use it for exploring the medium-term distributive income trend in a local area.

The following two Paragraphs will present the main aspects of the administrative source used, describing first the fiscal report dataset and its main variables, and then the characteristics of the municipal registry and its informative content.

2.5.1   Fiscal reports dataset

Our fiscal information is based on the Italian personal income tax[2] (Irpef) for the period 2002-2017.

Table 2.1 lists variables from the tax reports. Together with a detailed set of income components, the dataset contains information on the total income assessed, the net personal income tax and the

---

[2]For further details about the Italian personal income tax, see Appendix A.

Table 2.1: List of variables coming from the MEF dataset on personal income tax returns, available for the the fiscal years 2002-2017.

Variable	Label
ANNO	Year
MODELLO	Type of tax report presented by the taxpayer (730, 770, UPF)
Income components:	
IMP_TERR_DOM	Dominical income
IMP_TERR_AGR	Agricultural income
IMP_ALLEV	Breeding farm income
IMP_FABB	Buildings income
IMP_CAPITALE	Capital income
IMP_LAV_DIP	Income from dependent work and "assimilated"
IMP_LAV_DIP_NDETR	Income "assimilated" to dependent work for which deductions are not due
IMP_ATT_PROF	Self-employment income
IMP_LAV_AUTO_ALT	Other incomes from self-employment
IMP_SPORT	Income from amateur sports activities
TAX_SEP	Income where separate taxation is applied
IMP_PARTECIP	Participation income
IMP_CONT_ORD	Ordinary accounting business income
IMP_CONT_SEMP	Simplified accounting business income
DIVERSI	Other incomes
Taxable income:	
IMP_TOT	Taxable income for personal income tax
IMP_ADD_CMN	Taxable income for municipal income surtax
Net taxes:	
IMP_NETTA	Net personal income tax
ADD_CMN_DOV	Municipal income surtax due
Others:	
ADD_CMN_CERT	Municipal income surtax resulting from the certification
ADD_CMN_SOSP	Municipal income surtax suspended
CRED_ADD_NRIMB	Credit for municipal income surtax not reimbursed

Source: Data from MEF.

municipal surtax. The type of form[3] presented is also disposable.
The total individual taxable income[4] constitutes the logical corner-
stone of many analyses conducted later. For this reason, it requires
some clarifications. The first relates to the definition of income for
tax purposes. In our dataset, the main income variable is the taxable
income, defined as the total income net of deductions. Therefore,
our definition of income is that of gross income, including all in-
come items reported in the personal income tax returns (salaries
and pensions, self-employment and business income, farm income,
real estate income, and other smaller income items), after all deduc-
tions are subtracted.

The taxpayers share for which the difference between total income

Table 2.2: Composition of total income in the period 2002-2017 for the Munici-
pality of Bologna.

Category	2002	2007	2009	2017
Real Estate	8.3	7.1	7.5	3.1
Dependent Work	71.7	75.6	81.1	85.2
Self-Employment	7.8	7.9	7.8	7.8
Business	9.3	8.1	2.5	2.3
Capital	1.5	0.5	0.5	0.9
Other	1.3	0.7	0.6	0.7
Total	100.0	100.0	100.0	100.0

Source: Data processing on MEF data.

and taxable income is most significant is that of self-employed, since
the possibility of deducting all forms of social security and com-
pulsory assistance contributions from the total income is envisaged
only for those taxpayers. However, the share of income from self-

---

[3]For further details about the procedures for declaring the taxable income to the Revenue
Agency, see Appendix C.
[4]For further details about the identification of the tax base for the Italian personal income
tax, see B

Figure 2.1: Dynamics of taxable income for personal income tax and taxable income for municipal surtax in the period 2002-2017 (nominal values).
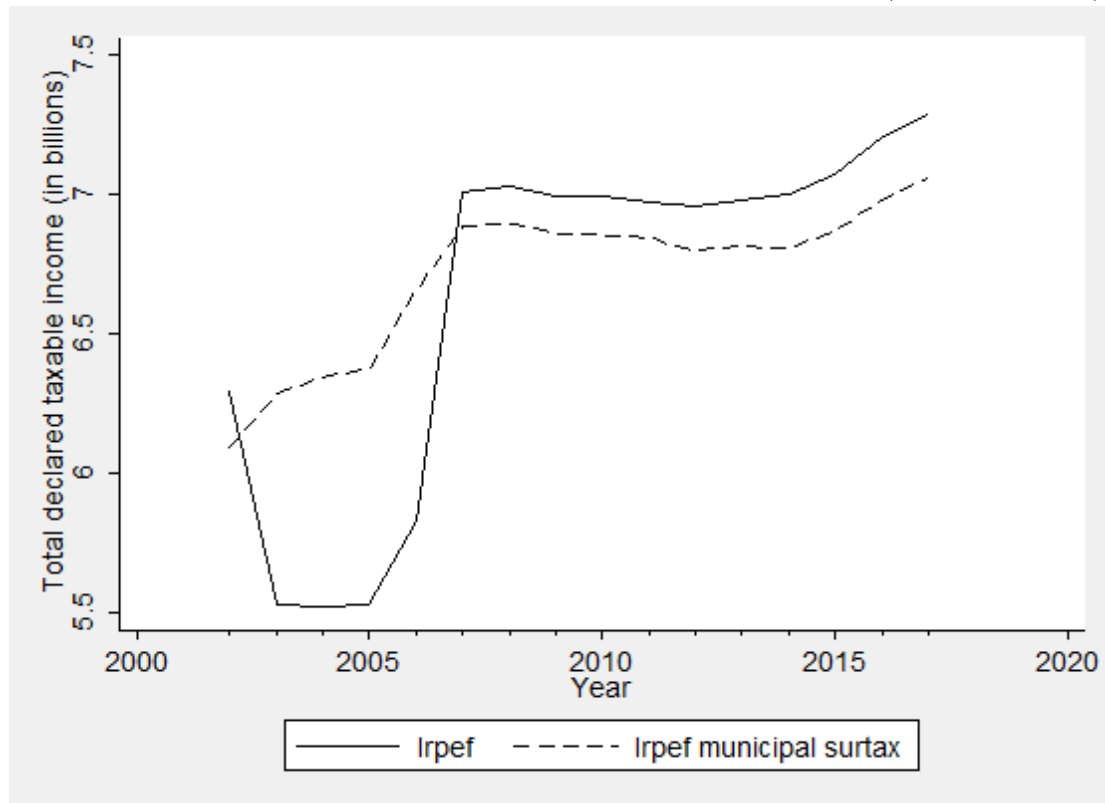
Source: Data processing on MEF data.

employed is approximately 8 per cent of the total income declared in the Municipality of Bologna. Right from the beginning of the introduction of the Italian personal income tax, 80 per cent of the total declared income in Italy derives from income from dependent work and retirement benefits[5], similar to what happens today in the Municipality of Bologna. Table 2.2 confirms the progressive trend, accentuated in more recent years, to concentrate the levy on income from employees and retirement benefits, which constituted, in 2017, 85 per cent of the total income declared. In 2017, the remainder of the taxable income reported by the taxpayers consisted of self-employment income (7.8 per cent), which has remained constant

---

[5]For further details, see Pellegrino and Panteghini (2020)

since the beginning of the observed period; real estate income (3.1 per cent), which has declined sharply since 2002 due to the exclusion of certain components of building income from total taxable income; and business income (2.3 per cent), which in turn has declined markedly due to the exclusion since 2011 of investment income. Realized capital gains went mostly un-taxed and not reported in the tax rates before the 2000s; since then, gains from qualified equities have been reported at varying degrees. Consequently, the total taxable income covers capital income incompletely and excludes most capital gains. In fact, in 2017 the share is equal to 0.9 per cent.

However, our dataset contains two possible definitions of taxable income (see Table 2.1): the first refers to the personal income tax, whereas the second refers to the municipal surtax. Figure 2.1 presents the temporal evolution of the total taxable income for the national and local personal income tax. Since there is greater continuity in local income[6], we choose the taxable income for the municipal surtax as main variable for our analyses. However, changes occur anyway due to the evolution in the Italian legislation during the 2000s; for this reason, data must be taken with a grain of salt, especially those prior to the 2007 tax year.

2.5.2   Registry Office dataset

The second part of the dataset comes from the Registry Office of the Municipality of Bologna, which collects demographic information on registered individuals living or fiscally domiciled in the Municipality

---

[6]Tremonti reform (L. 7 April 2003, n. 80) changed the structure of the personal income tax in the period 2003-2006, reversing the role of deductions and allowances for income from work and family charges. Afterwards, government reformulates the tax rules again, applying the structure that is still in force today. However, this change in the legislation do not affect taxable income for the municipal surtax purpose.

of Bologna, as well as information regarding family composition (see Table 2.3 for the complete list of variables in our dataset).

Specifically, three groups can be distinguished. The first one re-

Table 2.3: List of variables coming from the Registry Office dataset of the Municipality of Bologna available for the period 2002-2017.

Variable	Label	Notes
CF_CRIT	Encrypted tax code	
MATRICOLA	Individual serial number	
RESIDENZA	Residence	1. Resident at 31/12 of the reference year; 2. Resident during the year; 3. Resident during another year; 4. AIRE; 5. Never resident.
ETA	Age	
SESSO	Gender	
CITTADINANZA	Citizenship	
COMUNE_NASCITA	Municipality of birth	
PROV_NASCITA	Province of birth	
STATO_CIVILE	Marital status	
CENTRO_PERIF	Place of residence	1. City centre; 2. Suburbs; 3. Homeless.
QUARTIERE	Neighbourhood	
ZONA_AMM	Administrative area	
AREA_STAT	Statistical area	
ANNO_IMMIGR	Year of immigration	
COMUNE_IMMIGR	Municipality of immigration	
PROV_IMMIGR	Province of immigration	
NUMERO_FAMIGLIA	Family serial number	
COMPONENTI	Number of components of the family	
PERCETTORI	Number of taxpayers in the family	
TIP_FAMIGLIA	Type of family	

Source: Data from MEF.

gards personal information, which is usually highly reliable. This group includes information which identifies the individual, i.e. the tax code and the individual serial number. The tax code constitutes the linkage key between the registry and fiscal archive. After hav-

ing identified the corresponding information in tax return form, the
Municipality of Bologna replaced the tax code with an anonymous
individual identifier, in order to ensure the anonymity. On the other
hand, the individual serial number, who constitutes the linkage key
between the registry files belonging to different years, allows to set
the longitudinal dimension.

The second category is formed by certifiable information, which
normally is sufficiently reliable, since they are being updated and
checked periodically, for example when issuing certificates or renew-
ing the identity cards. The variables of this category regards resi-
dence, age, gender, citizenship, place of birth, marital status, and
other detailed information about the residence. For this category it
is necessary to provide some further explanation.

The key variable in this group is the residence of the taxpayer. It

Table 2.4: Shares of taxpayer by residence variable in the period 2002-2017.

Residence	2002	2009	2017
Resident al 31/12	93.3	93.4	92.8
Resident during the year	2.7	2.3	2.8
Resident during another year	1.8	2.2	2.0
AIRE	0.1	0.2	0.4
Never Resident	2.0	1.9	2.1
Total	100.0	100.0	100.0

Source: Data processing on MEF data.

consists of a general variable that indicates if the individual is resi-
dent or not in the municipality or if it is resident abroad. Since this
information is imported from the Municipal registry lists, detailed
information is available only for the taxpayers resident in the Mu-
nicipality of Bologna, that is those who have been resident for all
the preceding twelve months or only partially during the reference

year, or those who were resident during one of the preceding years. For those who are only fiscally domiciled in the municipality, demographic variables are not available. However, the share of taxpayers never residents in the Municipality of Bologna is only a small portion of the total (see Table 2.4), nearly the 2 per cent for all years from 2002 to 2017.

Only for those residing in the Municipality of Bologna, detailed information about residence are available. Specifically, three more variables, that indicates the neighbourhood, the administrative area and the statistical area of residence, are available. The presence of such a territorial dimension for each single record makes possible to carry out analyses at a lower level of territorial granularity than those conducted with the survey data.

The variable related to the neighbourhood is made up of the nine historical districts: Borgo Panigale, Navile, Porto, Reno, San Donato, San Vitale, Santo Stefano, Saragozza and Savena. On the other hand, the variable related to the administrative area let a more depth-in analysis possible, since it is composed by the eighteen administrative area: Barca, Bolognina, Borgo Panigale, Colli, Corticella, Costa-Saragozza, Galvani, Irnerio, Lame, Malpighi, Marconi, Mazzini, Murri, Saffi, San Donato, San Ruffillo, San Vitale and Santa Viola. Finally, the variable related to the statistical area of residence divides the Municipality in 92 different sections[7], reaching

---

[7]Aeroporto, Agucchi, Arcoveggio, Bargellino, Battindarno, Beverara, Bitone, Borgo Centro, CAAB, CNR, Cadriano-Calamosco, Canale di Reno, Caserme Rosse-Manifa, Casteldebole, Cavedone, Chiesanuova, Cirenaica, Corelli, Croce Coperta, Croce del Biacco, Dagnini, Ducati-Villaggio Ina, Due Madonne, Emilia Ponente, Ex Mercato Ortofrutticolo, Fiera, Fossolo, Galvani-1, Galvani-2, Giardini Margherita, Guelfa, Irnerio-1, Irnerio-2, La Birra, La Dozza, La Noce, Laghetti del Rosario, Lavino di Mezzo, Lazzaretto, Lungo Reno, Lungo Savena, Malpighi-1, Malpighi-2, Marconi-1, Marconi-2, Mengoli, Mezzofanti, Michelino, Monte Donato, Mulino del Gomito, Ospedale Bellaria, Ospedale Sant'Orsola, Osservanza, Paderno, Pescarola, Piazza dell'Unità, Pilastro, Ponte Savena-La Bastia, Pontevecchio, Prati di Caprara-Ospedale Maggiore, Ravone, Rigosa, Roveri, San Donnino, San Giuseppe, San Luca, San Michele in Bosco,

a very detailed level of territorial dimension, which on the other hand could lead to a complicated interpretation of the outcomes. In addition, information relating to immigration, if it was occurred from another Italian municipality, is available. These variables concern year, municipality and province from which the individual moves to the Municipality of Bologna.

Another group of variables related to the category of certifiable information available in the dataset regards the family. In particular, the registry list includes information about the family serial number, the type of family, the number of components and the number of taxpayers within the family. For this group, the presence and quality of the information available is strictly dependent on the quality of the information in the personal data lists from which they are derived. These variables are affected by a coverage error, maybe derived from a failure to regulate personal data of changed family situations or by opportunistic behaviours, such as pretended main residences.

The information available with a good level of quality is only those for taxpayers residing in the Municipality of Bologna at 31/12 of the reference year, for whom the share of missing information is less than 1 per cent (see Table 2.5). For the taxpayers who were resident during the year of reference or those who were resident during another year, the information at our disposal contains an high percentage of missing values. For those belonging to AIRE and for the taxpayers fiscally domiciled, information about the family is totally missing. For this reason, the subsequent analyses exclude those who are not resident in the Municipality, for whom it is not possible to inspect

---

San Savino, Savena Abbandonato, Scalo Merci San Donato, Scalo Ravone, Scandellara, Siepelunga, Stadio-Meloncello, Stradelli Guelfi, Tiro A Segno, Triumvirato-Pietra, Velodromo, Via Arno, Via Ferrarese, Via Larga, Via Mondo, Via Toscana, Via Vittorio Veneto, Via del Genio, Via del Lavoro, Via del Vivaio, Villaggio della Barca, XXI Aprile, Zanardi

Table 2.5: Shares of missing information by residence variable for the year 2017 (percentage values).

Residence	Family serial number	No. of components	No. of taxpayers	Type of family
Resident at 31/12	0.9	0.9	0.9	0.9
Resident during the year	0.1	63.8	66.5	63.8
Resident during another year	0.7	79.7	81.9	79.7
AIRE	0.0	100.0	100.0	100.0
Never Resident	100.0	100.0	100.0	100.0
Total	2.9	6.6	6.8	6.6

Source: Data processing on MEF data.

the aforementioned areas of investigation with certifiable quality.

Finally, the third group of variables is composed by non-certifiable information, such as qualification and profession, whose quality is unsteady because it depends on the citizen's declaration and its information content is not subject to periodic updates. For this reason, we have decided not to use these information for our subsequent analysis.

Although the administrative nature of the data makes the set of socio-demographic variables available more subtle than a sample survey, the available variables have however made possible interesting analyses.

2.6 Dataset quality assurance

One of the main obstacle in using administrative sources for statistical purposes relates to the quality of processes and outputs. The Quality Guidelines for Statistical Processes published by the Italian National Institute of Statistics (Signore et al., 2012) aim at describing the principles to be followed when planning, running and

assessing a statistical process, as well as at illustrating quality requirements of statistics.

For the purpose of measuring the quality of the statistics products,

Table 2.6: Quality dimensions for the assessment of the quality of statistical outputs.

Quality Dimension	Definition
Relevance	The degree to which statistics meet the need of users and correctly describe the phenomenon they were designed to measure.
Accuracy	The degree of closeness of statistical estimated results to unknown true values.
Reliability	The trustworthiness of the data supplied and how they are reliable over time and geographies.
Timeliness	The length of time between published data and the event they refer to.
Puntuality	The time lag between the actual publication and the planned publication dates.
Accessibility	The actions taken in order to help the user find the data he or she is interested in.
Clarity	The presence of sufficient and appropriate metadata to help users understand the data.
Coherence	The degree to which data referring to the same event but derived from different sources are similar.
Comparability	The degree to which data can be compared over time, space and domain.

Source: Eurostat, 2003.

ISTAT adopted the definition of quality issued by Eurostat in 2003 (Eurostat, 2003), later taken up by the European Statistics Code of Practice (European Commission, 2011), and the Italian Code of Official Statistics (Official Gazette 13 October 2010, n. 240). In this

sense, the quality of the statistics produced and disseminated must be assessed with reference to the criteria of Table 2.6.

Measuring the quality of statistics produced with the use of administrative data according to the dimensions listed above is not straightforward. In fact, a direct quantitative measurement is possible only for some components, while for the other dimensions often only judgments can be formulated. Besides, these standard quality dimensions are not always applicable in order to assess the quality of administrative data, when there is not sufficient supporting information about the population and the collection process, or some sort of quality survey to determine the correct values of certain variables are not available. The difficulty about this process is that there is often no "gold standard" measure (Oberski et al., 2017). Given that surveys do not contain measurement errors, some authors have suggested to link administrative data to a survey (Yucel and Zaslavsky, 2005). Anyway, it is well-known that even in survey data errors are widespread. In fact, researcher often measure errors in surveys using administrative data as "gold standard" (Kapteyn and Ypma, 2007). Regarding the quality of the measurement process, when the input comes from a collection process that is exogenous from the National Institute of Statistics control, it is of utmost importance to check the quality of the administrative data source acquired, taking into account the specific statistical production objective. In the case of direct production of statistics, through the replacement of units and survey variables with administrative data, the target population, the variables, the geographical and time dimensions assume particularly importance. With respect to these elements, the National Institutes of Statistics recommended to carry out an objective analysis on the

usability of the administrative data in the statistical production process for the specific purposes.

Firstly, starting from the definition of the target statistical population, the time references and the geographical boundaries, a careful analysis should be made on the ability of the available administrative dataset to properly cover the units of the target population. Likewise, the correspondence between concepts and definitions that concern the statistical variables of interest and those regarding the variables inferred from administrative data should be evaluated. In this regard, the quality aspects on which to focus attention concern the coverage levels expected for the population of interest, the validity of the variables used, and the evaluation of possible bias arising from inappropriate use of the administrative data.

Given that definitions and main aspects of our administrative dataset together with the advantages and potential drawbacks of using it for statistical purposes have been clearly laid out in the preceding pages, in the next Section the general assessment of the quality of our administrative data source is supplemented with a comparative analysis with a sample-based survey.

2.6.1   A comparison with the IT-SILC sample survey

One of the most widely used sample surveys for the analysis of income distribution is the Italian Survey on Household Income and Living Conditions (hereafter IT-SILC). It is part of a wider European information system[8], which has as the priority purpose to produce, using harmonized definitions and methods, comparable data, both at a transversal and longitudinal level, for the analysis of in-

---

[8] "European Statistics on Income and Living Conditions" (abbreviated EU-SILC) approved by the European Parliament Regulation, n. 1177/2003 and coordinated by Eurostat.

come distribution, well-being, and quality of life of families, as well as of the economic and social policies adopted at national and/or European level. The survey mainly measures the different types of income received by the various members of the household in the calendar year preceding the interview; participation in the labour market in the twelve months preceding the interview, and the living conditions (e.g. housing situation, any situations of hardship or material deprivation, etc.) at the time of the interview.

In order to assess the quality of our administrative dataset, we make use of data from 2017 IT-SILC survey covering the tax year 2016. The 2017 edition of the Survey reached on the national territory an actual sample of about 22 thousand households, for a total of almost 49 thousand individuals of at least 16 years old, spread over more than 880 Italian Municipalities of different demographic size. Families are randomly extracted from the national registry files or from the list of names already involved in the permanent census of population and households, according to a sample design that makes them statistically representative of the resident population in Italy. The sample design of the SILC dataset is planned to provide reliable cross-sectional estimates up to the regional level, and longitudinal estimates at the level of geographical distribution. Hence, the use of the IT-SILC data to produce estimates at a lower territorial level (i.e. province or municipality) is not foreseen except under the direct responsibility of the applicant.

Considering weights[9], in the IT-SILC dataset the universe of house-

---

[9]In the EU-SILC data, the calibrated household weight is assigned to each member who responds to the survey. For further details about the sample selection and the dataset description, see European Commission (2017).

Table 2.7: Distribution of individuals and households in the weighted IT-SILC sample survey and in the MEF dataset in 2017 by age, sex and marital status.

Variable	IT-SILC		MEF	
	Abs	*Percentage*	*Abs*	*Percentage*
Individuals	297,164	100	278,071	100
Sex				
Women	157,079	53	145,521	52
Men	140,085	47	132,550	48
Age groups				
Less than 40	115,100	39	71,696	26
40-60	76,845	26	99,454	36
60-75	49,094	17	57,331	21
Greater than 75	56,125	19	49,590	18
Marital status				
Single	100,318	34	95,631	34
Married	146,602	49	138,641	50
Divorced	9,428	3	13,372	5
Widowed	40,816	14	30,427	11
Citizenship				
Italian	234,762	79	248,808	89
Foreign	62,402	21	29,263	11
Households	174,057	100	187,387	100
Sex				
Women	86,650	50	86,587	46
Men	87,407	50	100,800	54
Age groups				
Less than 40	58,047	33	45,296	24
40-60	47,779	27	68,370	36
60-75	29,150	17	37,697	20
Greater than 75	39,081	22	36,024	19
Marital status				
Single	58,399	34	69,284	37
Married	75,317	43	80,378	43
Divorced	5,473	3	11,376	6
Widowed	34,868	20	26,349	14
Citizenship				
Italian	136,560	79	166,718	89
Foreign	37,497	21	20,669	11

Source: Data processing on MEF and ISTAT data.

holds is composed by 174,057 households[10], whereas in our administrative dataset there are 187,387 households. Table 2.7 shows the composition for both datasets considering different individual variables: sex, age, marital status and citizenship. Values are reported both for individuals and households. The variables considered seem to show rather similar values, apart from age and citizenship. In this case, there appear to be higher values in the sample survey referring to those under 40 years old, in spite of those in the 40-60 age group, and higher values for the foreign citizens, whether individuals or households are considered. The reason can be attributed to the fact that the calibrated weights obtained from ISTAT are designed to produce reliable estimates at the regional level. This is no longer true when smaller sub-samples, i.e. provinces and municipalities, are considered. In fact, the calibration adopted requires that the regional estimates of the variables (sex, age, etc.) reproduce exactly the known regional totals, but does not guarantee that these totals are met at the sub-population level.

Aware of the fact that using IT-SILC data at the municipal level is not recommended by the guidelines, we compared our data with the IT-SILC ones on household taxable income.

A picture of the Bolognese households income distribution in 2016 is provided by Figure 2.2, which depicts the percentage of Bolognese individuals reporting different levels of gross household disposable income in both the IT-SILC and MEF dataset[11]. Note that the distribution is slightly higher for the MEF dataset at the bottom; ceteris paribus, this would imply a lower average income than the

---

[10]To make the comparison more consistent, we selected from the survey sample only individuals with a non-missing taxable income, and capital incomes are excluded.

[11]MEF income values are restricted on a range of values between 0 and 200,000 Euros to make the comparison more clear.

Figure 2.2: Income distribution in the IT-SILC and MEF dataset in 2016.
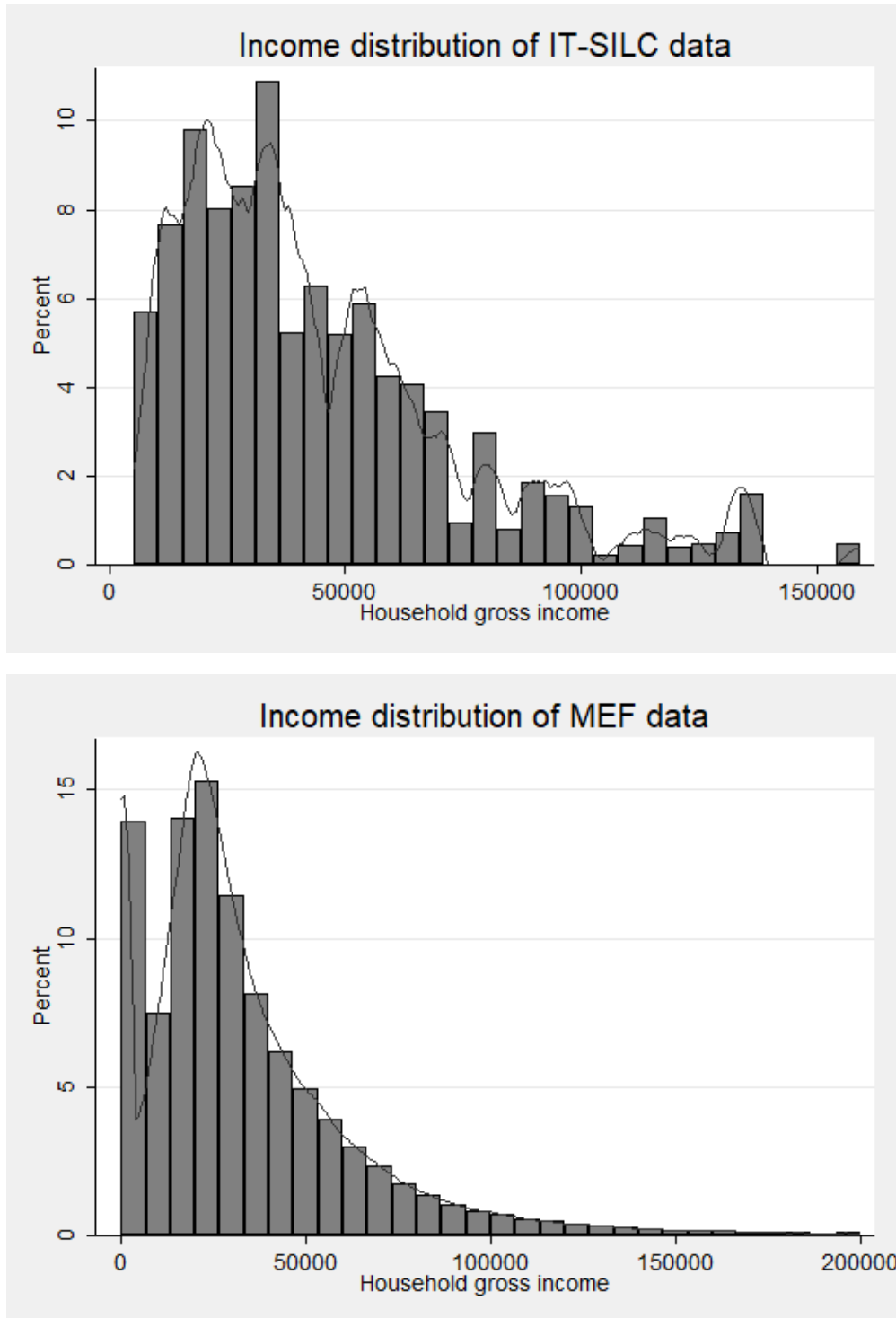
Source: Data processing on MEF and ISTAT data.

IT-SILC one. The average MEF income in 2016 tax year (41,350 Euros) is, however, very close to that calculated with the IT-SILC data (41,754 Euros). These two results are compatible with each other only if the MEF frequency density function is, for high income levels, higher than the IT-SILC one. Indeed, it is, even if it does not appear clearly from the Figure 2.2.

However, the shape of the distribution is substantially the same: asymmetrical with a greater concentration of incomes in the lower-middle classes and few high incomes. A consequence of the positive asymmetry of the distribution is a higher level of the average income compared to the median one, which is not influenced by the presence of any high values, and, generally, tends to be lower than average income for this type of distributions. In fact, the MEF median income is equal to approximately 29,000 Euro, while for the IT-SILC dataset it is equal to approximately 33,000 Euro.

Table 2.8 also shows a comparison performed on household income

Table 2.8: Average and median household taxable income and personal income tax in the IT-SILC and MEF dataset in 2016 (values in Euro).

Variable	IT-SILC	MEF	% Difference
Gross income			
Average	41,754	41,350	-1.0
Median	33,313	29,012	-12.9
Net income			
Average	32,384	31,687	-2.2
Median	26,765	23,838	-10.9
Personal income tax			
Average	9,476	9,948	5.0
Median	6,913	5,107	-25.8

Source: Data processing on MEF and ISTAT data.

net of social contributions, both gross and net of the personal income tax, and the average and median national personal income tax

for the tax year 2016. Evidences shows that median income is consistently higher in the IT-SILC data than our administrative based measure. Average income is also higher, even if to a lesser extent, in the IT-SILC estimates, with the gap between mean and median values larger for the administrative based measure across both the comparison.

The above mentioned pattern was consistent when comparing in-

Figure 2.3: Average income by quantiles of income distribution in the IT-SILC and MEF dataset in 2017.
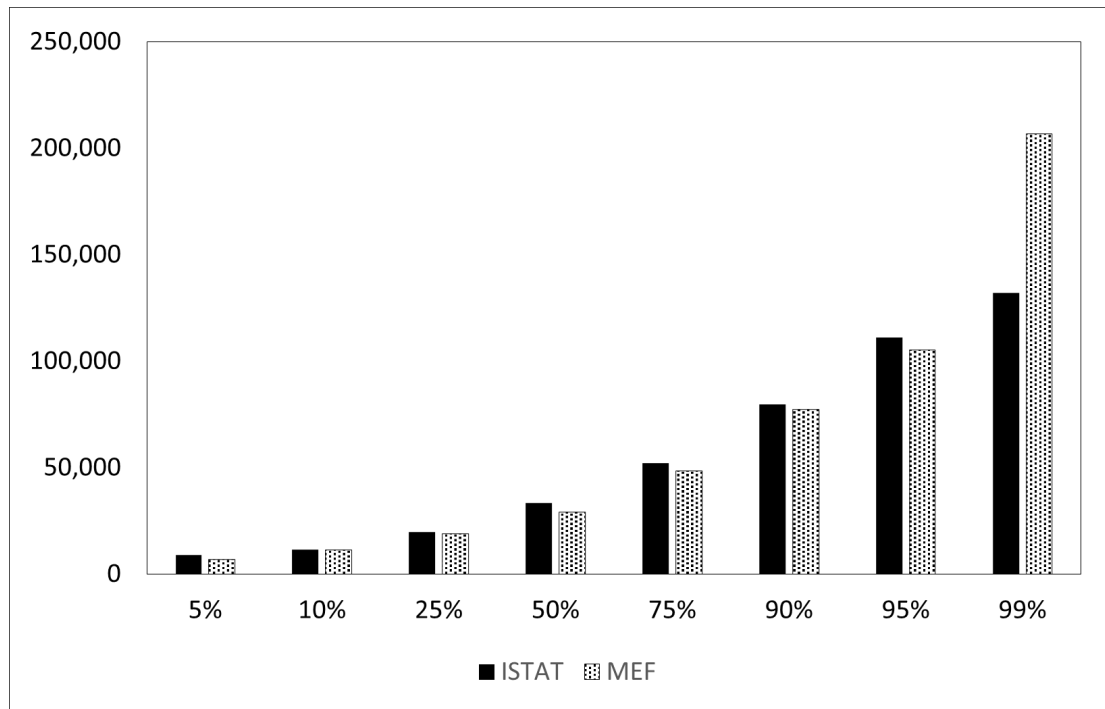
Source: Data processing on MEF and ISTAT data.

come within income quantiles. In fact, the analysis by percentiles of income shows how the average income held by the highest percentile is significantly higher in the administrative data than the survey data (see Figure 2.3). The fact that the average income of the 99th percentile reaches much higher values in the administrative data than the sample data and then implies that the share of income owned by

the highest percentile is more than double in the MEF data (9.5 per cent) compared to the IT-SILC data (4.5 per cent). In comparison with survey data, administrative data are much better at capturing very high earners, with the highest household gross income being over 180 times higher in the MEF data than IT-SILC. While this does not explain why IT-SILC estimates of income are higher than the MEF ones, it could explain why the differences between average income are smaller than those between medians.

The fact that through administrative data is possible to examine

Figure 2.4: Percentage of non-taxable incomes by percentiles of the IT-SILC gross household income distribution in the tax year 2016.
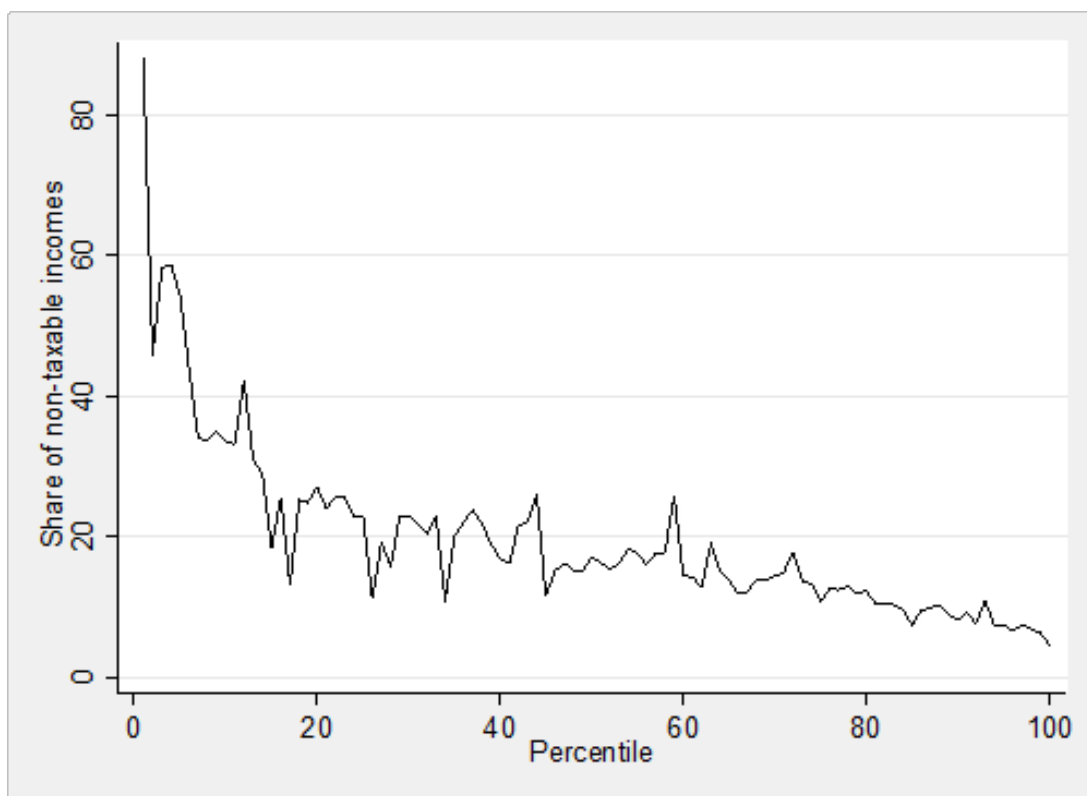
Source: Data processing on ISTAT data.

the appropriate share of income of the richest is a very important aspect. Until very recently, nothing was known about this feature in

the distribution of income. This gap has begun to be filled thanks above all to the work of T. Atkinson and T. Piketty who published the first studies on top income shares (Piketty, 2014). This form of ignorance was largely due to a lack of data, and secondly to the scant attention that the economic literature has devoted to the very rich. The examination of the top income shares is important because it enriches the analysis of inequality, allowing to observe the trends relating to the high tail of the distribution with a much greater precision than that ensured by the sample surveys, which, due to their limited number of observations, run the risk of not adequately representing high-income individuals in the sample (Acciari and Mocetti, 2013).

In the administrative dataset, individuals are not legally required to submit a tax return for given categories and values of income[12]. However, our analysis shows that 13 per cent of IT-SILC respondents reported an income less than Euro 3,000, while 17 per cent of individuals submitting tax returns in 2016 reported incomes below this threshold. After having removed individuals with zero income, this proportion is approximately the same in both dataset: 4 per cent in the survey data and 3 per cent in the MEF data.

In this regard, another factor to be taken into consideration is about non-taxable incomes, i.e. all income components not subject to the personal income tax, which, given the origin, escape from our dataset. In fact, one of the advantage of the survey data is that it contains information about non-taxable incomes, e.g. social, maternity, attendance, and pension allowances and contribution for rentals and household expenses. These components have a relatively im-

---

[12]For a detailed explanation of exempt incomes, see Appendix B.

portant weight, especially in the left tail of the income distribution. Figure 2.4 shows the share of non-taxable incomes on the distribution of total household income of the IT-SILC interviewees. The weight of this component reaches very high levels for the lower percentiles of the distribution (between 50 per cent and 80 per cent in the first 5 percentiles), and then reduces to less than 10 per cent in the higher percentiles.

In light of the potential distorting effect and the advantages of administrative data, in the next Chapter the main results of the analysis of income distribution and inequality indicators are shown using the administrative database presented.

Chapter 3

Income distribution in the Municipality of Bologna, 2002-2017

3.1 Introduction

Government plays an important role to reduce inequality and encourage a more inclusive growth. Drawing on the research about income inequalities and its trends over time, it can address adequate policies to ensure a more even distribution of income and opportunities. Against this background, this Chapter exploits tax data with the aim of responding to the informative needs of local decision-makers for the programming of the service plan and municipal taxation.

The sample surveys typically used for the analysis of income distribution, such as IT-SILC data, do not allow for sufficiently fine partitions of the territory. On the contrary, thanks to tax data linked to the municipal Registry Office data, in this work we can study income distribution and build inequality indicators at the municipal level for the period 2002-2017. Over time, a socio-economic profile

of the individual residing in the Municipality of Bologna emerges, which represents an important informative element to improve the effectiveness and equity of welfare policies, and to guide the distribution of economic and social support, and urban redevelopment interventions in the different areas of the Municipality.

Analysing income inequality and seeing how it has changed over the years is of particular importance since the national average data can hide very different situations at the territorial level. This is particularly true for Italy where, for many economic indicators, the territorial gaps are much wider than those recorded in comparison with other countries (Ciani and Torrini, 2019).

One of the main findings show that, in real terms, the average rate of per capita income, displayed that the sixteen years taken into consideration can on the whole be considered as a period of substantial stagnation. However, there are important hidden changes in particular subsets of the population. First, gender difference among taxpayers decreased between 2002 and 2017. However, the gap at the end of the period was still present to the detriment of women, for whom we find average and median income levels lower than those of men. Secondly, during the sixteen years of the observed period, the aging of the Bolognese population was accompanied by a reallocation of the taxable income decidedly favourable to those over 55 years old. On the other hand, the situation of the youngest clearly worsened: their income saw in fact a significant negative variation, and the reduction in the average and median income affected both genders. Perhaps, the most worrying aspect of the trend, which is very strong in this re-composition, is that it persists even once the data are cleaned off the pure demographic effect. The worsening of

the relative economic condition of younger individuals should be a source of attention especially in a phase of moderate growth, such as the one experienced by the Municipality of Bologna.

The rest of the Chapter is organized as follows. Section 2 give an overview of some related literature, in particular it reviews works published by the Ministry of Economy and Finance on tax returns at the national level together with works carried out by other Italian municipalities which analysed income distribution trends at the local level by applying record-linkage techniques to administrative data of fiscal origin. After some methodological issues needed for the subsequent analyses assessed in Section 3, in Section 4 a detailed description of the income distribution is presented. Section 5 presents a list of indices of inequality, their characteristics and their variation over the period 2002-2017. Lastly, Section 6 analyses tax levy trends and focuses on the role played by the Italian personal income tax an its surtaxes in the redistribution of income at a local level.

3.2   Related literature

Since the beginning of 2000s, tax records have been used in the international context to study income distribution within and across countries, to explain the redistributive capacity of a national tax system, as well as to compare income inequality, in particular by looking at the evolution of the top incomes shares[1].

In Italy, in the past two decades, several papers that exploits tax administrative data have been published[2]. These works, among the

---

[1]Among the others, see Kopczuk and Saez, 2004; Bonhomme and Hospido, 2013; Atkinson et al., 2016.

[2]Among the others, see Di Nicola and Pellegrino, 2017; Barbetta et al., 2018; Di Caro, 2018.

others, employed fiscal data to study the redistributive effect of the personal income tax in the Italian context and how this have changed during the years. Indeed, the acceleration of the statistical validation procedures by the Ministry of Economy and Finance, initiated by the tax administration in 2015, has made available data on tax returns every year a few months after the deadline for the submission of the declaration to the Revenue Agency (Ministero dell'Economia e delle Finanze, 2017).

The tax data provided by the Ministry of Economy and Finance makes a general assessment about the economic situation of the Italian taxpayers. In 2017, 41.2 million individuals fulfilled the reporting obligation. While the total number of taxpayers increased by 340,000 individuals (+ 0.83 per cent) compared to the previous year, the total income dropped to 838 billion Euro as compared to the previous year (-0.6 per cent). The 84 per cent were incomes from employment and retirement benefits. The average income was 20.670 Euro for the 2017 tax year. Analysis of per capita income also showed that the distance between the North and the South is increasing. At the regional level, in the upper part of the ranking, there were no major changes compared to the previous year: the region with the highest overall average income was Lombardia (24,720 Euro), followed by the autonomous Province of Bolzano (23,850 Euro) and Emilia-Romagna (22,870 Euro), whereas Calabria had the lowest average income (14,120 Euro). Despite the decline of total income, the amount of the net Irpef increased of 0.9 percent over the previous year. It was equal to 5,140 Euro on average and it was paid by 30.7 million individuals, equal to nearly 75 per cent of the total taxpayers. Over 10.5 million subjects had a null net tax. These

were mainly taxpayers with income levels included in the exemption thresholds, or those whose gross tax was zeroed as a result of the deductions. The regional surtax of Irpef remained unchanged from the previous tax year, and in 2017 it amounted to about 11.9 billion Euro. The average regional surtax was 410 Euro. The highest value was recorded in Lazio (610 Euro), whereas the lowest value was found in Basilicata (270 Euro). On the other hand, the municipal surtax amounted to a total of 4.8 billion Euro compared to 2016 (+ 0.8 per cent), with an average amount of 190 Euro, ranging from the maximum value of 250 Euro in Lazio, to the minimum value of 60 Euro in the autonomous Province of Bolzano.

As already pointed out, the analysis of income distribution through sample data is reliable up to a level of regional granularity. Indeed, the sample surveys typically used for the analysis of income distribution, i.e. the Survey on Households and Living Conditions (EU-SILC) and Household Finance and Consumption Network (HFCN), do not make possible to carry out sufficiently fine partitions of the territory. More detailed analyses of income distribution trends at sub-national level are possible only thanks to the exploitation of tax data. An in-depth analysis of the geography of inequality in Italy was conducted for the first time by Acciari and Mocetti (2013). They provided new insights about spatial inequality in Italy exploiting the tax data at the provincial level for the period 2000-2011.

In this context, it also worth mentioning the work of the metropolitan City of Rome (Carrozzi and Rauco, 2019), who has recently launched the annual publication of statistics on income distribution, which includes a comparison of 10 metropolitan cities in the national

context[3]. Evidences in this work shows that in the ten metropolitan cities as a whole there were 12,399,896 taxpayers in the fiscal year 2017 (representing 67.1 per cent of residents of the ten municipalities, and 30.1 per cent of the Italian taxpayers). They produced a total taxable income of approximately 270 billion Euro, corresponding to 33.6 per cent of total national taxable income. The average income was of 21,799 Euro, which was higher than that computed at a national level. The number of taxpayers registered an average increase equal to 0.77 per cent as compared to the previous year, roughly equal to the increase registered in the entire country. On the other hand, the taxable income calculated for the ten metropolitan cities experienced an increase equal to 0.78 per cent, in contrast to the national trend which was slightly down. Evidences also shows that Rome ranked first for the number of taxpayers, proving to be the first city in Italy also in terms of total taxable amount produced. Bologna ranked sixth in terms of the number of taxpayers (775,666) and fifth in terms of declared taxable income (18 billion Euro). In this context, Bologna held the record, together with Genova, for the degree of participation of residents at local Irpef: over 76 per cent among residents are taxpayers. Besides, Bologna also stood out for the value of the average individual income (23,414 Euro in 2017), ranking second, preceded only by Milano (26,737 Euro).

Recently, another step was done on tax data front, i.e. the record-linkage of tax data with data from the municipal Registry Offices, with the aim to conduct more specific income distribution analyses on even finer partitions of the territory: municipality and its neighbourhoods. One of the first work worth mentioning in this context is

---

[3]Bari, Bologna, Firenze, Genova, Milano, Napoli, Reggio Calabria, Roma, Torino, Venezia

the AMeRIcA (Anagrafe Milanese e Redditi Individuali con Archivi) project (Comune di Milano, 2005), which constitutes an important contribution offered by the Municipality of Milano to the integration, at a local area, of administrative sources with an economic content. This project was born and conducted within a collaboration between the Statistics Sector of the Municipality of Milano and the Statistics Department of the University of Milano-Bicocca for the period 2000-2004. It aimed to build a statistical information system through periodic acquisition and confluence of a plurality of administrative sources in a single database, in order to return descriptive information on Milanese individuals and households from an economic and social point of view.

Another important noteworthy work in this context was conducted by the Municipality of Brescia together with the University of Brescia (Palamenghi et al., 2005). In 2002, they performed a study for the reconstruction of household income in the Municipality of Brescia in the year 2000, using a plurality of sources, mainly administrative or, in the absence of these, of a statistical type. The aim was to respond to the information needs of local decision-makers for the service plan and municipal taxation assessment. This work, as the AMeRIcA project, is substantially based on the combination of data from the municipal registry and the Revenue Agency's Tax Registry Interchange System of local authorities (SIATEL)[4]. However, this project differs from that of the Municipality of Milano because of different methodologies used. First of all, in this experience not all the resident families registered at the municipal registry were considered, but a sample of families was extracted from the registry of

---

[4]Each local authority has access to this registry in order to carry out its functions in accordance with L. 31/12/1996, n. 675 (with subsequent modification and integration).

the resident population as of December 31, 2000, for which, at a later stage, individual information on declared income was retrieved from the tax archives.

In recent years, the tendency to build new sources of integrated data to study the income trends at a municipal and sub-municipal level has grown, and an increasing number of municipalities in Italy have started to exploit administrative micro-data derived from tax returns[5].

Another very recent noteworthy innovation concerns the publication by the Ministry of Economy and Finance of detailed information on taxable incomes with a distinction by ZIP codes. This level of analysis does not allow to accurately distinguish incomes between the various neighbourhoods, which almost always have a different perimeter than the ZIP codes, but it is still useful for drawing a more detailed income map than the simple municipal area.

In the Municipality of Bologna there are 19 zip codes. Evidences show that for 2019 tax year (the first available) the areas with the highest number of taxpayers and the highest amount of declared income were those related to the zip codes number 40139 and 40133. They are located in the east and west part of the municipality, and they are afferent to the administrative areas Mazzini and Santa Viola, respectively. As for the average income of the nineteen zip codes, the richest areas turned out to be those with a zip code equal to 40136 and 40125, which correspond partially to the administrative area of Colli and the historic centre, respectively.

---

[5]To our knowledge, the municipalities who conducted this study are Bologna, Brescia, Firenze, Genova, Modena, Trento and Trieste.

3.3   Methodological issues

This Section has the aim to set some additional information required to better understand the subsequent analyses presented in this Chapter.

The first one relates to the use of net taxable income rather than the gross one. This choice is made since the net taxable income is more expressive of the effective monetary availability of resources for individuals and households. The net income is obtained as the difference between the gross taxable income and the personal income tax. Considering that the total taxable income reported in the database are pre-tax, but net of the taxpayers' social security contributions, in order to compute the net disposable income, the difference between the gross income and the net taxes paid are made.

Another clarification needed is related to the choice of the unit of analysis. The following investigations are conducted both with respect to the individual and the family dimension of income. In fact, both offer important insights. The analysis of household income is useful as the well-being of each subject is strongly correlated to the economic condition of the family unit to which they belong, since in this context the economic-financial resources are pooled and decisions are made regarding their destination, determining the actual standard of living of its members.

In this circumstance, first of all it is important to underline what is meant by family. Since the data processed was obtained from the municipal registry office, the definition adopted is the one contained in the Registry Regulation of the Resident Population. According to this rule, the family is defined as a "group of people linked by ties of marriage, kinship, affinity, adoption, protection or affection,

cohabitants and having habitual residence in the same municipality (even if they are not yet registered in the Registry Office of the resident population of the same municipality)"[6].

Given the characteristics of families, such as type and size, the same amount of disposable income will determine different living standards for families with distinct number of members, all other things being equal. Therefore, comparing the welfare levels of two families on the basis of their respective total incomes is not properly correct. To take into account the size of the family, one could refer to the per capita family income, obtained by dividing the total income of each family by the respective number of members. However, this solution does not take into account the fact that the contribution of individual members to the overall needs of the family varies considerably due to socio-demographic characteristics, and living together involves significant economies of scale. In fact, the consumption of certain goods and service by any family member does not necessarily reduce the availability for another member, such as for residence and durable goods, for example bicycles and cars, which can be shared by several members of the family at the same or at different time. Therefore, the cost to have the same level of utility does not increase in direct proportion to the number of people of the family, so the per capita measure tends to underestimate the well-being of larger families than that of smaller ones[7]. For this reason, in order to make homogeneous comparisons of well-being between heterogeneous units (in terms of income, number of members, etc.), an appropriate corrective is used: the equivalence scale. We use the modified OECD

---

[6]D.L. 24 December 1954, n. 1228; D.P.R. 30 May 1989, n. 223.
[7]For further details, see Baldini and Toso (2009)

scale[8]. The equivalent income calculated in this way can then be attributed to the family in question or to each member of that family, depending on whether it is conducted an analysis at a family or individual level. In our analyses the equivalent household income with the modified OECD scale is attributed to each member of the household.

3.4   Income trends of the residents

In this Section, income distribution trends of the taxpayers residing in the Municipality of Bologna are analysed in the period 2002-2017. One of the main result of the analysis shows that the annual growth rate of average real income is approximately equal to zero. Therefore, we are analysing a period in which, similarly to what happened in the rest of the country, income dynamics were essentially stagnant. However, this element foreshadows the possibility of the emergence of distributive tensions, and an increase in the risk of poverty for some socio-demographic figures.

Evidence shows that taxpayers who fulfilled the obligation to submit the tax income return in 2017 were about 280 thousands, slightly down (-3.8 per cent) compared to the first year of the period (see Table 3.1). Overall, resident taxpayers remained approximately 93 per cent of the total number of taxpayers of the Municipality for the entire period examined. The remainder is made up of those who are only fiscally domiciled in the Municipality, for whom we do not have detailed information, so, as already pointed out, they have been excluded from the analysis. At the beginning, the total number

---

[8]The OECD scale assigns a value of 1 to the household head, 0.5 to each additional adult member, and 0.3 to each child (under the age of 14).

Table 3.1: Summary data related to taxpayers resident in the Municipality of Bologna in the tax years 2002 and 2017 (values of income in Euro at constant 2017 prices).

	2002	2017	Variation (%)
No. of taxpayers	288,925	278,071	-3.8
of which with income greater than zero	268,925	240,407	-10.6
of which with net tax paid	252,162	227,164	-9.9
of which with net surtax paid	249,167	193,613	-22.3
Total taxable income (billion)	7.54	7.06	-6.4
Total Irpef (billion)	1.78	1.65	-7.3
Total municipal surtax (million)	29.7	53.6	80.5
Average taxable income	28,021	29,364	4.8
Median taxable income	20,850	22,364	7.3
Net average tax	7,070	7,242	2.4
Net average municipal surtax	119	277	131.9

Source: Data processing on MEF data.

of taxpayers followed a negative trend that seemed to stop between 2007 and 2008, when a slightly increasing value was recorded. After several years of contraction, a slight recovery began in 2016 and continued in 2017.

On the other hand, when we do not consider those who have a null taxable income, the total number of taxpayers underwent a greater negative variation over the entire period: -10.6 compare to -3.8 per cent (see Table 3.1). In 2002, those reporting positive income accounted for 93 per cent of the total resident taxpayers. Then, this percentage dropped until reaching the absolute minimum value in 2017, when residents with a positive taxable income corresponded to 86 per cent of the total.

Consequently, the percentage of those who paid a positive tax also fell sharply. Among these, we distinguish between those who have a positive Irpef from those who have a positive municipal surtax,

which overall underwent different percentage changes (-9.9 per cent for the first, against -22.3 per cent for the second). Between 2002 and 2006 these percentages had a very similar trend: in 2002, 87 per cent of total taxpayers owed a positive Irpef, while 86 per cent had to pay the municipal surcharge. Between 2006 and 2007, there was a significant decrease among taxpayers with a positive municipal surcharge: the percentage fell from 85 per cent to 72 per cent. This drastic decline was due to the introduction of a new tax policy at the municipal level, which provided for the exemption from the payment of the municipal additional income tax for those who had a taxable income less than 12,000 Euro. At the end of the period, those who did not have a positive Irpef were 18 per cent of the total, against 30 per cent which did not pay the municipal income surtax. Both the total taxable income declared and Irpef recorded very similar percentage changes between one year and the next in the period examined. In 2002, Bolognese taxpayers declared 7.54 billion of Euro in real prices, whereas in 2017, total income amounted to over 7 billion (-6.4 per cent). The total net tax, calculated in real prices, went from 1.78 billion of Euro in 2002 to 1.65 billion of Euro in 2017 (see Table 3.1).

The total amount of the municipal surtax recorded a clearly different variation, which highlight an increase in the weight of local taxation during the period 2002-2017. In real prices, it went from 30 million Euro in 2002 to 54 million Euro in 2017 (see Table 3.1), recording an increase of 80.5 per cent in real values, despite the reduction in the number of those who owed the municipal surcharge. In particular, the municipal income surtax recorded two significant increases: the first between 2006 and 2007, and the second between 2014 and 2015.

These were due to changes in tax legislation introduced. In 2007, the single rate underwent a change from 0.4 per cent to 0.7 per cent, concurrently with the introduction of the range of exempt taxpayers. In 2015, the single rate was raised to 0.8 per cent, keeping the exempt threshold unchanged at 12,000 Euro. Overall, the weight of local taxation underwent a change from 1.6 per cent in 2002 to 3.2 per cent in 2017 over the national one, which indicates a still rather low weight of the local taxation.

Both the average and median income show a positive variation, especially if considered in the light of the relative length of the time horizon considered. The average national tax paid also grew, even if to a lesser extent than income. The increase in the average municipal surtax was instead more intense. However, it started from a significantly lower value. In real prices, the average income of taxpayers showed a total positive change of 4.8 per cent in the 2002-2017 period, moving from 28,021 Euro in 2002 to 29,364 Euro in 2017 (see Table 3.1). On the other hand, the variation in median income was slightly more consistent, going from 20,850 Euro in 2002 to 22,364 Euro in 2017 (+7.3 per cent). At the beginning of the period, the trend of average and median taxable income was similar: from 2002 to 2010 both incomes underwent a modest positive change. Subsequently, both registered a contraction, greater for the average income rather than for the median, probably due to the effects of the double recession of Italian economy. In the last period, both recorded positive, albeit rather limited, growth.

3.4.1   Gender and age

In order to make a more in-depth analysis of the distribution trends, two important socio-demographic characteristics are analysed: gender and age. Evidence shows that, although the gender gap was decreasing, in 2017 it was still significant. The presence of a not negligible reduction in the average and median income of young taxpayers of both gender is also noteworthy. Between 2002 and 2017, overall women experienced an increase in average and median incomes far greater than that of men. The share of women who declared medium-high incomes was growing, but at the same time those who declared very high incomes were confirmed as less than 1 per cent of the total. Specifically, female taxpayers aged over 55 recorded positive variations in median income between 2002 and 2017, whilst women belonging to the younger classes did not enjoy this income improvement. On the contrary, they recorded larger negative changes in median income than those recorded by the male peers.

 The share of women taxpayers fluctuated around 53 per cent for the entire period, while that of men varied around 47 per cent. These percentages are in line with the demographic trend of the Bolognese population at the gender level, which is characterized by a slightly higher female percentage than the male one.

Despite the slightly higher percentage of taxpayers, the total taxable income declared by women is still lower than that of men. However, this difference decreased over the years: while in 2002 men reported incomes 51 per cent greater than women, in 2017 the percentage dropped to 38 per cent. If we observe the growth percentages of total taxable income in real prices, we note that it decrease of 10 per

Table 3.2: Summary data related to taxpayers resident in the Municipality of Bologna by gender in the tax years 2002 and 2017 (values of income in Euro at constant 2017 prices).

	2002	2017	Variation (%)
No. of taxpayers			
Women	152,206	145,521	-4.4
Men	136,719	132,550	-3.1
Total taxable income (billion)			
Women	3.00	2.97	-1.1
Men	4.54	4.09	-9.8
Total Irpef (billion)			
Women	0.60	0.61	0.5
Men	1.18	1.04	-11.9
Average taxable income			
Women	21,579	24,170	12.0
Men	34,905	34,778	-0.4
Median taxable income			
Women	17,949	20,203	12.6
Men	24,133	24,937	3.3

Source: Data processing on MEF data.

cent for men, and only 1 per cent for women (see Table 3.2). This can be considered a first sign of the gender gap reduction affecting the income of the Bolognese taxpayers.

With regard to average and median income, the male values were higher than the female ones in all years. It should be also stressed that the percentage difference between the two genders decreased over the years. In particular, women recorded an overall increase of more than 10 per cent both in the average and median income. As regards the male counterpart, there was instead a slight negative change in the average income and a slight positive change in the median one (see Table 3.2). Specifically, the analysis of the average and median income by gender reveals that the income difference

decreased slightly but steadily until 2010; from this year onward, it remained unchanged. In particular, the average male income was 62 per cent higher than that of women in 2002, and decreased until 44 per cent in 2017. The gender gap in median income was less pronounced: it was 34 per cent at the beginning of the period, while it was 23 per cent at the end.

The reason why the percentage difference was smaller in case of median income is due to the fact that among men there was a greater concentration of income in the upper classes. On the other hand, the greater growth of taxable income for women could be linked to an improvement in their wages and/or to an increasing presence in top positions, although this was still rather low compared to the male one, as can be seen from Table 3.3. Here, the number of Bolognese taxpayers is divided into eight taxable income classes, expressed in real values revalued in 2017 prices, and for each of them the frequency of taxpayers by gender is considered for the years 2002 and 2017.

From a general assessment, it emerges that in 2002 the largest income bracket was the one including incomes between 15,000 Euro and 26,000 Euro for both genders, where the share of taxpayers is just over 30 per cent. In 2002, women who declared positive taxable incomes of less than 26,000 Euro were 103,735, equivalent to 68.1 per cent, whereas men were 71,843, equivalent to 52.6 per cent. What happens in the following ranges is noteworthy. The percentage of men who declare an income between 26,000 Euro and 55,000 Euro exceeded the share of women by more than ten percentage points. As for the three richest groups, in both cases we observe that the number of taxpayers decreased as the threshold increases. Once
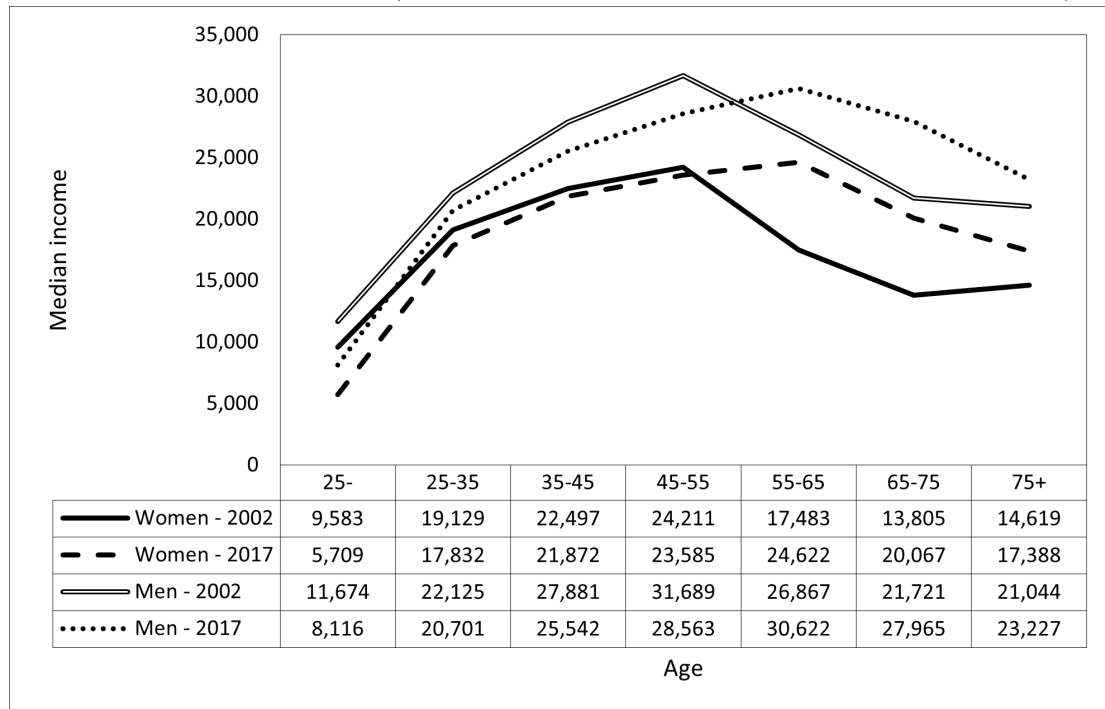
Table 3.3: Frequency distribution of taxpayers by taxable income range and gender for the tax years 2002 and 2017 (values of income at constant 2017 prices).

Income Range	Women		Men		Total	
	2002	2017	2002	2017	2002	2017
Less than 0	8.7	15.7	4.9	11.2	6.9	13.5
0-10,000	21.1	13.2	10.0	9.6	15.8	11.5
10,000 - 15,000	14.8	12.5	9.4	7.6	12.2	10.2
15,000 - 26,000	32.4	32.7	33.1	29.8	32.7	31.3
26,000 - 55,000	19.3	21.6	30.2	30.5	24.5	25.8
55,000 - 75,000	1.9	2.3	4.9	4.9	3.4	3.5
75,000 - 120,000	1.3	1.5	4.6	4.0	2.9	2.7
Greater than 120,000	0.6	0.6	2.8	2.4	1.6	1.4
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Data processing on MEF data.

again, however, the percentage of men was clearly higher than that of women: for income above 55,000 Euro, women were less than a third of men. In 2017, the situation was slightly different and, as we said, the gender gap smoothed out slightly. Women who declared positive incomes of less than 26,000 Euro are 84,897 (58.3 per cent). This value has decreased since the beginning of the period. The share of men is 47.0 per cent, which has also decreased from 2002. However, in 2017 the gap between the two genders was slightly smaller. Moreover, when we consider the three richest groups, the percentage of women increased, while that of men slightly decreased: 4.4 per cent of women declared incomes exceeding 55,000 Euro, whereas the share of men was equal to 11.2 per cent.

Analysing in detail the median taxable income of Bolognese taxpayers by gender and age groups, evidence shows that the female income was lower than male one for all age groups in all years (see Figure 3.1 for 2002 and 2017 values). The gap trend was similar

Figure 3.1: Median taxable income by gender and age groups for the tax years 2002 and 2017 (values of income in Euro at constant 2017 prices).
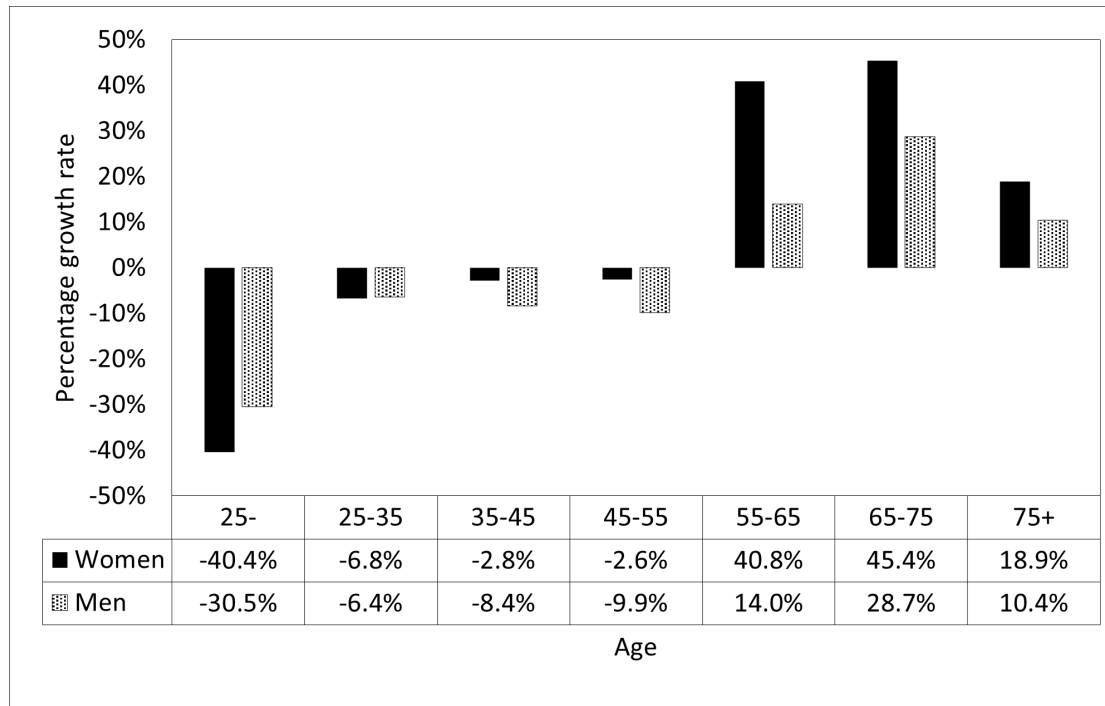

	25-	25-35	35-45	45-55	55-65	65-75	75+
Women - 2002	9,583	19,129	22,497	24,211	17,483	13,805	14,619
Women - 2017	5,709	17,832	21,872	23,585	24,622	20,067	17,388
Men - 2002	11,674	22,125	27,881	31,689	26,867	21,721	21,044
Men - 2017	8,116	20,701	25,542	28,563	30,622	27,965	23,227

Source: Data processing on MEF data.

in all the year: it was significantly lower in the younger classes; it grew as the age increase, and it decreased again in the oldest class. In 2002, the size of the gap was contained in the younger classes. Then, a marked increase is registered between 45 and 55 years old, in which the income difference was 30 per cent. For those between 65 and 75 years old, men had a median income 57 per cent higher than that of their peers. The gap smoothed out in the oldest class, dropping to 42 per cent. In 2017, the difference in median income between men and women was generally smaller, except for the very young class.

In order to compare the two-year values, Figure 3.2 depicts the growth rate of median real income by gender and age groups from 2002 to 2017. Although overall median income slightly increased

Figure 3.2: Growth rates of median taxable income in 2002 (revalued) and 2017 by gender and age groups.

	25-	25-35	35-45	45-55	55-65	65-75	75+
■ Women	-40.4%	-6.8%	-2.8%	-2.6%	40.8%	45.4%	18.9%
▨ Men	-30.5%	-6.4%	-8.4%	-9.9%	14.0%	28.7%	10.4%

Age

Source: Data processing on MEF data.

for both genders, the dynamics of taxable income was different in relation to the age group. Taxpayers over 55 showed markedly positive values. Moreover, among them women recorded a significantly higher growth rate than men. For younger taxpayers, the situation was definitely less favourable. It showed negative rates of change in all classes up to the age of 55. The most disadvantaged group was the under 25, with particular reference to women.

This data seems to confirm that, in the face of a substantial consistency in the dynamics of taxable income growth, measured on the whole population, we were witnessing a clear re-composition in the distribution, in favour of those aged over 55, which were basically retired or close to the retirement. Among these, the female component, which recorded the most significant improvement, was

probably the result of a generation that participated to the labour market more actively than previous generations and, therefore, enjoyed higher social security benefits in 2017 than in the past.

Moreover, an interesting question that we tried to answer in this work concerns the impact of the demographic change on the income shares held by taxpayers of different age. For this purpose, four broad age groups were considered: under 40 years old, from 40 to 60, from 60 to 75 and over 75. For all of them, the respective shares of income held in the tax years 2002 and 2017 were calculated (see Table 3.4). In 2002, the age group between 40 and 60 years old held

Table 3.4: Decomposition of the change in income shares owned by age groups between 2002 and 2017 into two components: demographic and income (percentage values).

| Age group	Share of total income			Effect	
	2002	2017	Demog.	Income	Total
Less than 40	25.1	18.1	-3.1	-4.0	-7.1
40-60	40.6	42.0	5.7	-4.3	1.4
60-75	23.1	24.3	-3.9	5.1	1.2
Greater than 75	11.2	15.2	2.5	1.5	4.0

Source: Data processing on MEF data.

the highest share of total income followed by the younger taxpayers, who owned 40.6 per cent and 25.1 per cent of the total income, respectively. Those aged between 60 and 75 years old were the 23.9 per cent of the total number of contributors and, at the same time, held the 23.1 per cent of the total income. At the bottom of the ranking, the elderly, who were numerically lower than the other classes, held just over 11 per cent of the total income. Between 2002 and 2017, there was a slight decrease in the number of younger taxpayers and those aged 60-75 years hold, in favour of the 40-60 years old group and the elderly. On the other hand, there was a substantial decrease

in the average income of the youngest in favour of the other classes. Table 3.4 shows the variation recorded in the income shares held by the four age groups between 2002 and 2017. These have been broken down into two components: the first is the demographic one, which takes into account the variation in the number of taxpayers over the total with the same average income; the second component is the income one, which instead takes into account the variation in average income level, with the same demographic condition. The sum of the two components is, by definition, equal to the overall change between the two years. The breakdown of Table 3.4 helps us to identify some important phenomena. For the youngest class, the negative change in the overall share of income held is the sum of two factors: the economic and demographic one, which have a substantially similar weight (-4.0 per cent and -3.1 per cent, respectively). In other words, young people were fewer in 2017 than in 2002, and their income fell compared to the average income of the population. The class of those aged between 40 and 60 recorded a slightly positive change in the share of income held. This is largely due to the increase in the share of taxpayers (5.7 per cent), while the average income suffered a decisively negative decline in the period examined (-4.3 per cent). Those aged between 60 and 75 years old also recorded a slightly positive change in the share of income owned. In this case, however, it was entirely due to the increase of average income (5.1 per cent), since demographically this segment of the population decreased compared to 2002 (-3.9 per cent). Finally, the older age group was the one that recorded the greatest increase in the share of income owned. This is due to a greater extent to the demographic change (2.5 per cent) and to a lesser extent to the

change in average income in the period considered (1.5 per cent).
Another interesting evidence useful to understand the evolution over
time of the average income of taxpayers by age is the *cohort analy-
sis*. It is an analysis of incomes on cohorts of taxpayers who, in this
case, share the year of birth (see Figure 3.3).

The cohorts are useful to compare how given quantities evolve over

Figure 3.3: Income trends for age cohorts in the period 2002-2017.
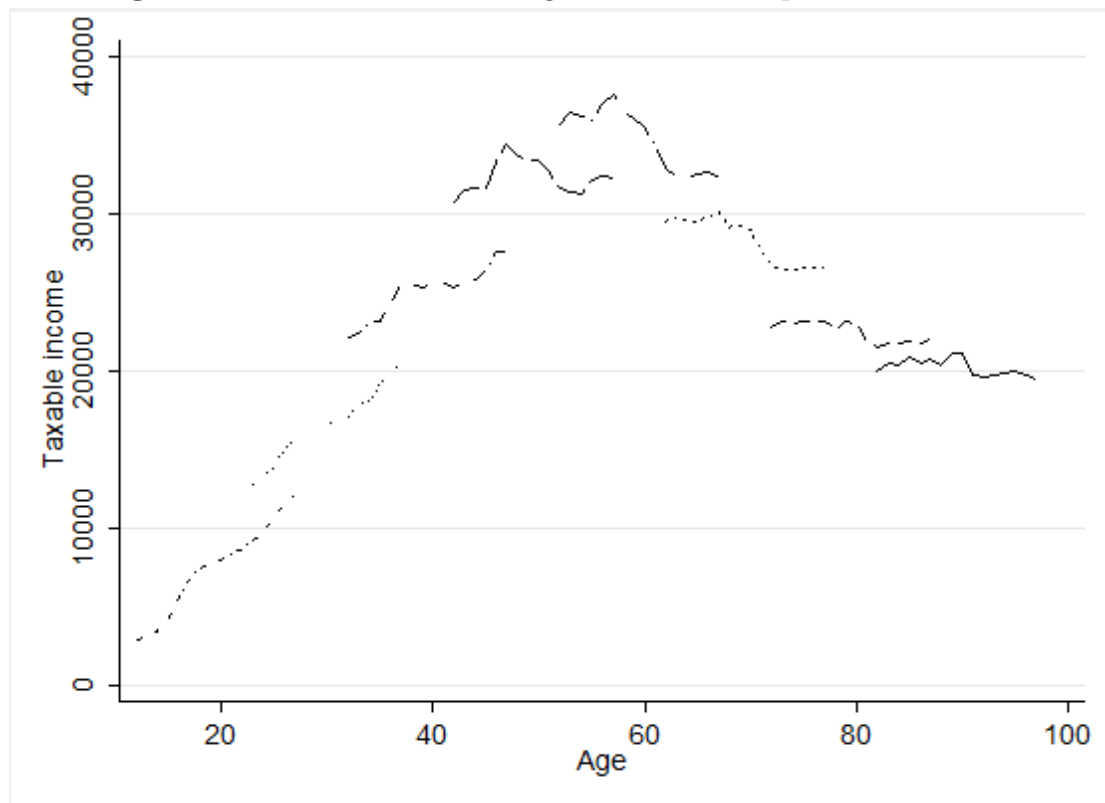
Source: Data processing on MEF data.

time in homogeneous by year of birth groups: in this case, how the
average taxable incomes of taxpayers evolve in the period between
2002 and 2017. The average income per cohort is on the ordinate
axis, while the age of the taxpayers is on the abscissa axis. The
dashed lines of the graph represent the income trends of eight dif-
ferent cohorts. For example, the first line on the left represents the

average income of taxpayers who were 12 in 2002 and therefore were 27 in 2017. A graph of this type allows to describe the evolution of income over their life cycle. Figure 3.3 displays the bell shape of the profile by age of income. It tends to increase up to an age close to the average retirement age and then decreases thereafter. Following over time (16 years) groups of individuals belonging to different cohorts (10 years length) allows another interesting exercise: the comparison between income levels of different cohorts at the same age. In this way, the data on which this analysis is built makes it possible to measure what, in the economic and demographic literature, are called *cohort effects.* Evidence shows that for the last four cohorts the sign of the effect is positive: in other words, the average income of the relatively younger cohorts is always, for the ages on which a comparison can be made, greater than that of the immediately older cohort. This result is consistent with a growing economy. Conversely, the sign of the cohort effect becomes negative for the cohorts that are on the left of the graph, i.e. the younger ones, born after 1956. In this case it is evident that being born after, on average, penalizes taxpayers compared to individuals immediately older, for the same age. Wanting to attempt a premature interpretation, this is a proof of the generational division taking place in the Italian economy and empirically documented both on sampling and administrative data also in other parts of the country.

3.4.2 Geography

In this sub-Section we analyse the distribution of income along the geographical dimension. A first analysis defines the income trend within the nine historic districts, while the ensuing more in-depth

analysis highlights the evolution of incomes in smaller portions: the eighteen administrative districts in force until the 1980s. Our results show that the situation across the neighbourhoods has been stable since the beginning of the period, and there have not seem to occur any notable changes. Nevertheless, the gap in the average and median level of taxable income between the northern and southern part of the municipality persist.

Evidence shows that the neighbourhood in which the greatest num-

Table 3.5: Number of residing taxpayers in the Municipality of Bologna by neighbourhood in tax years 2002 and 2017, and percentage change.

	No. of taxpayers		
Neighbourhood	**2002**	**2017**	**Variation (%)**
Borgo Panigale	19,183	19,030	-0.8
Navile	49,335	48,341	-2.1
Porto	24,729	23,930	-3.3
Reno	25,470	24,844	-2.5
San Donato	24,129	22,066	-9.3
San Vitale	34,622	33,908	-2.1
Santo Stefano	36,955	35,398	-4.4
Saragozza	27,661	26,472	-4.5
Savena	46,826	43,891	-6.7

Source: Data processing on MEF data.

ber taxpayers reside were Navile and Savena (see Table 3.5). These values remained constant throughout the all the period, although they both experienced a decrease in the absolute number of taxpayers. On the other hand, the least populated neighbourhood was Borgo Panigale both in 2002 and in 2017. All the nine districts registered a decline in the number of taxpayers: the smallest affected Borgo Panigale, whereas the largest occured in San Donato, which recorded a decrease of 9 per cent.

San Donato was also the neighbourhood in which the taxpayers

Table 3.6: Ratio between average/median taxable income to highest average/median taxable income by neighbourhood for the tax years 2002 and 2017 (values of income at constant 2017 prices).

	Average income		Median income	
Neighbourhood	**2002**	**2017**	**2002**	**2017**
Borgo Panigale	54.6	55.6	76.6	79.1
Navile	55.7	55.6	78.5	77.8
Porto	72.6	72.2	88.7	89.6
Reno	56.7	58.7	80.0	81.2
San Donato	51.7	52.7	74.9	75.4
San Vitale	71.9	71.7	86.0	86.3
Santo Stefano	100.0	100.0	100.0	100.0
Saragozza	83.6	83.8	91.7	91.8
Savena	64.5	64.3	84.0	84.1

Source: Data processing on MEF data.

owned the lowest median income for all years (see Table 3.6). On the other hand, at the top of the ranking, taxpayers residing in Santo Stefano were observed to have the highest median taxable income for all the period available. Concerning the average median income growth, Borgo Panigale registered the most sustained increase, followed by Reno, which have been merged into a single district since the last municipality revision[9]. At the bottom of the ranking we find Navile, whose income was higher than that of Borgo Panigale at the beginning of the period, but lower at the end. In general, at least observing median income, there does not seem to be any significant re-composition in the ranking of neighbourhoods between 2002 and 2017.

As regards average income, the ranking of the Bolognese nine districts appears unchanged at the end of the period available (see

---

[9]From 7 June 2016, the new administrative structure officially entered in force in the Municipality of Bologna led to a reduction of the districts from 9 to 6. Borgo Panigale was merged to Reno, so happened to Porto and Saragozza, and to San Donato and San Vitale.

Table 3.6): Santo Stefano remained at the top of the ranking, while San Donato held the lowest average income of the entire municipality.

In order to gain a more detailed picture of income trends, our dataset allows us to analyse income distribution of taxpayers at a even finer level of territorial granularity: the eighteen administrative areas which compose the Municipality of Bologna.

The thematic map is the solution commonly adopted to represent the territorial distribution of a phenomenon. It makes possible to examine the distribution of median taxable income, and to identify areas of greater or lesser criticality. Figure 3.4 displays the value of real median taxable income for each of the eighteen administrative areas in 2002. Evidence shows that the highest median incomes were concentrated in the historic centre and the areas in the southern part of the municipality. On the other hand, the north and west part, which had significantly lower values, were the most disadvantaged. Colli ranked first in 2002, with a median real taxable income of 28,509 Euro; on the other side of the ranking, San Donato registered the lowest value (18,650 Euro), equal to the 65.4 per cent of Colli. Figure 3.5 displays the values of real median taxable income for each of the eighteen administrative areas in 2017. The situation has remained unchanged with respect to the previous years: the area with the highest value recorded was Colli, with a value of 29,606 Euro (+3.8 per cent in real values since the beginning of the period). At the bottom of the ranking, we found San Donato with a median taxable income of 20,089, whilst a growth of 7.7 per cent since 2002.

Clearly, average income had a much higher value with respect to

Figure 3.4: Median real taxable income in the eighteen administrative areas of the Municipality of Bologna in the tax year 2002.
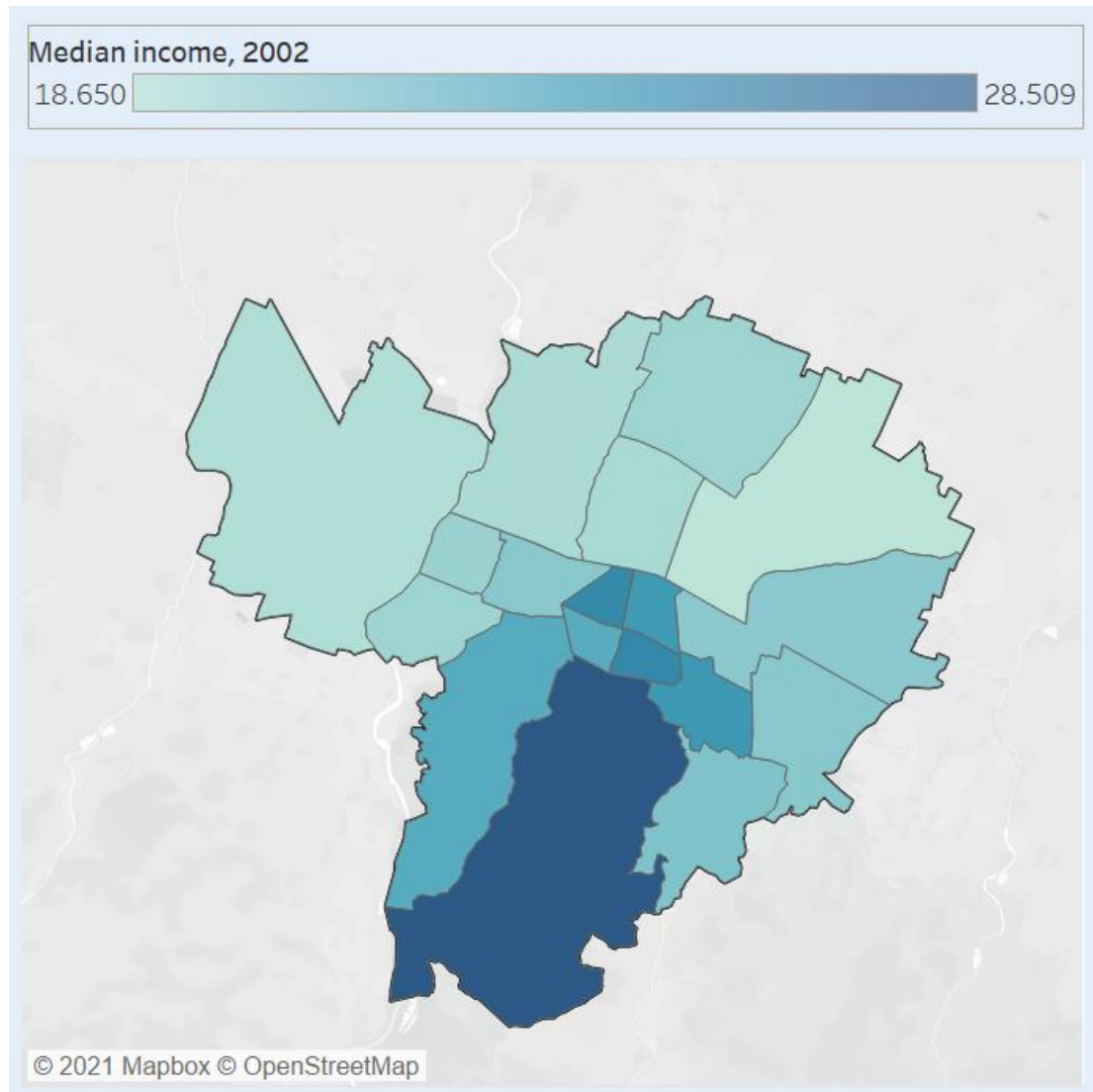
Source: Data processing on MEF data.

median one. Overall, top and bottom positions in the ranking did not differ from the observed situation. Colli was at the top of the ranking with an average income of more than 55 thousand Euro both in 2002 and 2017, and it also recorded a slight positive change over the years. At the bottom of the ranking, San Donato was confirmed to be the most critical area for the entire period available. In

Figure 3.5: Median real taxable income in the eighteen administrative areas of
the Municipality of Bologna in the tax year 2017.
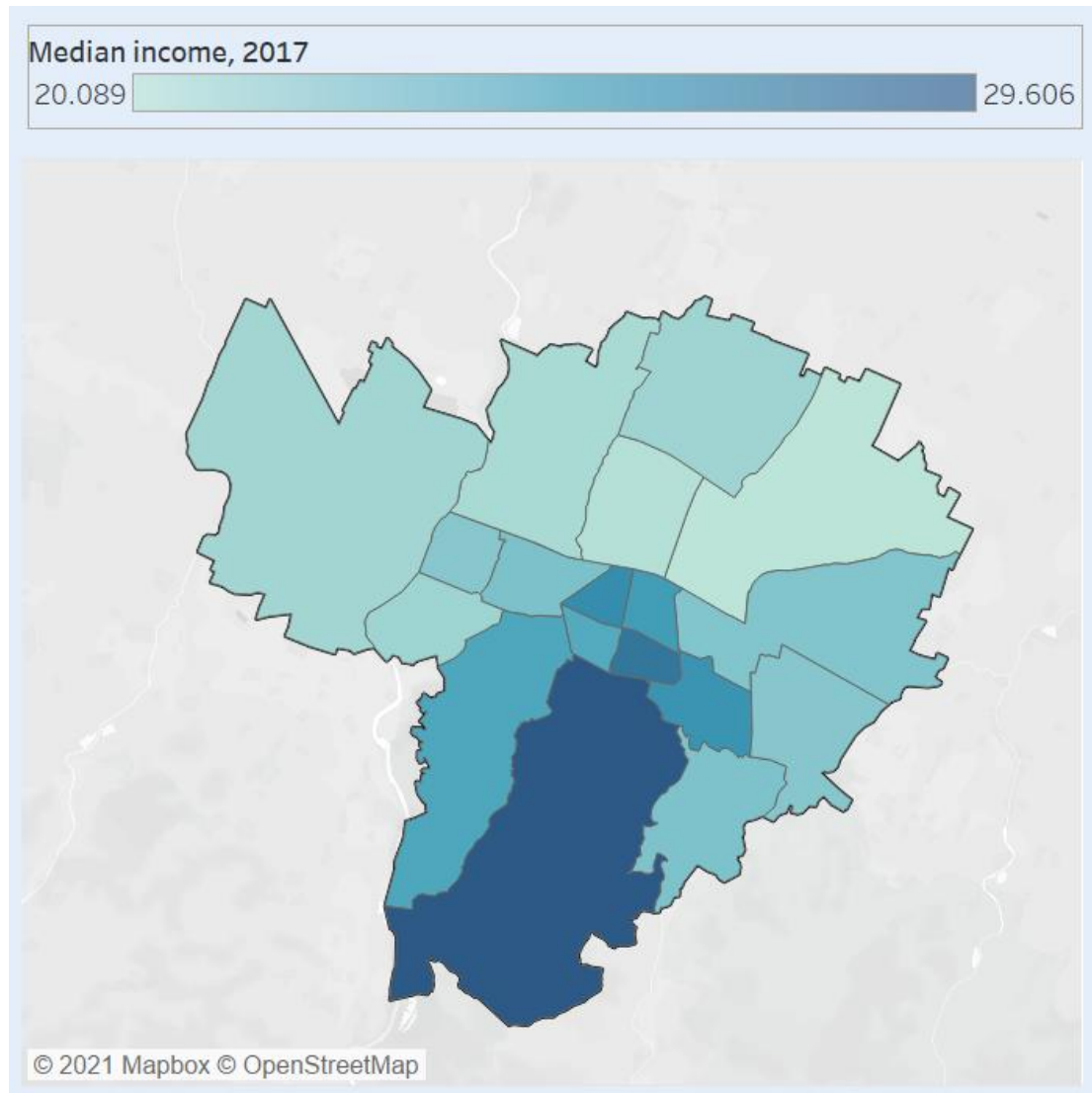
Source: Data processing on MEF data.

the central part of the ranking there were slight but not significant
changes.

3.4.3   Nationality

This sub-Section is aimed to study the income difference between
Italian and non-Italian taxpayers. Evidence shows that the number

of foreign taxpayers has more than tripled in the period examined, and, in real terms, their income modestly decreased since the beginning of the period. From an age point of view, the presence of foreign taxpayers was numerically greater in the younger classes, which in 2017, surprisingly, gained higher median income than the Italian counterpart. Moreover, as expected, the majority of foreign taxpayers resided in the poorest neighbourhoods.

The number of non Italian taxpayers grew significantly, especially

Table 3.7: Summary data related to taxpayers residing in the Municipality of Bologna by nationality in the tax years 2002 and 2017 (values of income in Euro at constant 2017 prices).

	2002	**2017**	**Variation** (%)
No. of taxpayers			
Italian	279,764	248,808	-11.1
Foreign	9,161	29,263	219.4
Total taxable income (billion)			
Italian	7.42	6.75	-9.0
Foreign	0.12	0.31	160.5
Total Irpef (billion)			
Italian	1.76	1.60	-9.1
Foreign	0.02	0.05	131.1
Average taxable income			
Italian	28,410	30,797	8.4
Foreign	15,093	14,589	-3.3
Median taxable income			
Italian	21,148	23,325	10.3
Foreign	12,583	12,585	0.0

Source: Data processing on MEF data.

in relation to the substantial migratory flows from Eastern Europe. The values has more than tripled between 2002 and 2017, so the incidence increased from 3.2 per cent to 10.5 per cent. Our results show that the foreign structure was very heterogeneous within the

municipality: the nationalities of taxpayers were more than a hundred. It is observed that Europeans underwent a significant increase between 2002 and 2017. It should be remembered that in this period Romania and Bulgaria became part of the European Union, which contributed to an increase in the share of European immigrants. Looking at the main nationalities, in 2002 Philippine was at the first place, whereas in 2017 the Romanian taxpayers were the most extended group.

At constant prices, the total taxable income declared by foreigners was 120 million Euro in 2002 and 310 million Euro in 2017, equivalent to 1.6 per cent and 4.4 per cent of the total taxable income declared for the respective years (see Table 3.7).

The total personal income tax paid by foreigner taxpayers was 20 million in 2002 and 45 million in 2017, respectively 1.1 per cent and 2.8 per cent of the total income tax paid by all taxpayers residing in the Municipality of Bologna. The share of tax rate paid by foreigners was lower than that of declared taxable income, which derives from the fact that the non-Italian taxpayers declare average incomes lower than the Italian counterpart.

The average taxable income, computed in real values, of foreigners was 15,093 Euro in 2002, while for the Italian counterpart it was 28,410 Euro, 46.9 per cent higher than the former. In the period between 2002 and 2017, average incomes underwent an overall variation of 8.4 per cent and -3.3 per cent for Italians and foreigners, respectively. In 2017, the average income of Italians was 30,797 Euro, while that of foreigners was 14,589 Euro, 52.6 per cent less than that of the former (see Table 3.7). Therefore, this variation led to a slight increase in the gap already existing between Italian and

non-Italian taxpayers.

As for median incomes, in 2002, the foreigners one was 12,583 Euro, which has remained stable until 2017. On the other hand, the Italian taxpayers median income underwent a positive variation (10.3 per cent): it was 21,148 Euro in 2002, and 23,325 Euro in 2017 (see Table 3.7). The gap has increased over the years: in 2017 it was 5.5 percentage points higher than in 2002 (40.5 per cent versus 46 per cent).

Faced with a slight increase in total average income in the period

Figure 3.6: Average real taxable income for the foreign citizens by year of arrival in the Municipality of Bologna at 2017 constant prices.
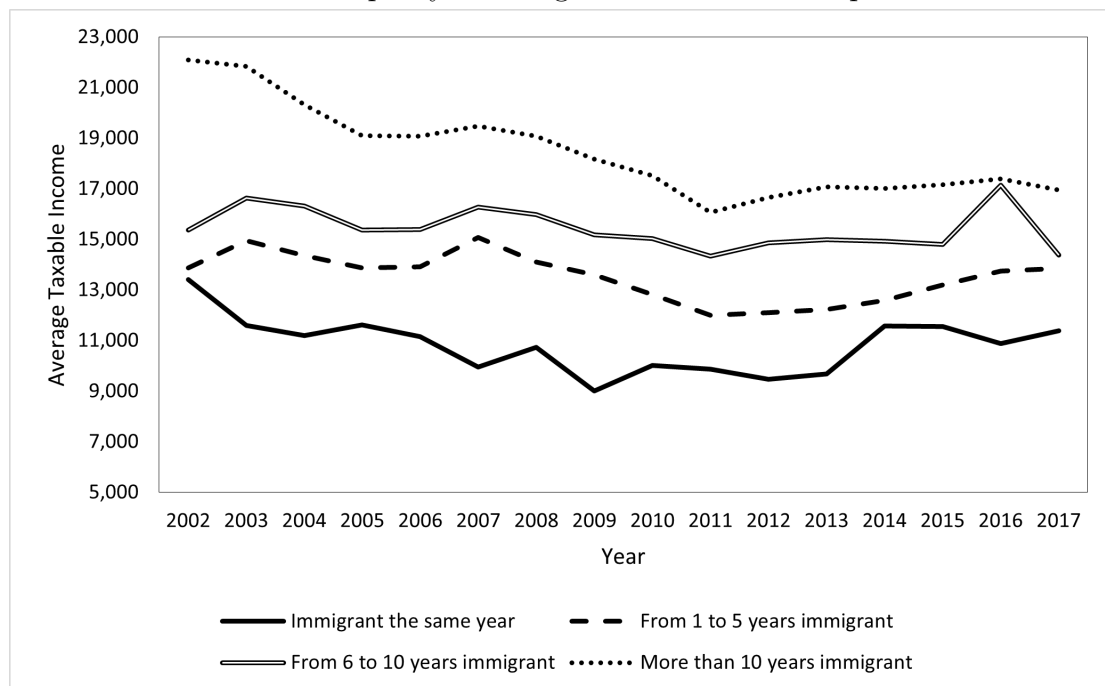
Source: Data processing on MEF data.

examined, the reason for the decrease in the average taxable income of foreign taxpayers between 2002 and 2017 could be ascribed to the influx of new immigrants with particularly low incomes. To investigate this point, an analysis on foreign taxpayers was conducted differentiating them by year of arrival in the Municipality of Bologna.

Firstly, it is observed that the average income of foreign taxpayers is directly proportional to the number of years in Italy: the longer the presence in Italy, the higher the average income of taxpayers. For example, the average income of newly arrived is equal to 13,404 Euros in 2002 and 11,381 Euros in 2017. On the other hand, the income of those who have been in the Municipality for more than 10 years is equal to 22,096 Euros in the 2002 and 16,935 Euro in 2017. Secondly, it emerges that the average income of taxpayers decreased substantially (-23 per cent) between 2002 and 2017 for the class of those who have been in Italy for more than ten years, followed by the class of those who have just arrived, whose income decreased by 15 per cent. On the other hand, the income decline for the other two classes, i.e. the taxpayers present in the Municipality from 1 to 9 years, is less influential, or even zero.
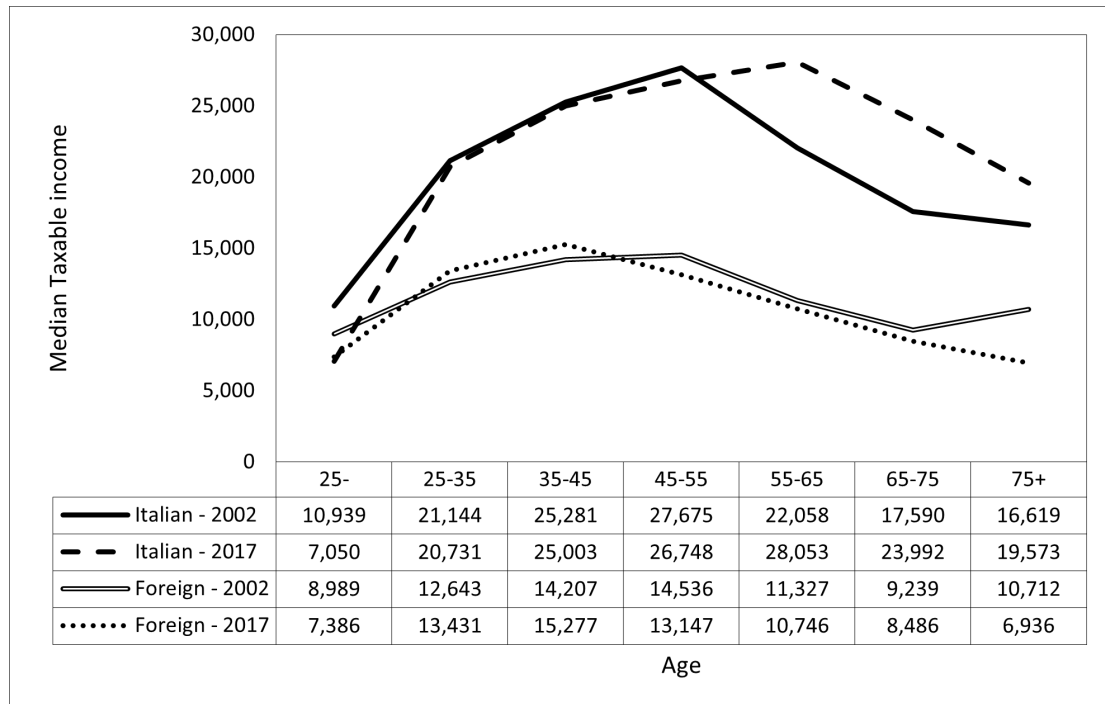
The analysis of the plate of foreign contributions at gender level shows that the number of male was greater than that of women, and it has remained constant between 2002 and 2017. Overall, how-ever, women grew numerically more than the male counterpart.

From the age point of view, evidence shows that in 2017 the average age of foreigners was 40.3 years old. Although there has been an aging compared to 2002 (where the average age was 35.4), foreign taxpayers were significantly younger than the Italians. Their average age in fact went from 54.2 in 2002 to 56.7 in 2017. For this reason, clearly, the presence of non-Italian taxpayers was numerically greater in the younger classes. The trend has remained constant over the period: the greatest incidence of foreign taxpayers was among under 25 years old, and the share of non-Italians constantly decreases as the age increases up to the over 75 age group, whose share was equal

to 0.2 per cent.

Foreigners median income was lower than that of Italians for all

Figure 3.7: Median real taxable income by citizenship and age groups for the tax years 2002 and 2017 (values in Euro at constant 2017 prices).
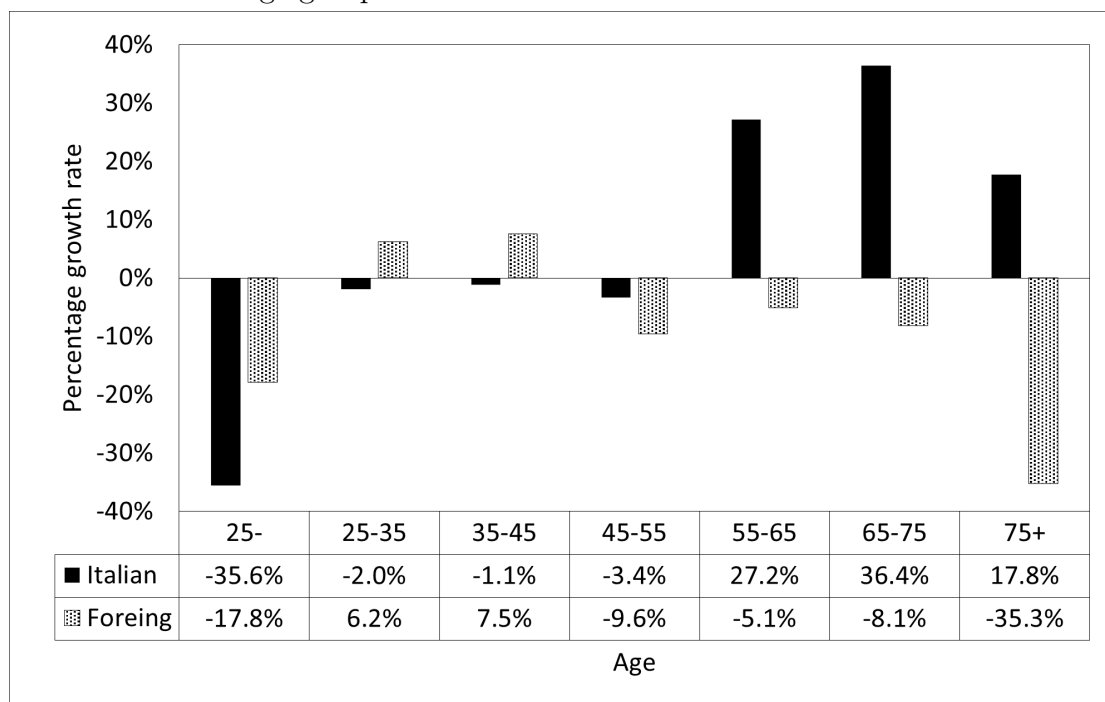

	25-	25-35	35-45	45-55	55-65	65-75	75+
Italian - 2002	10,939	21,144	25,281	27,675	22,058	17,590	16,619
Italian - 2017	7,050	20,731	25,003	26,748	28,053	23,992	19,573
Foreign - 2002	8,989	12,643	14,207	14,536	11,327	9,239	10,712
Foreign - 2017	7,386	13,431	15,277	13,147	10,746	8,486	6,936

Source: Data processing on MEF data.

the period considered (see Figure 3.7). The gap, initially limited, grows with increasing age until it reaches its maximum extent in the class between 45 and 55 years old. In this age group, in 2002, both foreigners and Italians reached the maximum value of the real median income, but foreigners declared about half that of the Italian peers. In 2017, the dynamic was different: for non-Italian taxpayers the maximum value is reached in the class between 35 and 45 years old, while Italians reached the maximum median income between 55 and 65 years old. Therefore, the older Italians age groups registered a greater improvement in income than the younger taxpayers, whereas, among the foreigners, the younger saw a slighter improve-

ment in the median income than the older counterpart. However, it is surprisingly how in 2017 the median income of foreign taxpayers under 25 is, albeit by a few hundred Euro, higher than that of their Italian counterpart.

Figure 3.8 shows the growth rate of median real income by age

Figure 3.8: Percentage growth rate of real taxable median income by citizenship and age groups between 2002 and 2017.

	25-	25-35	35-45	45-55	55-65	65-75	75+
■ Italian	-35.6%	-2.0%	-1.1%	-3.4%	27.2%	36.4%	17.8%
▨ Foreing	-17.8%	6.2%	7.5%	-9.6%	-5.1%	-8.1%	-35.3%

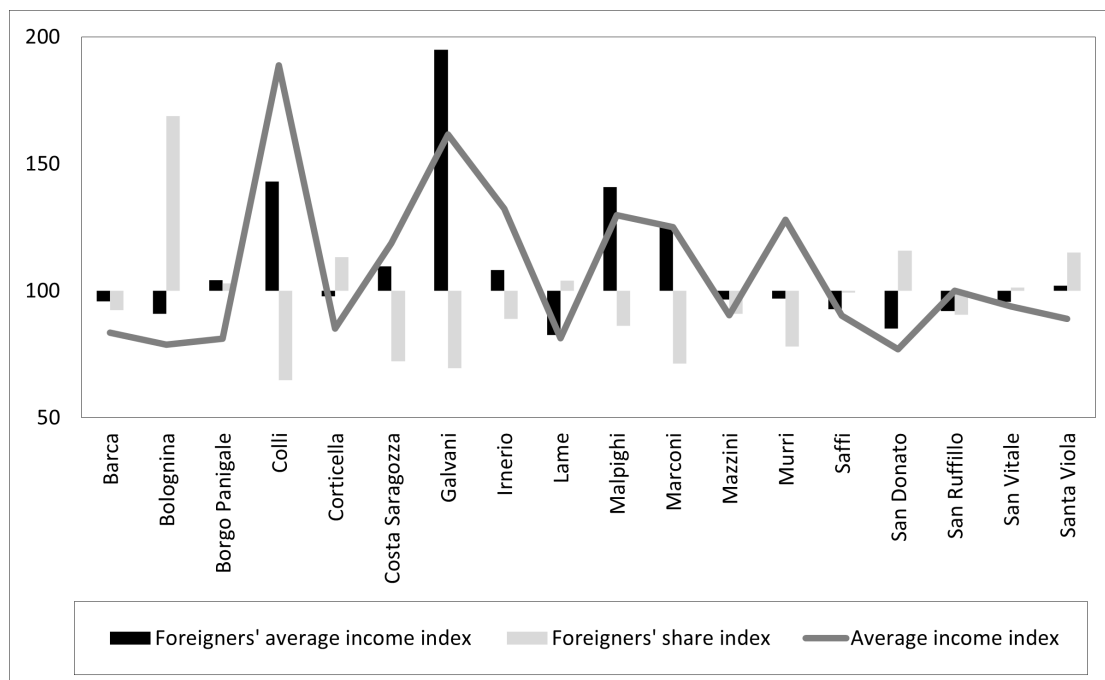*Source:* Data processing on MEF data.

groups and citizenship between 2002 and 2017. Evidence shows that the dynamic of taxable income was different in relation to age. Italian taxpayers over 55 registered markedly positive values, while the income situation of the foreigner counterpart worsened, in particular for over 75 years old. The situation was definitely less favourable for younger Italian taxpayers. The most disadvantaged age group was individuals under 25, in which the Italian taxpayers underwent a negative variation twice higher than that of their foreign peers.

These data highlights, once again, how the increase in real terms of taxable income mainly affected the older taxpayer group, thanks to the dynamics of social security treatments. On the other hand, for the youngest a state of not negligible difficulty emerged, certainly due to the growing difficulties of entering the labour market.

The following analysis highlights the territorial distribution of in-

Figure 3.9: Average income index, foreigners share index and foreigners average income index by administrative areas of the Municipality of Bologna in 2017.

Source: Data processing on MEF data.

comes by nationality. In 2002, the highest component of foreigner taxpayers was resident in Navile compared to all taxpayers residing in the district, followed by San Vitale and Porto. At the bottom of the ranking, there was Savena, with 2 per cent of foreigners with respect to the total. In 2017, the situation changed slightly: Navile still ranked first, with a percentage of foreigners tripled compared to 2002, followed by San Donato and Borgo Panigale, where the

percentages were almost four times higher. Santo Stefano, the richest neighbourhood, held the lower percentage of foreign taxpayers: nearly 8 per cent out of all non-Italian residents.

Figure 3.9 shows three different indices for each administrative area of the Municipality of Bologna in 2017. The average income index assumes values above or below 100 depending on whether the average income declared in the respective administrative area is greater or less than the average income recorded in the entire Municipality (29,364 Euro). The foreigners average income index refers only to non-Italian taxpayers and shows values higher or lower than 100 for each administrative area depending on whether their income is greater or lower than the average income of foreigners in the entire Municipality (14,589 Euro). Finally, the foreigners share index assumes values proportionally higher or lower than 100 in relation to the fact that more or less foreigners reside in the corresponding administrative area than the average percentage of residents in the Municipality (10.5 per cent). Evidence shows that the neighbourhoods in which the foreign average income values were the highest (i.e. Colli, Galvani, Malpighi and Marconi) also had a relatively low share of foreigners. On the other hand, neighbourhoods with a non-Italian average income below the municipality average had a greater share of foreigners taxpayers (i.e. Bolognina, Lame and San Donato).

Internal immigration

This sub-Section aims to analyse the taxpayers residing in the Municipality of Bologna by the geographical origin.

For each taxpayer, it is observed whether he has always resided

within the Municipality, or whether he has moved there from another place. Specifically, it is examined whether the transfer took place from another municipality of the Province of Bologna, from another Italian province, or from a foreign State. When the immigration comes from another Italian province, these have been divided between North, Centre, and South or Islands.

In 2002, about two thirds of the residents in the Municipality of

Table 3.8: Percentage composition and median real taxable income of taxpayers residing in the Municipality of Bologna by immigration area in 2002 and 2017 (values of income in Euro at constant 2017 prices).

	Percentage of taxpayers		Median income	
	2002	**2017**	**2002**	**2017**
Not immigrant	31.9	27.9	21,469	21,672
Immigrated from:				
Province of Bologna	32.7	27.9	18,331	19,762
North	17.6	16.1	20,620	21,956
Centre	4.6	5.1	22,797	22,917
South and Islands	11.1	15.9	19,448	21,428
Foreign State	2.1	7.2	10,206	7,992

Source: Data processing on MEF data.

Bologna always resided there, or were immigrants from another municipality within the Province of Bologna (see Table 3.8). In 2017, the percentage of these individuals dropped to 55.8 per cent. The two most dynamic components were individuals coming from a foreign state and those coming from South and Islands, which registered an increase in the period examined.

Regards median taxable income in real term, note that taxpayers coming from the Centre had the highest median income for all the fiscal years examined. They declared a median income of 22,797 Euro in 2002, and 22,917 in 2017 (see Table 3.8). All the median incomes had the same trend: they increased until 2007, slowed down

until 2012, and then it grew again in the last part of the period. At the second place of the ranking, in 2017 we find immigrants from the North of Italy, with a median income of 21,956 Euro. Non-immigrants were at the third place, declaring a median income of 21,672 Euro.

Our results showed that there is no significant difference between those who have always been resident in the Municipality of Bologna and those who have moved there during their life from the same Province, or from other parts of Italy.

3.4.4   Households

Cross-referencing between our dataset and data from the Registry Office of the Municipality of Bologna makes it possible to analyse fiscal data using the household as the unit of measure. This analysis is useful since the well-being of each subject is strongly correlated to the economic condition of the family unit to which he or she belongs. Within the family the economic and financial resources are pooled and decisions made regarding their destination, thus determining the actual standard of living of its members.

In general, the analysis at household level confirms the trends seen on individual data, however it makes possible to better qualify gainers and losers. Our results shows that the number of households increased over the years. Among these, single-member household substantially increased, if compared to the other categories. Besides, evidence shows that households composed of at most two components are those who saw their income increase the most. On the contrary, households with children and other cohabiting members registered a reduction in their household income over the years.

Since within the first group there are mainly retired individuals and in the second group individuals of younger age, this split allows us to better qualify the generational issue already evident in the analysis of individual data.

Evidence shows that the number of households increased of approx-

Table 3.9: Summary data related to households resident in the municipality of Bologna in the tax years 2002 and 2017 (values of income in Euro at constant 2017 prices).

	2002	**2017**	**Variation** (%)
No. of households	173,743	187,386	7.85
of which with income greater than zero	167,741	170,181	1.45
of which with net tax paid	162,540	164,699	1.33
of which with net surtax paid	161,694	147,105	-9.17
Average no. of components	2.04	1.92	-5.94

Source: Data processing on MEF data.

imately 8 per cent over the entire period examined (see Table 3.9). Compared to the total number of households registered in the Registry Office of the Municipality of Bologna, those declaring income decreased from 94 to 91 per cent between 2002 and 2017. This means that the number of households in which there is no income earner increased. When those who had a null taxable income are not considered, the number of households underwent a slightly positive variation over the entire period (+ 1.45 per cent). In 2002, households with a positive income were 97 per cent of the total. Numerically, this number increased until reaching the maximum value in 2011, when households with positive income amounted to 174,713 out of a total of 185,297. Since 2011, the absolute frequency of families declaring a positive income decreased by 4,532 units: in 2017, 91 per cent of total households had a positive taxable income. When we consider households that owed a positive Irpef, the total

number dropped even further. In 2002, they were 162,540 (94 per cent) (see Table 3.8). The percentage change over the entire period was positive but minimal: in 2017, families who owed an non-null Irpef were 88 per cent of the total.

The trend of households with a positive municipal income surtax was similar to that observed for individual taxpayers: between 2006 and 2007 there was a significant decrease due to the introduction of the range of exemption from the payment of a municipal tax. In 2002, 93 per cent of the total paid a municipal surcharge greater than zero, in 2017 this value dropped to 78 per cent.

Figure 3.10 shows the number of households by household type[10].

Figure 3.10: Number of households residing in the Municipality of Bologna by household type in tax years 2002 and 2017.
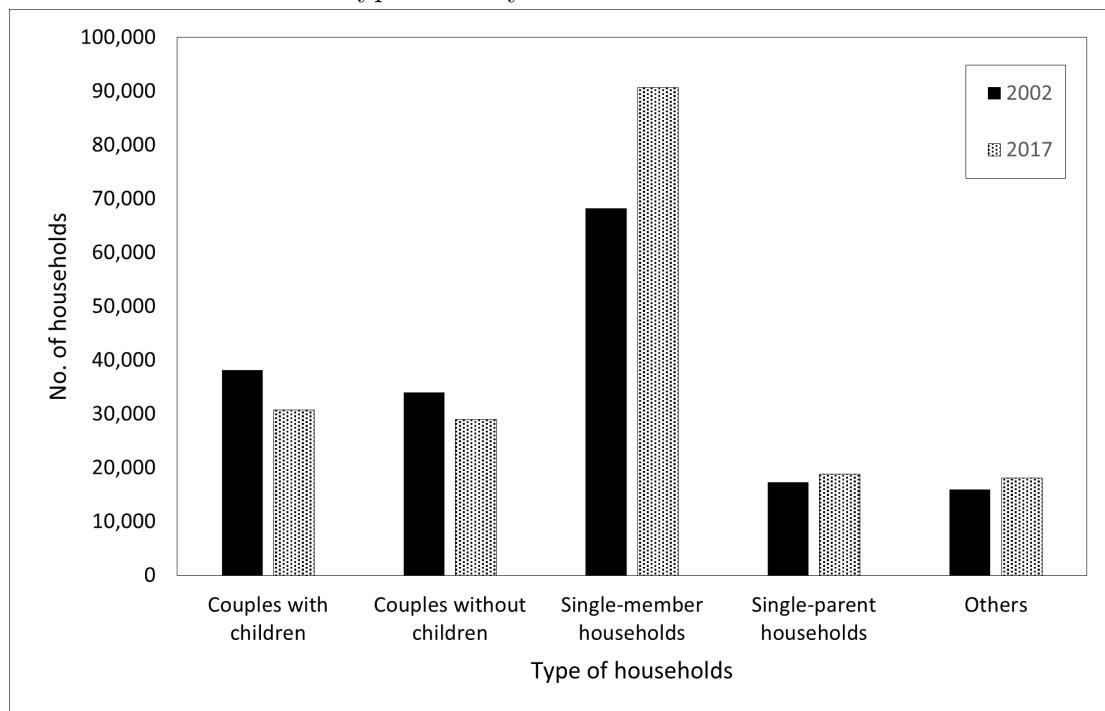
Source: Data processing on MEF data.

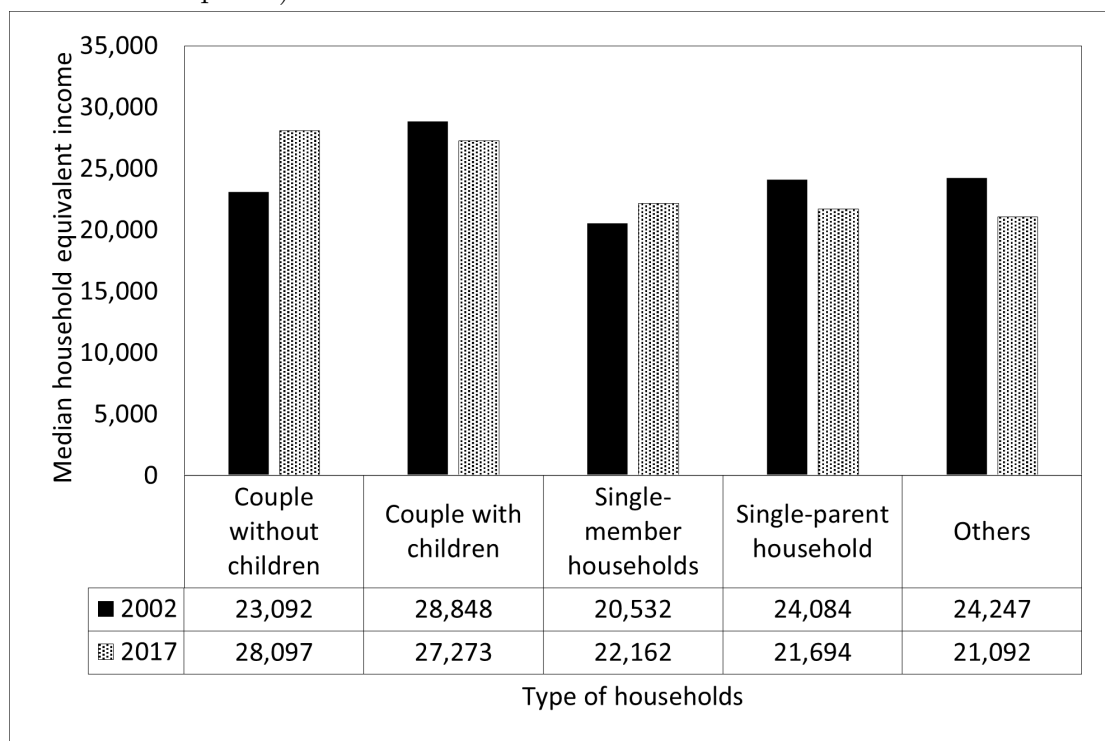In 2017, single-member households accounted for just under half of

---

[10]The category "Others" grouped all households in which, in addition to the couple with children, there are other cohabiting members

the total: their incidence increased from 39 to 48 percent between 2002 and 2017. They are followed by couples with children and couples without children (approximately 16 percent). They experienced a negative change of 25 percent and 21 percent, respectively. Single-parent households remained fairly stable since 2002 (10 percent of the total).

As already pointed out in Section 3.3, the equivalent household income computed using to the modified OECD scale, and attributed to each member of the household, is used to analyse income distribution at the household level.

Figure 3.11 shows the median household equivalent income by

Figure 3.11: Median taxable equivalent income (modified OECD scale) by family type in tax years 2002 and 2017 (values in Euro at constant 2017 prices).

	Couple without children	Couple with children	Single-member households	Single-parent household	Others
■ 2002	23,092	28,848	20,532	24,084	24,247
▨ 2017	28,097	27,273	22,162	21,694	21,092

Type of households

Source: Data processing on MEF data.

household type for the tax years 2002 and 2017. Our results show

that in 2002, couples with children raked first with an equivalent household median income of 28,848 Euro. "Others" category and single-parent households followed. Childless couples and single-member households appear to be the most disadvantaged at the beginning of the period. On the other hand, in 2017, the situation changed slightly: at the first place we find couples without children (28,097 Euro), which recorded a sharply growing income (+21.7 per cent). On the contrary, couples with children equivalent household income decreased since 2002 (-5.5 per cent).

This tendency confirms the generational issue already evident in the analysis conducted at individual level: the income growth registered by the couple without children are probably due to the fact that this type of household is mainly constituted by retired individuals, which has registered the greatest positive income variation. On the other hand, couple with children are mainly made up by individuals of younger age and, for this reason, they were more likely to experience a decline in median equivalent household income in the period examined.

3.5   Inequality of income

In this Section, using tax data, different indexes of inequality are constructed and discussed in relation to their advantages and limitations: the Gini index, percentiles and share ratio, Theil index, together with a detailed description of the top income earners are presented.

3.5.1   Gini coefficient

The most commonly used summary measure of economic inequality is the Gini coefficient. The main attraction of this index is its immediate geometric interpretation in terms of the Lorenz curve[11]. In fact, the Gini coefficient is defined as the area between the Lorenz curve and the 45-degree line, divided by the total area under the 45-degree line. This inequality index takes values between 0, which refers to perfect equality, and the maximum value of 1, in the case of maximum inequality. Therefore, the lower the Gini coefficient is, the more equal a society is.

Figure 3.12 shows the trends of the Gini index before and after

Figure 3.12: Gini index pre and after personal income tax from 2002 to 2017.
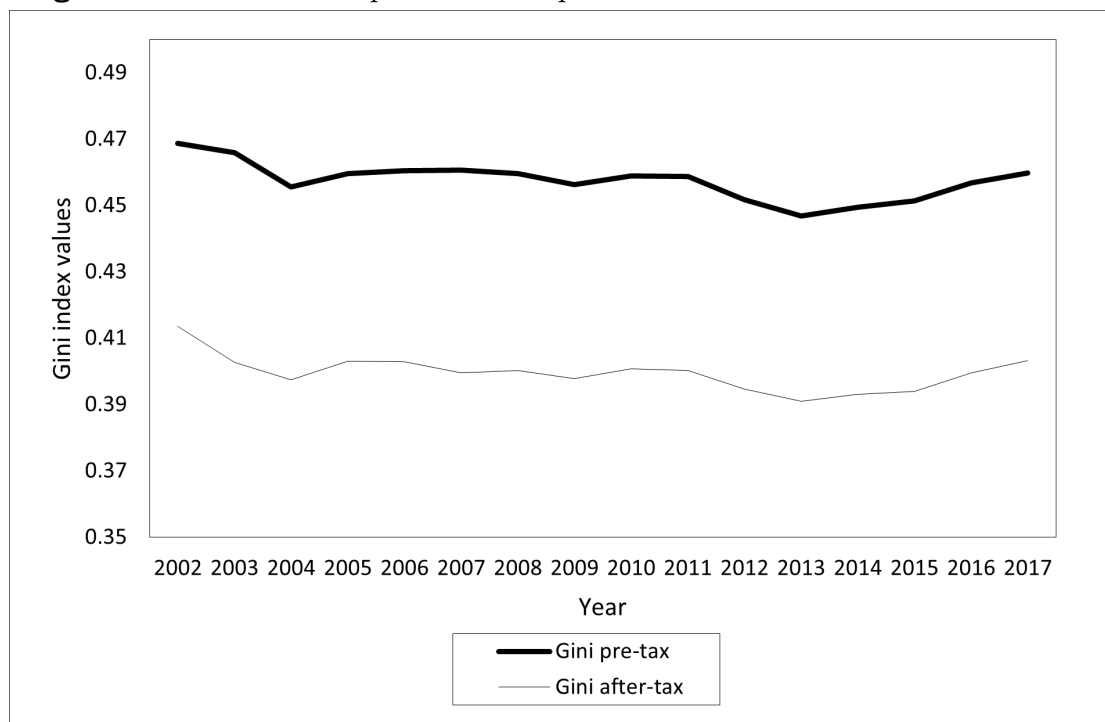
Source: Data processing on MEF data.

---

[11]The Lorenz curve is a graphical method of representing the degree of income inequality, that identifies the share of total income owned by cumulated fractions of the population, once this has been sorted by non-decreasing levels of income.

the collection of Irpef from 2002 to 2017, calculated on the equivalent household income. Evidence shows that the degree of inequality modestly reduced: in 2017 the pre-tax Gini index was 0.460, while in 2002 the value was 0.469.

Note also that the value of the Gini index will change depending on the variable, i.e. income before or after Irpef. Indeed, the Gini index calculated on income net of Irpef had values significantly lower than the one measured before tax, due to the redistributive effect induced by the application of marginal tax rates increasing by brackets and a system of deductions inversely related to the taxpayer total income. In the Municipality of Bologna, the Gini index after-tax was 0.403 in 2017, slightly lower compared to 2002, when the value of the index was 0.414 (see Figure 3.12).

Going beyond numbers and ignoring the fluctuations from one year to the next, note that the trajectories of the two indicators are parallel lines. The dynamics of the Gini index referring to the 2000s are characterized by a substantial stability[12].

An important aspects of the Gini coefficient is that it is independent of the size of economy and population. Since it respects the Pigou-Dalton principle of transfer[13], the sensitivity of the index to a transfer of income from a rich to a poor does not depend on the income levels of the two individuals, but on their difference in rank in the income scale. Therefore, regardless of the income level of the two subjects, a given redistribution will have a greater effect on the index the more distant from each other, in an ordinal sense, the two

---

[12]For further details about inequalities in the national context through the use of MEF data, see Acciari and Mocetti, 2013.

[13]The Pigou-Dalton principle states that, all other things being equal, a social welfare function should prefer allocations that are more equitable. In other words, it recommends a non-rank-switching transfer from someone with more of the income to someone with less, as long as no one else's holdings are changed.

subjects are. On the other hand, a limitation of the Gini index is that it does not provide information on the degree of asymmetry of a distribution, which implies that economies with a similar Gini coefficients can have very different income distributions. For this reason, it is good practice to combine this index with a broader assessment of inequality.

3.5.2   Percentiles and share ratios

Other commonly used measures of inequality focus on specific points or regions of the income distribution, such as the percentile ratios and the share ratios. Their appeal is that they are very intuitive and easy to calculate.

A commonly used percentile ratio, also called the interdecile ratio,

Table 3.10: Summary statistics for income inequality: interdecile and share ratios between 2002 and 2017.

Year	Interdecile ratios			Share ratios	
	90-10	90-50	50-10	80-20	90-40
2002	4.25	1.97	3.12	5.73	1.35
2003	4.03	1.96	2.83	5.36	1.30
2004	4.05	1.95	2.87	5.38	1.29
2005	4.29	1.99	3.03	5.75	1.37
2006	4.29	1.99	3.01	5.80	1.39
2007	4.10	1.96	2.99	5.60	1.35
2008	4.17	1.98	3.08	5.68	1.35
2009	4.23	1.97	3.21	5.77	1.35
2010	4.35	1.97	3.36	5.94	1.37
2011	4.48	1.96	3.58	6.12	1.38
2012	4.32	1.94	3.44	5.86	1.33
2013	4.25	1.93	3.42	5.76	1.31
2014	4.27	1.93	3.55	5.84	1.32
2015	4.27	1.92	3.57	5.88	1.32
2016	4.30	1.94	3.57	5.96	1.35
2017	4.38	1.94	3.62	6.08	1.37

Source: Data processing on MEF data.

is the 90-10 ratio, which is a measure that, as the Gini index, does not involve the calculation of a weighted sum of differences with respect to average income. The interdecile ratio shows the income level of individuals at the top of the income distribution (top 10 per cent) relatively to the income level of those at the bottom of the distribution (bottom 10 per cent). The 90-10 ratio, calculated on equivalent household income after taxes, was 4.25 in 2002 and 4.38 in 2017 (see Table 3.10). This means that the income of the richest decile was more than four times greater than that of the poorest decile for all the period available.

This measure can be split into 90-50 and 50-10 ratios to study income disparity separately between the upper and the middle, and between the middle and the lower end of the income distribution, respectively. For instance, evidence shows that there was more inequality among the poor than among the rich (see Table 3.10). In fact, 50-10 ratio assumed values far greater than the 90-50 ratio for all tax year. Besides, the gap between the middle and the lower end of the income distribution slightly grew, whereas the disparity between the upper and the middle of the distribution remained substantially stable.

A similar measure of income concentration can be obtained by looking at income shares of individuals at different part of the income distribution, for example, by dividing the population into quintile groups. Table 3.10 reports also the interquintile share ratio, 80-20, which shows the share of total income earned by the top quintile (top 20 per cent) relative to the share earned by the bottom quintile (bottom 20 per cent). This share equals 1 under a perfectly equal income distribution. In practice, however, the income share of the

top quintile was multiple times that of the bottom quintile, ranging from a factor of approximately six for all the available period.

Another commonly used share ratio is 90-40, called the Palma ratio. It represents a ratio of the income of the richest 10 per cent of the distribution to those in the bottom 40 per cent. The rationale for this measure is based on the empirical regularity that in most countries the upper-middle of the income distribution (from the 4th to the 9th decile) earn about 50 per cent of national income and that share is consistent over time and across countries (Palma, 2011). Therefore, changes in income inequality are almost exclusively due to changes in the share of the richest 10 per cent and the poorest 40 per cent. It is common to consider societies with a Palma ratio of 1 or below 1 to be relatively equal, meaning that the top 10 per cent does not receive a larger share of national income than the bottom 40 per cent. In our case, the Palma ratio maintained a constant value for the whole period analysed approximately equal to 1.3.

The interdecile ratios have the characteristic of concentrating on punctual sections of the income distribution, ignoring everything happens in other parts of the same. Therefore, it does not respect the principle of transfer. The benefit of using the percentile or share ratios is that they are transparent about what part of the distribution is driving the observed changes in the summary measure, which is more difficult to pinpoint when using the Gini coefficient.

We conclude this part of the analysis by taking an overview of the trend for each percentile of the percentage change between the average equivalent household incomes after-tax of 2002 and 2017 (see Figure 3.13). Each of these is compared to the overall average, which was contained and equal to -2 per cent. This substantial stability,

Figure 3.13: Percentage change between 2002 and 2017 for each percentile and overall total change.
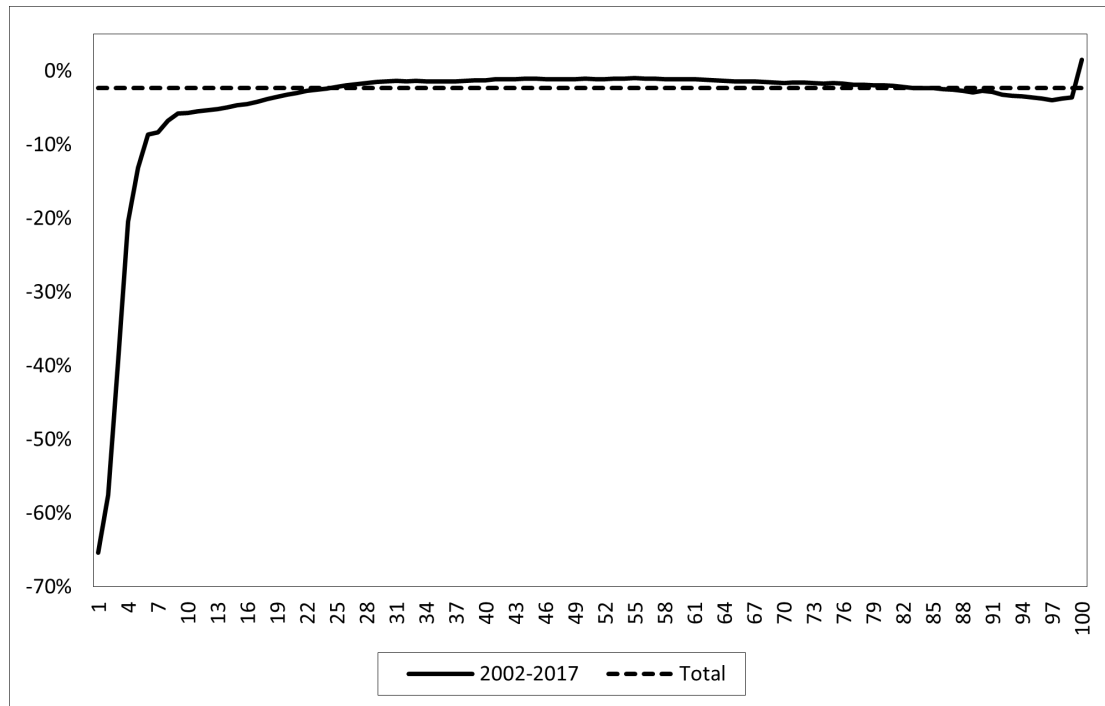
Source: Data processing on MEF data.

however, hides within it important changes in the income owned by each percentile of the distribution. For the first percentiles, data show a broad negative variation over the period examined, much more consistent for the poorest. Subsequently, limited variations are observed for incomes in the medium part of the distribution, which decreases as the percentiles increase until the 68th percentile. In the highest part of the distribution of income the percentage change grew again. On the other hand, incomes of taxpayers belonging to the richest percentile of the Bolognese population recorded increasing values.

3.5.3   Theil index

The Theil index is one of the most popular generalized entropy index that draws inspiration from information theory, in particular from the theory that measures the information value of a system of uncertain events. The decomposition property makes this index one of the most used tools in the study of the level and evolution over time of inequality within a given context (i.e. between social groups, geographical areas and so on). While the Theil index does not have an intuitive explanation, it is often used in empirical studies because of its decomposition property.

This tool is used in the literature when the population can be divided into several sub-groups (e.g. based on age, education, geographical regions) to quantify the relative weight of two components into which the inequality can be broken down: the between and within component. The first is a measure of the average gaps between individuals belonging to different groups, while the second takes into account the dispersion that are observed within each group of individuals with homogeneous characteristics. The Theil index, unlike the Gini coefficient, has the fundamental property of being able to be perfectly decomposed by subgroups. In other words, it can be expressed as the weighted sum of the between and within component. The first is calculated by reconstructing a distribution in which each member of a group has exactly the average income of the group and the within group differences are eliminated. Therefore, it reflects the average distances between groups within the income distribution. On the other hand, the second is expressed as a weighted average of inequalities within each group; in particular, in the breakdown of Theil index the weighting is carried out by attributing as a weight

of each group the share of income received on the total distribution. In other words, the between component, if expressed as the percentage share of the total inequality, expresses how much the overall inequality would be reduced if the average gap between the groups were eliminated, keeping the internal gap constant.

A picture of the Municipality of Bologna income inequality trends

Figure 3.14: Theil index on equivalent household income between tax years 2002 and 2017 across different population subgroups: citizenship and area of residence.
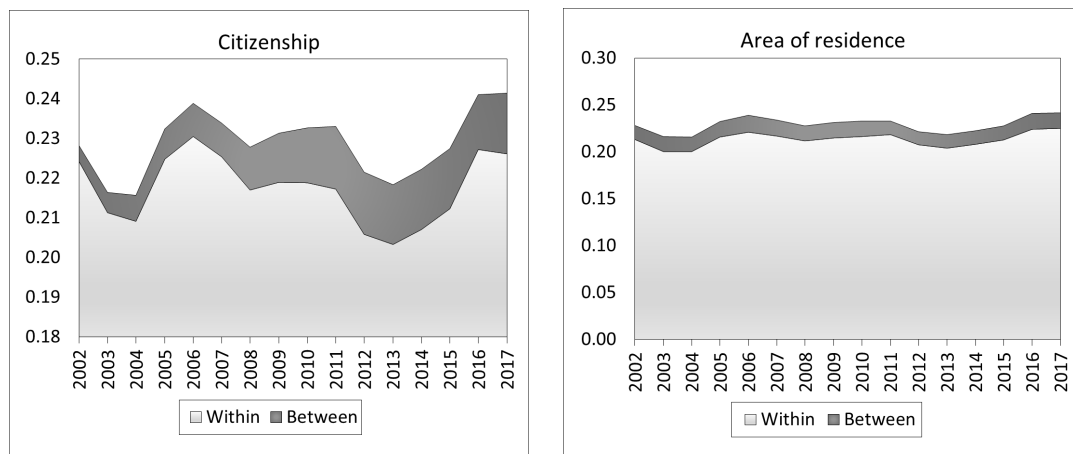
Source: Data processing on MEF data.

over time and its decomposition for two subgroups into between and within components is provided in Figure 3.14. The component capturing income inequalities within subgroups represents the lion's share of inequality. In fact, as regards to the citizenship decomposition, although the within-groups component has undergone significant fluctuations between 2002 and 2017, at the end of the period it still covered the 94 per cent of the total inequality measured with the Theil index. This change is certainly linked to the fact that the share of foreign taxpayers has tripled in the period examined. In 2017, the between-group component states about 6 per cent of the inequality, compared to the 2 per cent of 2002. As regards

the variations of Theil index components for the area of residence, evidence shows that they have been more restrained between 2002 and 2017. Note also that the component capturing income inequalities within different administrative areas of residence represents the greatest part of income inequalities (93 per cent) throughout the period analysed.

This outcome confirms also at a local level that the inequality patterns can be better explained by income disparities within regions and/or subgroups of a given area. Therefore, these results are relevant from a policy-making perspective in trying to identify the sources of inequality and assess the corrective measures. Firstly, they provide support for the implicit assumption of policy documents that economic integration should lead to convergence between groups. Secondly, given that within-subgroups inequalities currently explain an overwhelming proportion of the municipality income inequality, policies aimed at reducing disparities at subgroups level would offer the greatest prospect in the future, since they would tackle inequalities both within subgroups and consequently for the municipality as a whole.

3.5.4 Top income earners

The interest in income inequality has recently been renewed thanks to a literature that has shown, for a large number of countries over the last decades, a progressive concentration of incomes in the richest portion of the distribution. This part of the research focuses on the highest taxable incomes, the so-called "top incomes", i.e. those individuals whose taxable income is above the 90th percentile of the income distribution. This Section examines the performance of the

very high-income recipients share, and describe the evolution of top average income between 2002 and 2017 together with a description of income thresholds to enter the highest percentiles of the distribution, gender and age composition, and the source of income of this group.

Table 3.11 gives thresholds and average incomes for a selection

Table 3.11: Thresholds and average income in top income groups in the Municipality of Bologna in tax years 2002 and 2017.

Percentile threshold	Income threshold	Income groups	No. of taxpayers	Average income
2002				
Top 10%	50,097	Top 10-5	14,402	60,163
Top 5%	74,093	Top 5-1	10,646	99,906
Top 1%	156,362	Top 1-0.5	1,331	181,932
Top 0.5%	213,619	Top 0.5-0.1	1,064	298,890
Top 0.1%	475,722	Top 0.1-0.01	240	710,843
Top 0.01%	1,323,354	Top 0.01	27	2,247,665
2017				
Top 10%	51,635	Top 10-5	11,909	61,138
Top 5%	73,778	Top 5-1	9,527	97,919
Top 1%	154,269	Top 1-0.5	1,191	177,508
Top 0.5%	206,311	Top 0.5-0.1	952	287,085
Top 0.1%	474,327	Top 0.1-0.01	215	734,215
Top 0.01%	1,448,437	Top 0.01	24	2,896,019

Source: Data processing on MEF data.

of top fractiles in 2002 and 2017. The evolution, in real values, of taxable income thresholds to access different percentiles of the distribution had a rather stable trend over time, and only the top 10 per cent and the top 0.01 per cent slightly grew. More in detail, after a period of expansion between 2002 and 2010, the economic crisis was followed by moderate oscillations in real economic growth, resulting in an average income of top earners in 2012 lower than 2002. The richer the group considered, the higher the decrease in the average

income was: -7 per cent for the top 10 per cent and -28 per cent for the top 0.01 per cent between 2010 and 2012. This period was followed by a rising trend up to 2016: the top 10 per cent increased of 4 per cent, whereas the top 0.01 per cent increased of 98 percent from 2012 and 2016. On the other hand, the last period was characterized by substantial stability coming with a decrease occurred only for those in the top 0.01 percent (-17 per cent)[14].

As already pointed out, our tax records provide interesting insights on income concentration during the 2000s, which are not adequately caught by existent survey data. Results shows that the top income shares have remained rather stable since the beginning of 2000s: only the top 0.1 per cent share has registered a small increase between 2002 and 2017.

Figure 3.15 shows the share of total personal income owned by the top decile divided into three subgroups: the bottom half of the top decile (top 10–5 percent), the following 4 per cent (top 5–1 per cent), and the top percentile. The three series respond to two different patterns. The top 10–5 per cent share has displayed a rather stable pattern throughout the period. On the other hand, the top 5–1 per cent and the top 1 per cent have shown a fluctuating pattern until 2007, followed later by a decreasing trend which stopped in 2015, and finally a recovered trend in the last two years. Overall, the top 5-1 per cent decreased from 14.2 per cent to 13.3 per cent from 2002 and 2017, whereas the top 1 per cent share decreased from 10.6 per cent in 2002 to 10.1 per cent in 2017.

Figure 3.16 analyses concentration further by splitting the top 1 per cent into three groups: the top 1–0.5 per cent, the top 0.5-0.1

---

[14]Note that, in 2017, 24 individuals belong to the top 0.01 per cent income earners, therefore, the income change of a single taxpayer affects by far the average income of the whole group.

Figure 3.15: The top 10-5 %, top 5-1 % and top 1 % income shares in the Municipality of Bologna between 2002 and 2017.
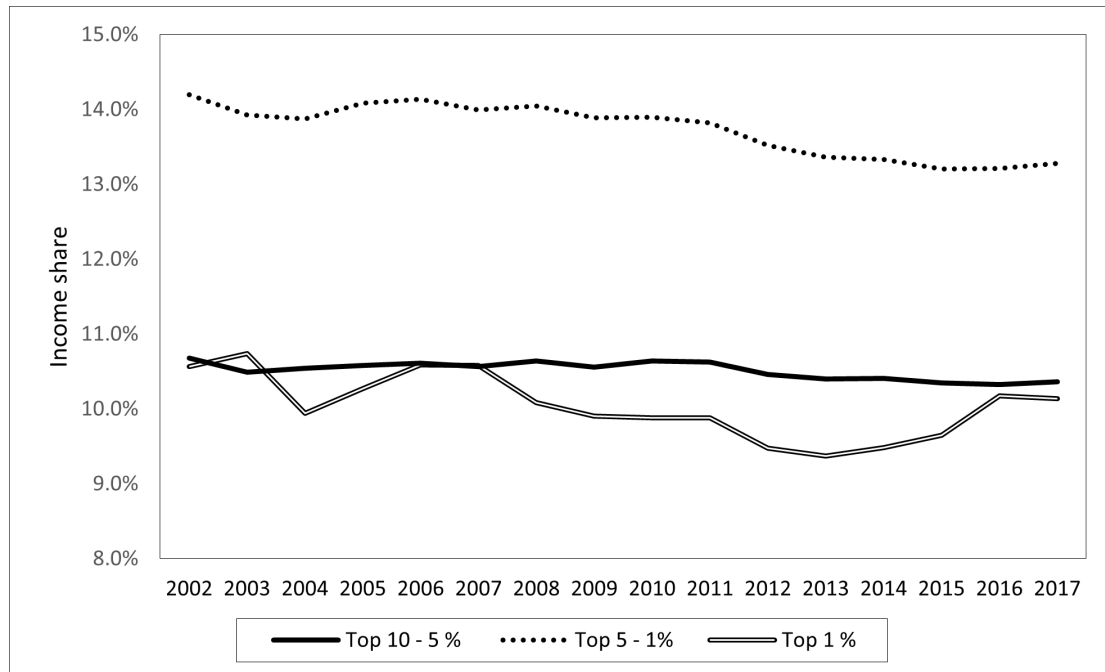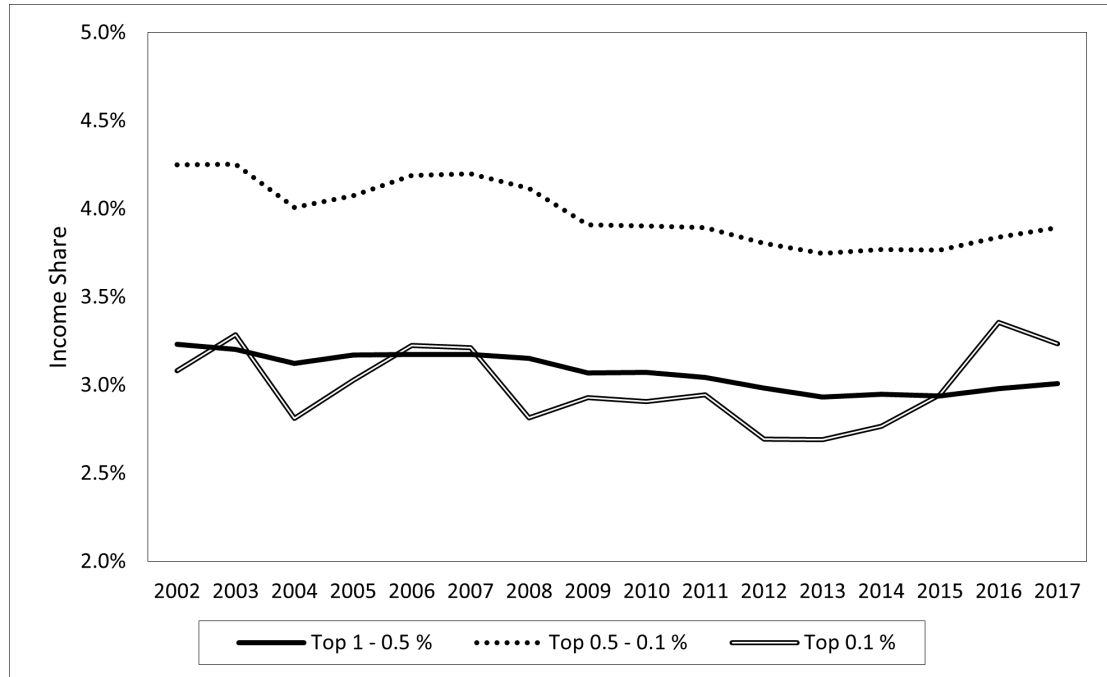

Source: Data processing on MEF data.

per cent, and the top 0.1 per cent. Overall, the decrease in the share in the 2000s is similar for the first two groups: the top 1–0.5 per cent decreased from 3.2 per cent to 3.0 per cent between 2002 and 2017, and the top 0.5-0.1 per cent decreased from 4.2 per cent in 2002 to 3.9 per cent in 2017. In contrast to the top 1-0.5 and 0.5-0.1, the top 0.1 fractile underwent more consistent fluctuations over the first decades of the 2000s. Overall, the share of total income received by the top 0.1 has registered a small increase from 3.1 per cent in 2002 to 3.2 per cent in 2017.

In countries with independent taxation of couples, as for Italy, it is possible to investigate the proportion of women in the top income groups. Therefore, we examine the gender divide at the top of the income distribution. There is a strong suspicion that women are under-represented in the highest percentiles of distribution and the

Figure 3.16: The top 1-0.5 %, top 0.5-0.1 % and top 0.1 % income shares in the Municipality of Bologna between 2002 and 2017.

Source: Data processing on MEF data.

Figure 3.17: Share of women in the top incomes groups in the Municipality of Bologna between 2002 and 2017.
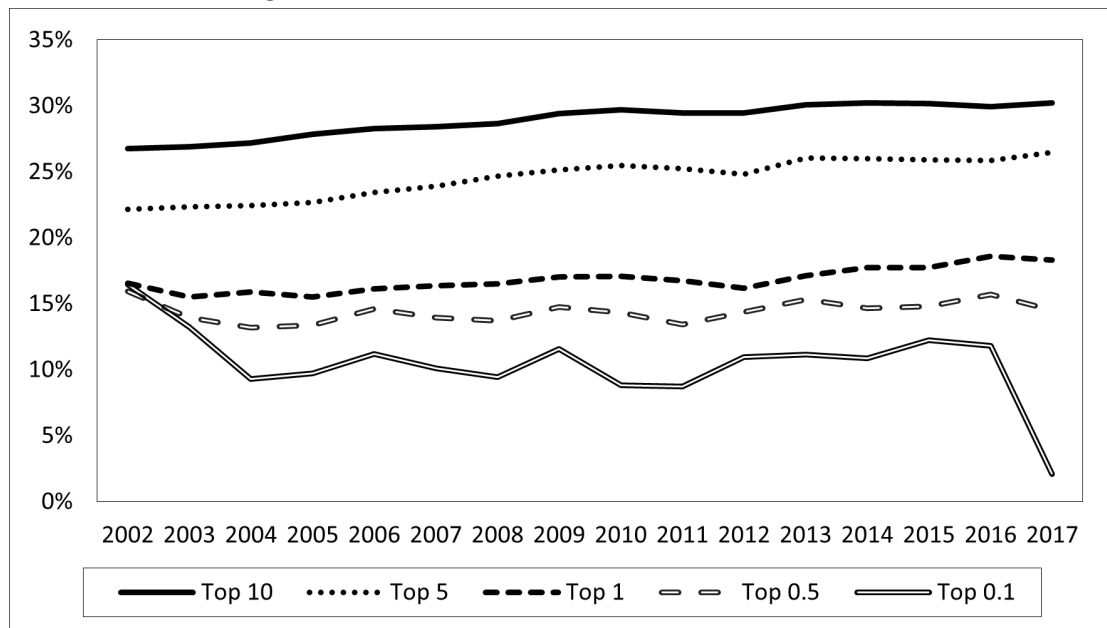
Source: Data processing on MEF data.

degree of under-representation increases as one approaches the top. However, there is little evidence in the literature to confirm this[15].

Evidence from Figure 3.17 is interesting in its own right, suggesting that the gender divide narrowed but it remained large: in 2017 there were 4.5 men in the top 1 per cent for every woman. However, the proportion of women in top income groups has, in general, rose over time. Looking at the most recent year available, evidence shows that the proportion of women in the top 10 per cent is 30 per cent, compared to the 27 per cent of 2002. Another striking feature is the decline in the proportion of women as one approaches the highest part of the income distribution: in 2017 the share of women in the top 10 per cent is 1.7 times the share of women in the top 1 per cent. Besides, dividing the highest percentiles in different groups, it is shown that the share of women in the top 0.1 per cent of the income distribution is only the 2 per cent of the total.

The access of women to high-pay job in the labour market and the increase in their presence in top positions is a key feature to promote the reduction of gender divide which, despite the decreased between 2002 and 2017, still has a remarkable value.

Another discussed issue is the situation of young taxpayers. It was observed how the work discomfort of the youngest worsened during the period available, leading the incomes of this age group (less than 40 years old) to decrease and the age gap to increase. Figure 3.18 shows the trend of under 40 age group share among the highest percentiles. Evidence suggests that the proportion of young taxpayers in top income groups decreased over time. Looking at the most recent year available, the proportion of the youngest in the top 10 per

---

[15]For further details, see Atkinson et al., 2016.

Figure 3.18: Share of under 40 age group in the top incomes groups in the Municipality of Bologna between 2002 and 2017.
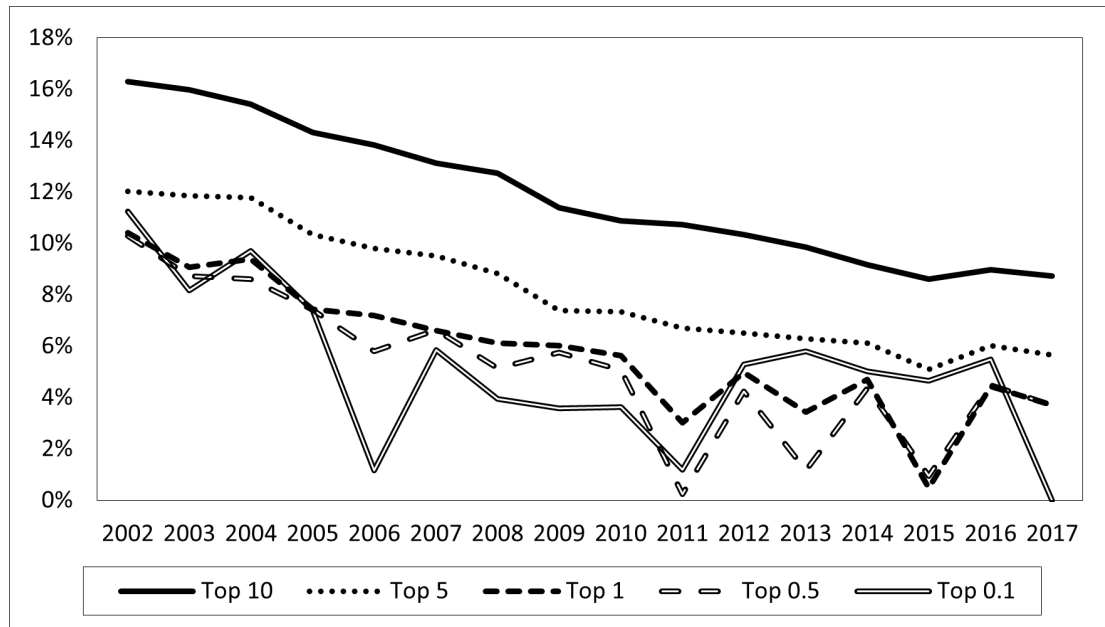
cent is 9 per cent, compared to the 16 per cent of 2002.

Another evidence from Figure 3.18 is that the proportion of young taxpayers is lower as one rises higher on the income scale: in the top 1 per cent it is equal to 4 per cent . Besides, the rate of decrease increases as one approaches the top: the share of young taxpayers in the top 1 per cent decreased of more than 68 per cent between 2002 and 2017, compared to a decrease of 52 per cent for the top 10 per cent.

Lastly, Figure 3.19 displays the income share composition of two income fractiles from 2002 to 2017: the top 10 per cent and the top 0.1 per cent. Evidence shows that employees stated the majority of taxpayers for the top 10 per cent between the 2000s with also slightly increasing values. An analogous situation was observed for the top 0.1 per cent. Our results show that as one approaches the top of the income distribution it is observed a decline of employees

Figure 3.19: The top 10 per cent and top 0.1 per cent income share composition in the Municipality of Bologna between 2002 and 2017.
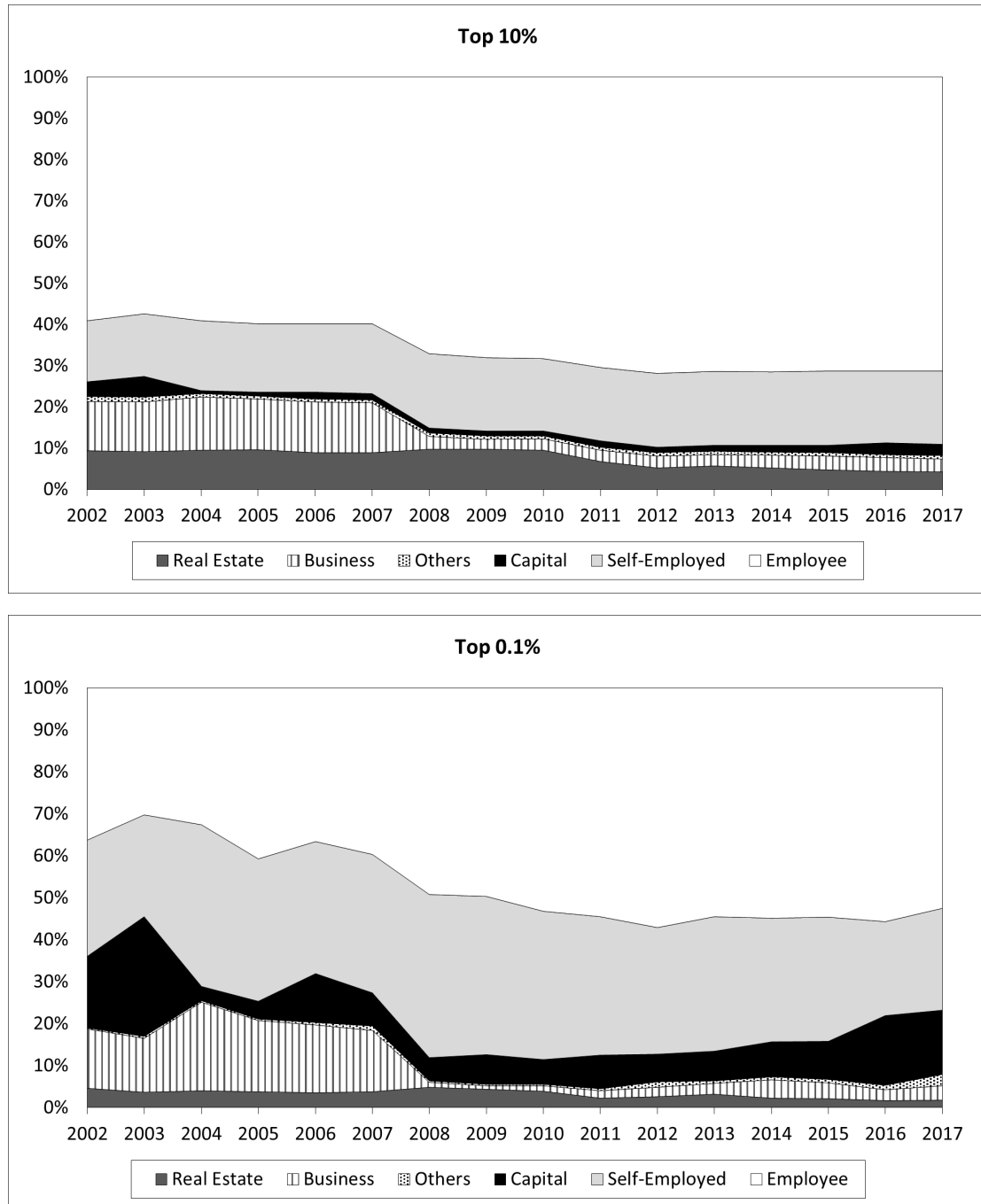
Source: Data processing on MEF data.

and retired in favour of self-employed and capital incomes. In fact, in 2017 income from dependent work and pensions accounted for the 69 per cent in the top 10 per cent, whereas it is 46 per cent in the top 0.1 per cent group. On the other hand, in 2017 the proportion of self-employed in the top 10 per cent is 17 per cent, whereas it is 21 per cent in the top 0.1 per cent. Similarly, capital incomes accounted for the 3 per cent in the top 10 per cent and 13 per cent in the top 0.1 per cent.

In conclusion, the presence of employees and retired in the highest percentiles of the income distribution is prevalent and rather stable over the 2000s: at the end of the period they counted for almost 70 percent of top incomes earners. This category includes the incomes of the most important executives and managers, but also of super-stars in sport and entertainment.

Alongside with the analysis of the top income shares, it is also useful to pay attention to the evolution of the share of the total income tax payed by the top incomes earners, as well as the other sections of the income distribution. Firstly, the analysis shows that in the Municipality of Bologna the portion of Irpef paid by the bottom 40 is less than 10 percent. On the other hand, the top 40 per cent contributes to more than 80 per cent of the total personal income tax paid (see Table 3.12). Moving to the highest part of the income distribution, data shows how the top 10 pays about half of all personal income tax of the Municipality of Bologna. Scaling the income distribution up to the highest percentile (top 1 percent), it can be noted that the taxpayers belonging to it pay about one fifth of the total income tax of the entire Municipality.

The trend of the total income tax shares paid during the 2000s shows

Table 3.12: Share of total income tax payed for different sections of the income distribution between the tax years 2002 and 2017 (percentage values).

Year	Bottom 40	Top 40	Top 20	Top 10	Top 5	Top 1	Top 0.01
2002	8.6	81.0	64.4	50.2	37.6	18.0	5.7
2003	8.5	81.6	65.2	50.9	38.4	19.1	6.8
2004	8.9	80.6	63.3	48.5	35.6	16.3	4.8
2005	9.0	80.3	63.7	48.9	35.9	16.7	5.2
2006	8.9	80.5	63.9	49.1	36.2	17.0	5.5
2007	8.4	81.4	65.1	50.2	37.1	17.3	5.5
2008	8.5	81.3	64.8	49.5	36.3	16.4	4.8
2009	8.7	80.8	64.0	48.8	35.7	16.1	5.0
2010	8.6	80.9	64.2	48.8	35.6	16.0	5.0
2011	8.6	80.7	63.8	48.6	35.4	16.0	5.0
2012	9.0	80.1	62.9	47.5	34.5	15.4	4.6
2013	9.4	79.7	62.4	47.1	34.2	15.3	4.6
2014	9.2	80.1	62.9	47.5	34.5	15.6	4.8
2015	9.1	80.2	62.8	47.4	34.5	15.8	5.1
2016	9.0	80.5	63.4	48.2	35.4	16.7	5.8
2017	9.0	80.5	63.4	48.2	35.3	16.6	5.6

Source: Data processing on MEF data.

that the share of tax paid by the higher sections of the income distribution underwent a slight decline between 2002 and 2017, in line with the results obtained from the analysis of the top income shares. On the other hand, it seems that the tax rate paid by the bottom 40 has slightly increased between 2002 and 2017, probably due to the increase in the tax rates for the lower brackets of the income distribution. Given the importance and novelty of the tax data at the municipal level, the next paragraph delves into all the aspect related to the personal income tax and its surtaxes.

## 3.6   Redistribution and progressivity of the personal income tax

The Italian personal income tax is one of the main instrument to achieve progressivity and redistribution at least de jure and a primary source of public revenues. Indeed, the Gini index assumed significantly different values when calculated with household equivalent income after tax (in 2017 it was 40 per cent) or before tax (in 2017 it was 46 per cent), reflecting the re-distributive role of taxation. To this regard, data from tax returns make it possible to assess the equalization effectiveness of the tax system and how this also varies locally, depending on the different distribution of the underlying income (Tomarelli and Acciari, 2011). Therefore, the aim of this Section is to study how the redistributive effect and progressivity of the personal income tax have changed at a local level through the use of some well-known statistical measures, i.e. the Reynolds-Smolensky and Kakwani indexes.

The initial design of Irpef imposed very high marginal tax rates on high incomes and a very large number of brackets (32), in line with the example provided by many advanced economies of the time. Thorough years, Irpef was the subject of profound structural changes, i.e. the reduction of the top marginal tax rate, the reduction of the number of tax brackets, and the expansion of income excluded from the tax, realized through greater tax credits or deductions[16]. Figure 3.20 summarizes the changes in the tax rates for the personal income tax occurred between 2002 and 2017. From a general assessment, changes on Irpef rates increased the amount of Irpef by 14 percent between 2002 and 2017, compared to a 16 per-

---

[16]For further details about Irpef, see Pellegrino and Panteghini (2020).

Figure 3.20: The tax rates for the Italian personal income tax in force from 2002
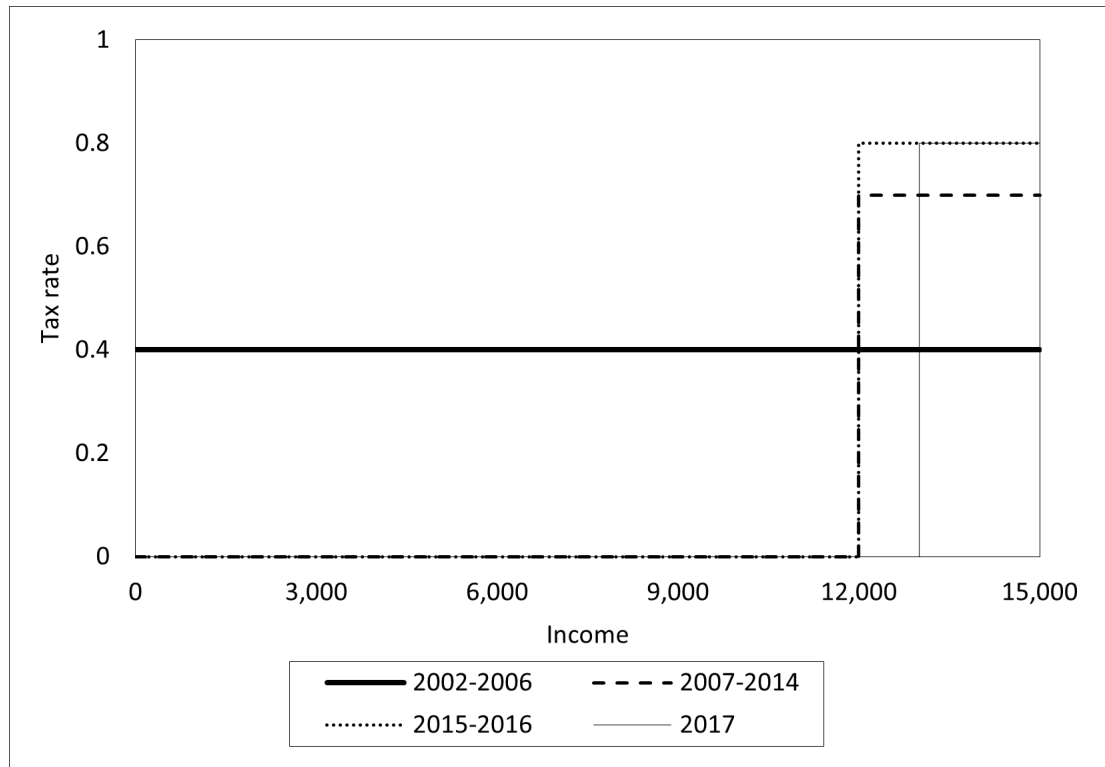to 2017.
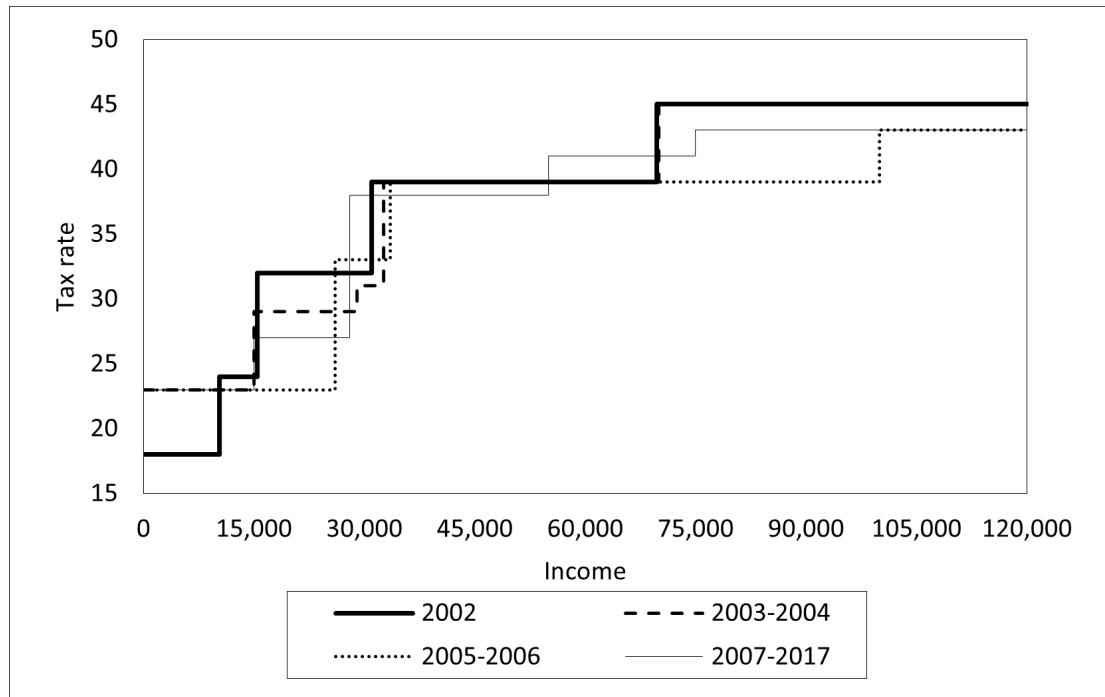
Source: Data processing on MEF data.

cent increase in total taxable income.

On the other hand, as an autonomous source of revenue, municipal-
ities have the legislative power to levy taxes on income. Indeed, the
municipal surtax on personal income constitutes the bulk of the mu-
nicipal tax revenue. In the Municipality of Bologna, several changes
occurred for the personal income municipal surtax, which resulted
in a considerable increase of the revenue from the municipal surtax
over the last decades (from nearly 24 million in 2002 to 54 million
in 2017), compensating for the notable reduction in central govern-
ment transfers over the analysed period, and underlining the process
of fiscal decentralization in progress (Rubolino, 2019).

Figure 3.21 summarizes the decisions taken by the Municipality
of Bologna between 2002 and 2017 regarding the additional surtax.

Figure 3.21: The tax rates for the personal income municipal surtax in force from 2002 to 2017 in the Municipality of Bologna.

The municipality chose not to apply the multiple tax rates, preserving the flat rate for the whole period examined. Between 2002 and 2006, the local average tax rate was the same for all the taxpayers regardless the amount of income declared, instead from 2007 the tax exemption for low-income taxpayers was introduced.

We now exploit our dataset in order to give an overview of the distribution of Irpef and its surtaxes[17] for different characteristics of the taxpayers (see Table 3.13). As already pointed out, the total taxable income declared by women in 2017 was still lower than that of men, although the number of female taxpayers was greater than

---

[17]In our dataset, the variable regarding the regional surtax is missing. For this reason, we estimated its values using the tax rates of Emilia-Romagna for the tax period 2002-2017, and the taxable income for the personal income municipal surtax purpose as tax base.

Table 3.13: Share of taxpayers, total taxable income and Irpef (including its additional surtaxes - regional and municipal) for the taxpayers residing in the Municipality of Bologna by sex, age, citizenship and area of residence for the tax years 2002 and 2017 (percentage values).

Category	2002			2017		
	Taxpayers	**Income**	**Tax**	**Taxpayers**	**Income**	**Tax**
Sex						
Women	53	40	34	52	42	37
Men	47	60	66	48	58	63
Age						
Less than 40	30	25	22	26	17	14
40-60	31	41	44	36	41	44
60-75	24	23	23	21	26	28
Greater than 75	15	11	10	18	16	15
Citizenship						
Italian	97	98	99	89	96	97
Foreign	3	2	1	11	4	3
Area of residence						
Borgo Panigale	7	5	4	7	6	5
Navile	17	14	12	17	14	12
Porto	9	9	9	9	9	9
Reno	9	7	6	9	8	7
San Donato	8	6	5	8	6	5
San Vitale	12	13	13	12	13	13
Santo Stefano	13	18	22	13	18	22
Saragozza	10	12	13	10	11	13
Savena	16	15	15	16	15	14

Source: Data processing on MEF data.

the male counterpart. In 2017 they contributed for the 42 per cent of the total taxable income, which was a slightly increasing value with respect to 2002 when they accounted for the 40 per cent of total income. This difference, obviously, had an impact on Irpef and its surtaxes. In fact, in 2017 women contributed to the total Irpef for the 37 per cent, whereas in 2002 the total Irpef due by women was 34 per cent. Evidence in the previous analysis showed that the youngest taxpayers were characterised by lower average income than

the other age groups and non increasing trends. This is confirmed also in this analysis: even if the under 40 numerically accounted for the 30 per cent of taxpayers in 2002, they contributed to the total Irpef only for the 22 per cent. This situation was even worse in 2017, when under 40 were the 26 per cent of the taxpayers and accounted for the 14 per cent of total Irpef. Evidence also shows that in the face of an increase in the number of foreign taxpayers of eight percentage points between 2002 and 2017, the total taxable income declared did not registered a proportional increase and, analogously, happened for Irpef. In fact, in 2002 foreign taxpayers accounted for the 3 per cent of the total, and contributed for 1 per cent to Irpef, whereas in 2017 they increased up to 11 per cent of taxpayers, and they paid the 3 per cent of Irpef. As for the area of residence, it is not surprising that Santo Stefano, the wealthiest neighbourhood in the municipality, was confirmed as the area that contributed the most to the total Irpef. It accounted for the 13 per cent of taxpayers for the entire period, and at the same time it contributed for the 22 per cent of the total Irpef.

We now focus on the progressivity and redistributive effect of the personal income tax and its surtaxes between 2002 and 2017.

It is useful to start with a graph showing the average tax rate and its change between the first and the last year of the period examined computed on the deciles of the pre-tax individual taxable income distribution (see Figure 3.22). Note that the average tax rate slightly increased for almost all the deciles of the distribution, a part from the first three, for which it remained rather stable.

Much more dynamic was the trend in the incidence of the municipal surtax considered separately. Also for Figure 3.23, individual

Figure 3.22: Incidence of the personal income tax on pre-tax income in the tax years 2002 and 2017 by deciles of the distribution of the pre-tax income.
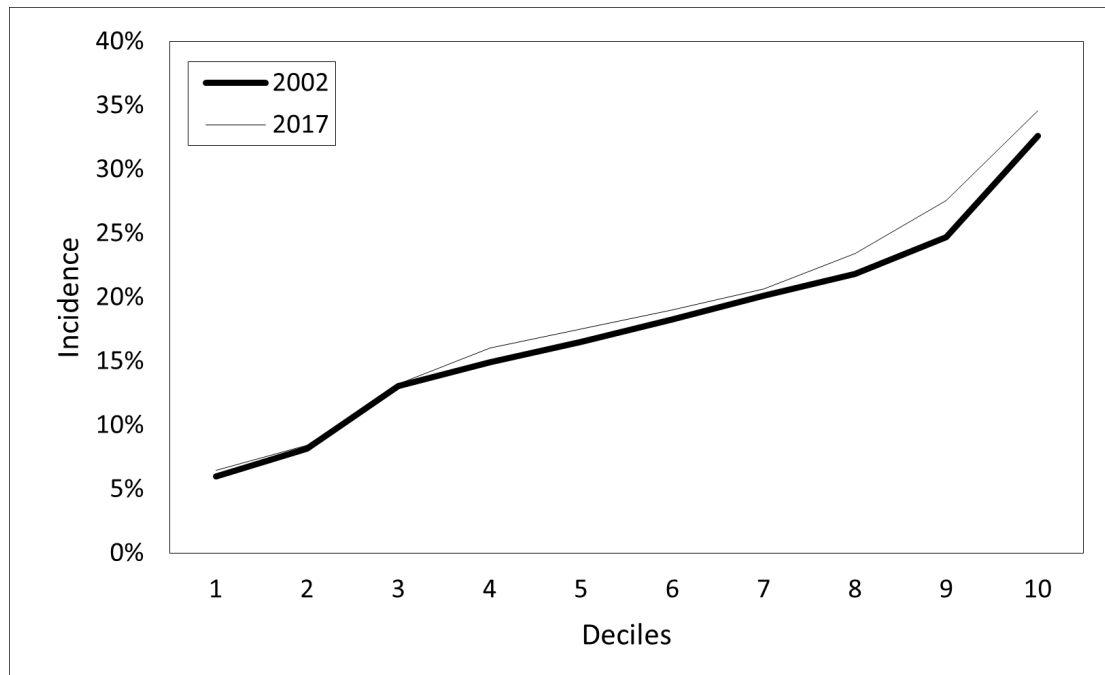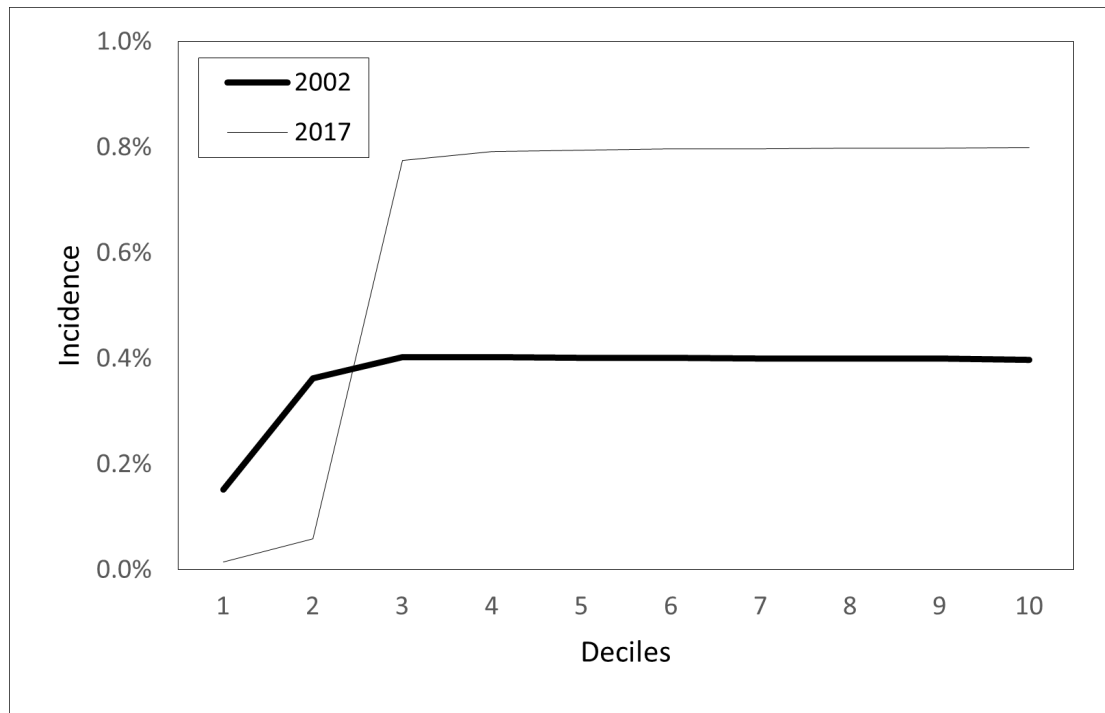

Source: Data processing on MEF data.

taxable income distribution is grouped into deciles of pre-tax income, and the incidence of the municipal surtax is computed for each decile in the tax years 2002 and 2017. Evidence shows that the introduction of the exemption threshold at the municipal level led to a decrease in the incidence of the local tax on the first two deciles. At the same time, the increase in the rate of the municipal surcharge led to an increase in the incidence in all the other deciles. Moreover, the trend of the average tax rate of the municipal surcharge shows the substantial proportionality of this tax from the fourth decile onward, unlike the personal income tax which instead has the characteristic of being progressive.

Overall, although some fluctuations incurred, the value of the incidence remained stable over the sixteen-year period when computed

Figure 3.23: Incidence of the personal income municipal surtax on pre-tax income in the tax years 2002 and 2017 by deciles of the distribution of pre-tax income.

Source: Data processing on MEF data.

alone (see Figure 3.24). On the other hand, when computed together with the additional surtaxes, the incidence of Irpef slightly increased, especially during years when a legislative change occurred, which resulted in higher marginal rates for the personal income surtaxes, i.e. 2007, 2011 and 2015. The graph showing the incidence on pre-tax income suggests that the personal income tax should have become both more redistributive and more progressive during the period examined.

However, a detailed assessment of these changes is provided by the Reynolds–Smolensky (RS) index of redistributive impact. This index is given by the change in the Gini index before and after the

Figure 3.24: Average overall tax rate (incidence) of Irpef and Irpef together with additional surtaxes in the Municipality of Bologna between the tax years 2002 and 2017.
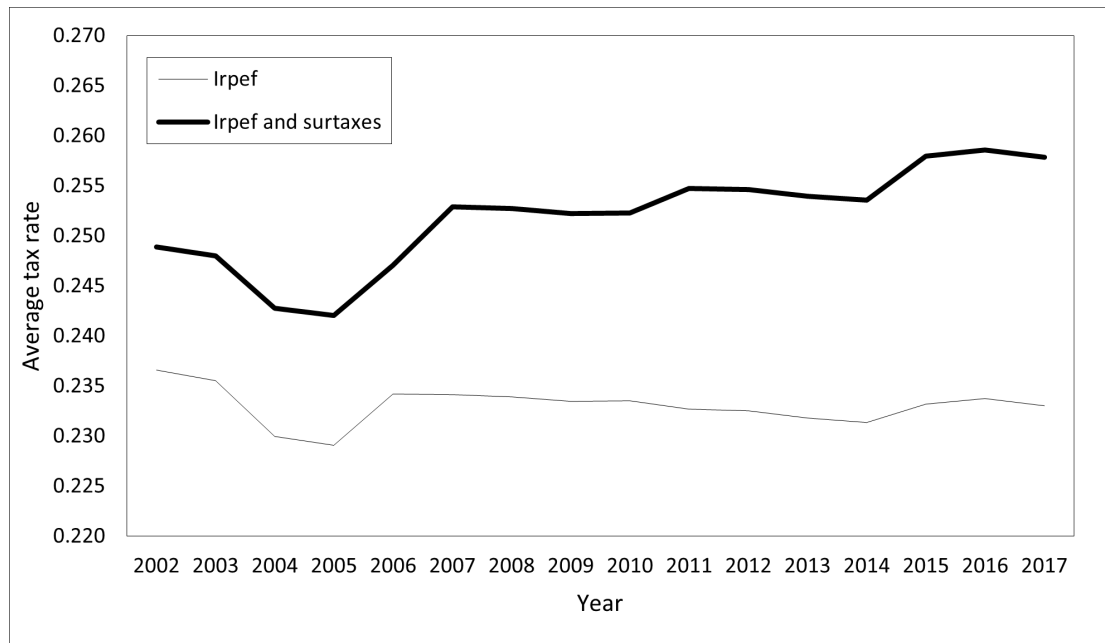
Source: Data processing on MEF data.

tax:

$$RS = G_Y - G_{Y-T},$$

where Y is the total income and T the personal income tax.
The higher the value of the Reynolds - Smolensky index, the greater is the redistributive effect of the taxation (Irpef and its surtaxes) . This measure can be usefully decomposed into the contribution of incidence and global progressivity, as follows:

$$RS = \frac{t}{1-t}K - D,$$

where t is the average tax rate, K is the Kakwani index of global progressivity, and D is a term that takes into account the reordering of units between gross and net income. This relation shows how the redistributive effect of the tax system, measured by the Reynolds-

Smolensky index, is given by the product between the level of progressivity, measured by the Kakwani index, and the factor $t/1-t$, which is an increasing function of the average tax rate.

The Kakwani index, in turn, is the difference between the concentration coefficient of the tax (C_T) and the Gini index on gross income itself:

$$K = C_T - G_Y.$$

The Kakwani index takes into account the fact that a progressive income tax has an higher concentration of the levy than on gross income. Therefore, the higher the Kakwani index, i.e. the concentration of the tax in relation to that of gross income, the greater is the degree of progressivity of the tax.

Table 3.14 reports the measures that describe the redistribution

Table 3.14: Reynolds - Smolensky index and its decomposition in the Municipality of Bologna for the tax years 2002, 2007, 2012 and 2017.

Measure	2002	2007	2012	2017
Reynolds-Smolensky index (RS)	0.058	0.065	0.061	0.062
Incidence (t/(1-t))	0.331	0.338	0.342	0.347
Kakwani index (K)	0.183	0.197	0.186	0.189
Gini coefficient for the gross income (G_Y)	0.443	0.433	0.418	0.419
Gini coefficient for the net income (G_{Y-T})	0.384	0.368	0.357	0.357
Concentration coefficient (C_T)	0.625	0.630	0.605	0.609
Reordering effect (D)	0.002	0.0002	0.003	0.004

Source: Data processing on MEF data.

and progressivity of Irpef in the Municipality of Bologna in the tax years 2002, 2007, 2012 and 2017. The RS index slightly increased from 0.058 to 0.062 between the first and the last year of the period. This small increase is due to both an increase in the incidence (t/1-t went from 0.331 to 0.347) and at the same time a small increase in the K index from 0.183 to 0.189.

Figure 3.25 shows in detail the dynamics of the Reynolds-Smolensky

Figure 3.25: Trends in Reynolds-Smolensky (left axis) and Kakwani (right axis) indices computed on individual taxable income for taxpayers residing in the Municipality of Bologna over the period 2002-2017.
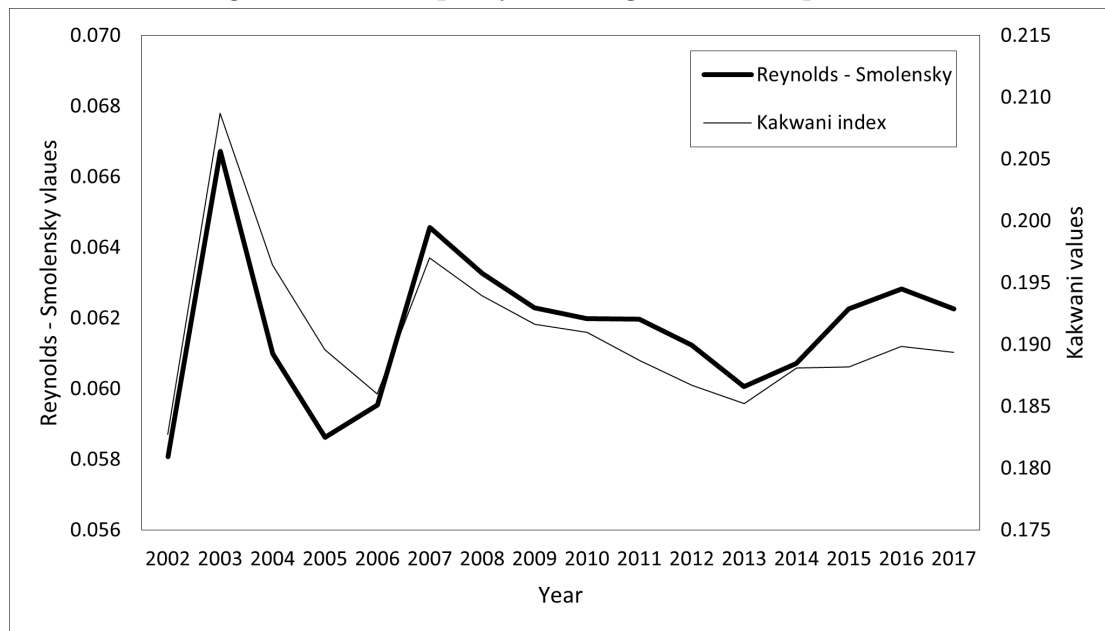
Source: Data processing on MEF data.

and Kakwani index for the period 2002-2017. The trend can be divided into different phases. Before 2007, the values of RS index show wide fluctuations, reaching a peak in 2003, as a result of the introduction of the no-tax area for low-income taxpayers. From 2007 to 2013, the index shows a progressive decrease. Finally, after a slight increase between 2013 and 2016, in the last year the index underwent a further decrease. As for the Reynolds-Smolensky index, up to 2007 the Kakwani index shows fluctuations together with a peak in 2003. From 2007, the Kakwani index shows a downward trend which stopped in 2013, reaching a value of 0.185. Consequently, in the last period, it shows a modest increase, that ended in 2017 with the value of 0.189.

The trend observed for these two indexes is consistent with the re-

lation between them. The RS index increased between 2002 and 2003, as a result of an increased level of tax progressivity measured by the Kakwani index. The increased K index is found to have a redistributive effect which is stronger than the adverse effect produced, in the opposite direction, by the decreased average tax rate. On the contrary, between 2003 and 2005 the RS index decreased, as the result of a decreased level of tax progressivity and of a decreased average tax rate. Between 2005 and 2006, an increased level of tax equalization is observed, due to an higher value of the tax incidence. At the same time, an increasing redistribution of the disposable income is also observed, which occurred despite the reduced level of tax progressivity. Another fluctuation incurred between 2006 and 2007, when an increase level of redistribution operated by the tax system is observed despite the decreased level of tax progressivity. From 2007, tax statistics were influenced by a new Irpef law, which have modified the income brackets and reintroduced family and income tax credits instead of tax allowances in force before this year. Between 2007 and 2013 the RS index decreased slightly, as the result of a decreased level of tax progressivity, which had a stronger effect than the effect produced by the average tax rate. Between 2013 and 2016 the increase in the RS index was due to a combined effect of an increase in the tax incidence and progressivity. Finally, between 2016 and 2017 the RS decreased, as a result of a decreased value of the K index and the tax incidence.

The analysis performed over the period 2002-2017 highlight that the reforms occurred had different impacts on the level of tax progressivity and redistribution operated by the Italian personal income tax, indicating an albeit little increase in the effects of redistribution and

progressivity of Irpef and its surtaxes.

It is difficult to compare our results with other studies, due to a scarce use of administrative sources at the municipal level, and also to differences in the period examined and/or the variables employed. However, in the national context several studies about the redistributive effect of the personal income tax were published. Tomarelli and Acciari (2011), using tax files for Italy, found similar trends for the two indexes, although rather different values: in 2002 the RS index was equal to 0.049, and the Kakwani index was equal to 0.215, whereas in 2009 the values were equal to 0.051 and 0.220, respectively.

Chapter 4

In-work poverty in the Municipality of Bologna, 2007-2017

4.1 Introduction

Poverty and its economic and social implications are a central topic in the research agenda both nationally and internationally: indeed, preventing poverty and reducing its incidence and persistence is a goal shared by all the governments. However, the attainment of this goal is not an easy task. Many of the analysis in the past focused only on the study of poverty at a point in time, unless it has been recognized that this approach is useful to state the extent of poverty and how it differs at two points in time, but it does not provide useful information about the individual experience of poverty (OECD, 2001). Thanks to the availability of suitable data, in the last decades the research has made substantial progress about the dynamics of poverty experiences, as the cross-sectional perspective has increasingly been complemented by the longitudinal dimension of the poverty problem at an individual level.

Alongside this strand, a growing concern about poverty among working individuals has been developed over the last decades. Indeed, the empirical research shows that the labour market is not always sufficient to ensure a way out of poverty. Nevertheless, with respect to in-work poverty, we find a high reliance on cross-sectional analysis, and relatively little is known about the dynamics of in-work poverty over time (Hick and Lanau, 2018).

Moreover, even though the importance of poverty measures at sub-national level is widely attested (Biggeri and Pratesi, 2017), official channels do not generally provide poverty statistics for detailed geographic areas since official surveys are not built to provide reliable estimates at such a fine level of spatial detail. In Italy, it is known, in fact, that the only surveys produced by the National Institute of Statistics, that officially estimate poverty (Statistics on Income and Living Conditions and Household Consumption), provide representative data only at the national and, at most, regional level.

The analysis of in-work poverty at the local level is of great interest in order to be able to characterise the complete low-income pattern of individuals along time. Indeed, some individuals experience only a single short spell of poverty, while others are kept in a long-term poverty status, which implies lower standards of living and a high risk of social exclusion over time. The advantage of distinguishing the characteristics of individuals that suffer longer period of poverty in contrast with those that experience poverty for a relatively short time, is of particular importance for the policy maker in order to address policy solutions that can successfully reduce working poverty. Short period of poverty may be assessed with short period money transfers working as income substitutes. On

the other hand, for longer period of poverty these strategies are insufficient: policy makers should design, for example, labour market policies promoting employment stability and different stable monetary transfers for poor adults (Arranz and Cantó, 2012). Design of anti-poverty strategies also depends on different types of events, namely "trigger events", that may influence the way to poverty or the escape from it. For example, knowing whether the transition into or out of poverty mainly involve changes in the household composition may be important for policy design.

This work focuses on the degree of in-work poverty in its entirety in the Municipality of Bologna for the period 2007-2017. The attention paid to the longitudinal dimension of the phenomenon at the local level represents an important novelty with respect to existing research, which has mainly focused on the spread and intensity of poverty at a given point in time. To achieve this goal, tax data at our disposal was exploited in order to estimate the relative poverty of individuals across an observation window of eleven years. We employed a poverty approach based on income declared in the tax return and referred to the local context, calculating a local relative poverty line, based on the equivalent income of households residing in the Municipality of Bologna. Indeed, Biggeri and Pratesi, 2017 showed the limits of the adoption of a single national threshold to estimate poverty at the sub-national level in Italy, where there is a pronounce income gap between the North and the South.

After the seminal works of Allison, 1982 and Bane and Ellwood, 1986, literature on poverty dynamics has largely focused on the extent and composition of persistent poverty through the study of the time spent in poverty or non-poverty, i.e. spell length or spell dura-

tion, and the estimation of entry and re-entry hazards[1]. Following this approach, we tackle poverty problem from a longitudinal point of view by introducing a new measure that takes into account the duration of the poverty spells, analysing the transitions to and from in-work poverty and estimating the hazard rates of exit and re-entry across the working population aged 16-64 using the longitudinal dimension of the administrative dataset for the period 2007-2017.

The case of the Municipality of Bologna is of particular interest since few works on poverty dynamics have been conducted for Italy, and none at the local level for such a long-time span. Thus, this analysis contributes to the literature extending the debate on the dynamics of in-work poverty at the municipal level using a new source of data. Our strategy also allows us to show the importance of individual and household characteristics to better explain the complete in-work poverty experience of individuals.

The main findings show that the incidence of in-work poverty, measured at cross-sectional level, in the Municipality of Bologna reached high level during the period examined, related to the interplay between low-income workers, mainly younger workers and non-Italian citizens, and the limited number of income recipients in the household. Moving to the longitudinal dimension, at a first glance it came out that the number of low-income workers in a single year is substantially smaller than that of taxpayers touched by in-work poverty over a long period. Our results show that during the 2007-2017 period only one out of ten experienced a single year in poverty. The remaining not negligible portion of working individuals registered recurrent low income over a longer period of time. Among

---

[1]For further details, see also Stevens, 1999, Devicienti, 2002; Biewen, 2006.

them, the groups with the most dramatic experience of persistent in-work poverty were non-Italian workers, younger taxpayers, and households with only one income earner; on the other hand, the subgroups that appeared to ward off poverty were couples with children, older taxpayers and households with three or more income earners. The analysis of transitions to and from poverty shows the importance of the trigger events in the routes to and from poverty: evidence shows that individuals who worsened their in-work poverty status reported an higher percentage of transitions from households with more than one taxpayer to families in which there is a single income earner.

The Chapter is organised as follows. In the next Section, we outline some of the key findings of the literature on the analysis of poverty dynamics, and we introduce the analysis of poverty at the local level making a sketch on what is already known about the Municipality of Bologna. Subsequently, we detail the data and the methodological approach employed (Section 3). Section 4 focuses on the cross-sectional analysis of in-work poverty. Finally, Section 5 broadly presents the working poverty from a longitudinal standing point, discussing our main findings on duration, recurrence, and transitions to and from in-work poverty.

4.2   Related literature

A strand of vast literature concerns the dynamic analysis of poverty for a single country or for many countries. In the United States during the eighties of the former century, mainly as a result of the availability of a mature and reliable longitudinal data survey - the Panel Survey of Income Dynamics (PSID) - ongoing since 1968, Bane

and Ellwood (1986) published one of the first and most relevant paper in this literature. Their approach is based on the definition of spells of poverty, and a key innovation of the study is the characterisation of the events which leads to the beginning or the ending of a spell of poverty. In the European context, only a decade later, Duncan et al. (1993) compared, for the first time, family income position at two points in time, using transition tables, in a group of eight countries, which had gathered the requisite longitudinal data for their population.

Unfortunately, Italy does not have a long tradition in longitudinal research. The introduction of the European Community Household Panel (ECHP) survey in 1994, replaced in 2004 by the Statistics on Income and Living Conditions (EU-SILC) survey was the first step to align with the growing methodological needs and demands of researchers to have accurate and comparable longitudinal data information for most countries in the European Union. The main reasons for this scarce presence of longitudinal surveys were certainly linked to the difficulties in the design and realization of the panels themselves, considering the high costs and the methodological and data processing complexity they present. The exploitation of this dataset, together with data from some nationally based panels available for some countries - the Household Panel Survey (BHPS) for the UK, the German Socioeconomic Panel (GSOEP) for Germany and the Encuesta Continua de Presupuestos Familiares (ECPF) for Spain - allowed researchers to study poverty dynamics in some European countries (e.g. Jarvis and Jenkins, 1997; S. Jenkins et al., 2001; Devicienti, 2002; Cantó, 2002; Biewen, 2006, to name but a few).

After the dissemination of panel data, the focus on the longitudinal dimension of poverty has started to represent an important innovation with respect to pre-existing research, in Italy as well. Poverty dynamics were primarily analysed in Italy in a work of OECD (2001), using data from the ECHP, with reference to the first three waves, within an analysis carried out primarily in a comparative key for 12 European countries, Canada and the United States. A similar analysis was also carried out in the Report on Policies against Poverty and Social Exclusion of the Commission of Inquiry on Social Exclusion (Commissione d'Indagine sull'Esclusione Sociale, 2003), again limited to the first three years of the ECHP. Devicienti and Gualtieri (2004) were the first to study characteristics and determinants of the longitudinal poverty experiences at the individuals level for Italy only, extending the period of observation to the first 8 waves of the ECHP (1994-2001), whereas Devicienti et al. (2014) used multiple spell hazard rate models on income and lifestyle deprivation to estimate individual poverty persistence in Italy over the same period. Finally, Giarda and Moroni (2018) used dynamic random effects models in a comparative analysis on poverty state dependence in Italy, France, Spain, and the UK, exploiting the longitudinal component of EU-SILC for the period 2009–2012.

Rather than resorting to traditional panel surveys, the growing availability of long panel data, provided from administrative sources, has allowed researchers to investigate social and economic phenomenon also from a longitudinal perspective. However, despite the widespread use of administrative data in the cross-sectional measurement of income dynamics and social indicators, the use of this data from a longitudinal standing point is still underdeveloped. Few papers have

exploited this type of data to measure poverty dynamics: examples are Finnie and Sweetman (2003), who used tax records covering the period 1992-1996 to explore poverty rates in Canada, and Larrimore et al. (2020), who were the first to consider the poverty rates and the persistence of poverty directly in U.S. tax data between 2007 and 2018.

Moreover, although increasing attention is now being paid by local public authorities to outline social policies aimed to alleviate poverty and income inequality in the local community, the analysis of poverty dynamics at the local level is still not well advanced, despite the potential of administrative data to support longitudinal studies also within small geographical areas. Fransham (2017) used administrative data about low-income families claiming support for paying their housing costs in the city of Oxford to explore the potential of administrative data on housing benefits and to construct a small area dynamic indicator of income poverty between 2010 and 2014.

To the best of our knowledge, in Italy few analyses have been conducted exploiting the potential of administrative data to investigate the income and poverty dynamics at sub-national level. The manuscript of Comune (2014) began to properly fill this gap, and to investigate income poverty in Italy at the local level, for the Municipality of Brescia, over the period 2005-2008, using appropriately linked registry and tax data. The need to know the state of indigence of households at the local level was motivated by the author by a total absence of official statistics on income and poverty at such a fine territorial level.

In fact, attempting to get a picture of poverty at the local level with

the information available, local government can dispose of accurate information on people living conditions at the national and, at most, regional level thanks to the available survey data produced by the National Institute of Statistics. Indeed, the Statistical Office of the Emilia-Romagna Region published every year a report on Income, Poverty and Social Exclusion at regional level, using data from the EU-SILC sample survey (Regione Emilia-Romagna, 2019). Evidence shows that in 2017, families residing in Emilia-Romagna received an average net income of more than 35 thousand Euro, which is one of the highest values recorded in Italy, after the autonomous Province of Bolzano. In 2017, in Italy the incidence[2] of the risk of poverty was 20.3 per cent, while in Emilia-Romagna was almost halved, standing at 10.5 per cent of the total number of residing individuals. Thus, Emilia-Romagna ranks among the most virtuous regions, although the indicator of risk of poverty has not returned to pre-crisis levels (8.7 per cent in 2008). So far, increasing the accuracy of information up to a finer territorial level has implied the requirement of small area estimation methods to supply the lack of survey data and the not yet widespread exploitation of administrative sources of data at the required geographical level. These methods use statistical models that relate the survey data with auxiliary information to produce estimations for detailed geographical areas. Exploiting small area estimation techniques, Ferrante and Fabrizi (2020) estimated poverty, deprivation and inequality parameters for Emilia-Romagna health districts and provinces, with reference to the period 2008-2017. Evidence shows that individuals at risk of poverty in the Province of

---

[2]An individual is considered at risk of poverty if its household equivalent income is below the poverty threshold, which is conventionally set at the 60 per cent of the national equivalent median income. In Italy, in 2017 the poverty line was 9,925 Euro.

Bologna were 7.82 per cent in 2008 and 8.78 per cent in 2017, which turned out to be the lowest value among all the provinces of the Emilia-Romagna Region for the last period. On the other hand, if we consider the health district, i.e. the one pertaining to the City of Bologna, the estimate of poverty risk was 10.05 percent in 2017 and 7.90 per cent in 2008.

Given that there is no information available at the municipal level and the reference to municipalities is important in order to offer the local policy makers reliable disaggregated information, which are more appropriate than the regional or even national ones, to design effective policies and actions against poverty, the analysis presented in this manuscript supply at the lack of this information at the municipal level exploiting the tax data integrated with registry data. Moreover, we borrow from the methodology on the analysis of the poverty dynamics, which allows us to improve the knowledge on the patterns and determinants of poverty incidence, transitions, and persistence referred to the population of the Municipality of Bologna for a window of eleven years (2007-2017).

4.3   Data and definitions

Although we are aware that the phenomenon of poverty has different types of manifestations and therefore has a multidimensional nature[3], the theoretical approach of reference adopted in this study is one-dimensional. It is based on a deprivation of a monetary nature without considering other aspects that may affect quality of life and well-being, which are generally included in multidimensional poverty approaches. In other words, in this study reference was made to eco-

---

[3]For further details, see Baldini and Toso (2009)

nomic poverty, that is, we focused only on a single indicator: income. Given the origin of our dataset, the definition of income used in this Chapter includes all forms of income declared by individuals in the tax return. We consider only working individuals, i.e. those who declare labor market earnings from employment and self-employment. This definition is more encompassing than the analysis conducted at the European level (Ahrendt et al., 2017), where the individual is included in the sample only if he or she has worked for at least seven months in the previous year, and it does not to exclude any category of working individuals. Besides, given that the analysis is focused on individuals working only, we consider people aged 16-64. Hence, our definition of income includes all forms of income, i.e. the annual taxable income net of the personal income taxes (Irpef and its additional surtaxes), but only for those who are in-work in the years examined.

We follow the conventional practice and assume that the individual lives within a family where resources and various economic and demographic events are shared. Then, the total household income – the sum of all the incomes of each household member - which is supposed to be distributed equally within the household, is considered the most suitable measure to assess the living standard of each member of the household. Following the assumptions made in other works on poverty dynamics (see Devicienti, 2002; Devicienti and Gualtieri, 2004), the unit of analysis adopted is the individual, but each individual's poverty status is assessed in terms of the adequacy of total income available to the household. In order to adjust for differences in family size and composition, annual taxable income is divided by the modified OECD equivalence scale. The resulting

equivalent income measure is an estimate of the potential consumption for each individual in a household.

Therefore, individuals are defined as being in-work poor if their equivalent household income falls below a particular low-income cut-off. The measure used in this work is closed to the EU indicator of in-work poverty risk, which is defined by combining individual and household economic conditions (Ahrendt et al., 2017; Raitano et al., 2019). According to this definition, in our analysis, an individual is considered in-work poor when he or she has worked in the year previous to the submission of the tax return, as explained above, and if he or she belongs to a household which has an income below a poverty threshold, defined as 60 per cent of the median household income of the reference population. Hence, this indicator is assessed according to the individual employment condition, which defines the population subgroup to be analysed, and the household income, which identifies the poverty status of the worker.

We adopt a local poverty line that is based on the equivalent income of households residing in the Municipality and, therefore, referring to the level of well-being of the population residing within the municipal boundaries. In the last decade, research literature has highlighted the limits of the adoption of a single national threshold to estimate poverty at sub-national levels (Biggeri and Pratesi, 2017; Comune, 2014), showing in our Country that the national threshold tends to overestimate the poverty levels of regions in the South and underestimate those in the Center and North of Italy (Chiappero-Martinetti and Pansini, 2009). Besides, poverty lines differentiated by territorial level have already been used in the European context, where the incidence of poverty in individual member states has been estimated

using of as many national poverty lines as there are EU countries (Chiappero-Martinetti and Bottiroli Civardi, 2002). Of course, the adoption of a poverty line relative to the local context let the specific inequality within the territorial boundaries be measured, but the results will be not comparable with other local realities. On the other hand, the adoption of a single national poverty line has the advantage of allowing comparisons between different realities, but the disadvantage of not considering the disparities present between them.

In order to analyse poverty dynamics, panel data are required since they allow the equivalent household incomes of a sample of individuals to be followed over a multi-year period. The choice of the length of the period over which we should observe an individual's income is not a clear cut. A relatively long observation window better enables us to judge whether, for example, a spell of poverty represents a temporary phenomenon, or whether it forms part of a pattern of recurrent or long-term poverty, and also model them better because their start dates are more likely to be observed. Due to concerns about our data archive before 2007, as exposed in Chapter 2, we take into consideration the period 2007-2017.

We require an individual not to have missing observations in between years to be included in our dataset, i.e. we select individuals present in all the consecutive years until either the observation window ends or individuals leave prior to the end of the period.

For the longitudinal analysis, we need a further adjustment: we consider working individuals that are present in the observation windows in at least four consecutive interviews. Clearly, if one also considers those who remain in the panel for fewer years, estimates

of the frequency of recurrence of in-work poverty spells would report higher frequency for those who have experienced shorter spell of poverty, since it is not possible to test over a longer horizon of years what happens to these individuals. Hence, we chose to consider only those who are present over the 2007-2017 period for no fewer than four waves.

We include both the left-censored and the right-censored spell of poverty or non-poverty. This implies that the exact length of the first spell is unknown because the start date might have occurred prior to the beginning of the observed time window. In a similar way, poverty and non-poverty spells may be still in progress at the end of the period. In this case, the duration of these spells could be longer than the interval between the spell start date and the last wave in which the individual is observed.

4.4   Cross-sectional analysis

The poverty problem has long been characterised in terms of the number of people poor at a particular point in time alongside with their socio-demographic characteristics. Following this approach, the aim of this Section is to give a picture of the cross-sectional in-work poverty in the Municipality of Bologna in the period 2007-2017 using tax data. First, evidence for the population is shown, then breakdowns by subgroups follow.

When computed with respect to a local poverty line, the Municipality of Bologna is characterised by a high rate of in-work poverty: in 2017 the incidence of in-work poverty was 28.3 per cent versus 12.3 per cent in Italy and the risk of in-work poverty increased by 2.5 percentage points (pp) in the Municipality versus 1.2 pp in Italy

in the period 2012-2017. The high rate of in-work poverty recorded
in the Municipality of Bologna, using a local poverty line, follows
the direction of other works[4], in which poverty estimated with a lo-
cal threshold register significant different values than that estimated
with a poverty threshold calculated at the national level.

The evolution of the poverty threshold and the in-work poverty

Figure 4.1: Evolution of the poverty threshold and the corresponding in-work at
risk poverty rate (IWP) in the Municipality of Bologna for the period
2007-2017.
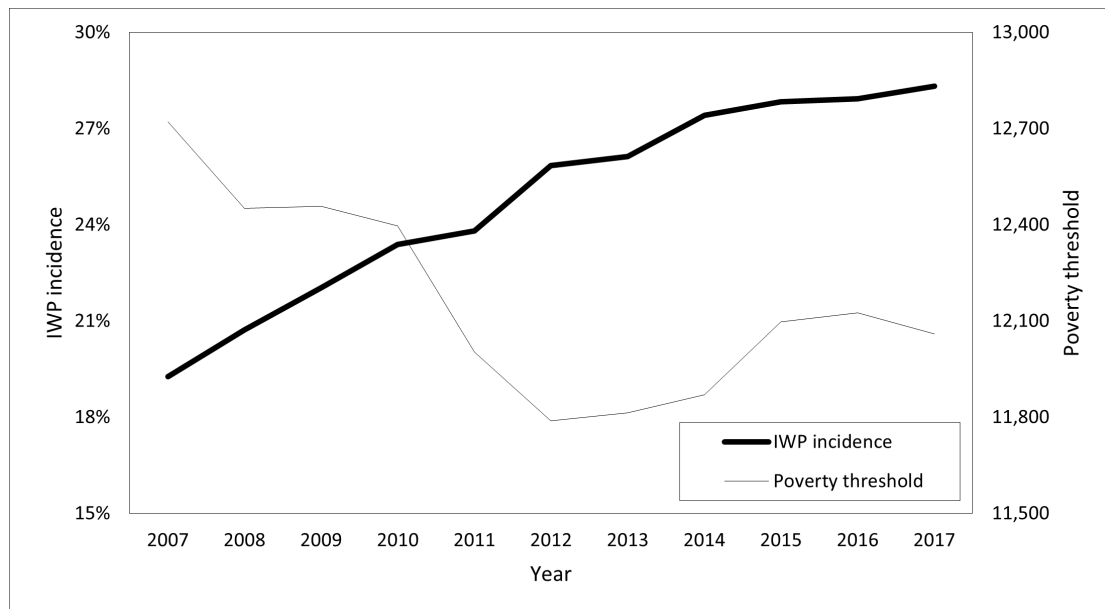
incidence is described in Figure 4.1 for the period 2007-2017. The
poverty threshold, in real terms, decrease from 12,720 Euro to 12,059
Euro between 2007 and 2017. In particular, in the first half it regis-
tered a consistent decrease until it reached the lowest value in 2012
and then, it slowly recovered in the second half; finally, in the last
year available, a further slight decline has been registered. On the

---

[4]To name but one, Chiappero-Martinetti and Pansini (2009) showed using EU-SILC data that
a sub-national line (equal to 60 per cent of the median equivalent income) led to an increase in
the estimate of poverty incidence in the area considered, i.e. Lombardia.

other hand, the incidence of the in-work poverty relative to the standard level of welfare of the local population increased by 9 pp in the period 2007-2017, moving from 19.3 to 28.3 per cent. In particular, the in-work poverty rate strongly increased between 2007 and 2012 (+6.5 pp), while in the second half of the observed period a modest increase was observed (+2.5 pp).

The comparatively high in-work poverty is mainly driven by the high mean income level of households in the Municipality of Bologna with respect to the Italian households. Indeed, in 2017, the national poverty threshold, computed as the 60 per cent of the median equivalent household income with modified OECD scale, was equal to 9.925 Euro, whereas the sub-national poverty threshold, computed with reference to the median equivalent income of the citizens of the Municipality of Bologna, was equal to 12,059 Euro. Therefore, if computed with respect to the national poverty threshold the in-work poverty incidence in the Municipality of Bologna would have been equal to 8.5 pp.

Referring to a wider classification of in-work poverty categories, useful to overcome the sharp dichotomy between the poor and the non-poor, which constitutes an oversimplification of reality, next to the standard poverty threshold, we computed the extreme poverty and the near-poverty thresholds. The first one, equal to the 80 per cent of the standard poverty line, identifies individuals who are in extreme in-work poverty conditions, as they have an equivalent income well below (20 per cent below) the standard poverty line, whereas the second, equal to the 120 per cent of the standard poverty line, makes it possible to identify individuals who, while not being poor, have an household equivalent income slightly higher (no more than

Figure 4.2: In work at risk poverty rate in the Municipality of Bologna computed with respect to the 80 and 120 percent thresholds for the period 2007-2017 (percentage values).
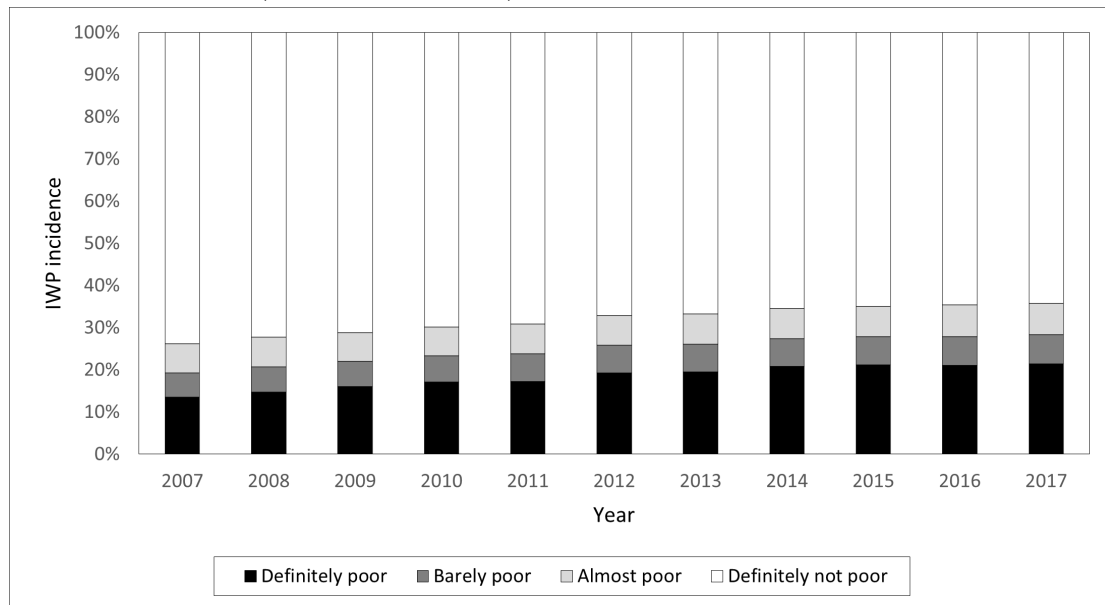
20 per cent) than the standard poverty line and, therefore, who may be at risk of falling into in-work poverty.

Figure 4.2 shows the evolution of the in-work poverty rate for the "definitely poor", i.e. workers in extreme in-work poverty conditions, the "barely poor", i.e. those with income between the 80 per cent and the standard poverty line, the "almost poor", i.e. those with income between the standard poverty line and 120 per cent and the "definitely not poor", i.e. those with income higher than 120 percent the standard poverty line. Evidence shows that the in-work poverty rate did not register significant changes for the "almost poor" category between 2007 and 2017, which are the 7.0 per cent of the total in 2007 and 7.4 per cent in 2017. Similarly, the "barely poor" category increased by 1.2 pp in the whole period examined, moving from 5.7 per cent in 2007 to 6.9 per cent in 2017. Therefore,

the major changes occurred in the "definitely poor" category, which increased by 7.9 pp (13.5 per cent in 2007 and 21.4 per cent in 2017). Consequently, the rate of individuals definitely not at risk of poverty decreased by 9.5 pp in the whole period, but they still remain the larger share of workers individuals in the Municipality of Bologna (73.7 per cent in 2007 and 64.3 per cent in 2017).

Moving to the analysis by subgroups, Table 4.1 shows, based on the cross-sectional analysis of in-work poverty, the incidence rate according to different socio-demographic characteristics of the in-work poor individuals in the Municipality of Bologna. Despite the gender income gap, in-work risk in 2017 was almost the same for female workers (29.1 per cent) than for male workers (27.6 per cent). Moreover, the in-work poverty risks for males and females increased by 8.9 pp and 9.2 pp, respectively, from 2007 to 2017.

Evidence also shows that the risk was lower for workers aged 55-64 (19.7 per cent in 2017 and 10.6 per cent in 2007), than for the two groups aged less than 55, which registered an in-work poverty rate of 30.0 per cent in 2017 and 21.0 per cent in 2007 for those aged 25-54 and 43.8 per cent in 2017 and 33.1 per cent in 2007 for those aged less than 25. This confirms the worrying situation of the youngest taxpayers in the Municipality of Bologna, whose median income registered significant negative variations during the observed period (see Chapter 3), which was accompanied by a reallocation of the taxable income that is decidedly in favour of to those over 55 years old. However, even those aged 55 or more have registered an increase in the in-work poverty rate between 2007 and 2017 of 9.1 pp, compared to 9.0 for those aged 25-54 and 10.7 for those aged less than 25.

Table 4.1: In-work at-risk poverty rate in the Municipality of Bologna by gender, age, citizenship, household type and number of income earners within the household and area of residence in the tax years 2007 and 2017 (percentage values).

Group	2007	2017
All individuals	19.3	28.3
Gender		
Women	19.9	29.1
Men	18.7	27.6
Age group		
Less than 25	33.1	43.8
25-54	21.0	30.0
Greater than 55	10.6	19.7
Citizenship		
Italian	14.7	20.9
Foreign	62.7	68.7
Household type		
Couples with children	13.1	21.7
Couples without children	9.0	14.9
Single-member households	28.2	33.5
Single-parent households	25.0	33.8
Others	23.1	34.9
No. of taxpayers in the household		
One	33.0	39.0
Two	13.6	21.4
Three of more	9.0	19.5
Area of residence		
Borgo Panigale	17.2	26.1
Navile	20.7	31.1
Porto	20.2	27.4
Reno	17.5	26.9
San Donato	22.9	33.1
San Vitale	20.4	29.0
Santo Stefano	18.6	25.9
Saragozza	19.9	27.9
Savena	16.2	26.1

Source: Data processing on MEF data.

Highly heterogeneous in-work poverty risks also emerge when we distinguish workers by citizenship, suggesting problems of labour market integration for non-Italian citizens. In-work poverty risks were indeed much greater for workers without Italian citizenship: in 2017, 68.7 per cent for foreign citizen and 20.9 per cent for Italians. Besides, the in-work poverty rate increased of 6 pp and 6.2 pp, for non-Italians and Italians, respectively.

Household characteristics are a crucial driver of in-work poverty. As expected, – given that the in-work poverty risk is closely associated with household composition – this risk is largely heterogeneous across household types, and rises when household needs (i.e. number and characteristics of members) are higher. In 2017, the incidence of in-work poverty was the highest for "Other households"[5], i.e. households with children and other cohabiting members (34.9 per cent). Then, single-parent workers with dependent children and single-member households follows, which registered an in-work poverty rate of 33.8 per cent and 33.5 per cent, respectively. At the bottom of the ranking, we find couples with children and couples without children, which have an in-work poverty rate well below the average rate for 2017 (21.7 and 14.9 per cent, respectively). An increase in the in-work poverty risks characterised all household types in the period 2007-2017; however, the increased was higher for households with children and other cohabiting members (12 pp) and lesser for households without dependent children (6 and 5 pp for couples and single-member, respectively).

As expected, given the characteristics of the in-work poverty indicator, the household number of workers is closely associated with the

---

[5]The "Other households" category includes all households with a couples or a single parent with children and other cohabiting members.

in-work poverty risk: the incidence of this risk varied in 2017 from 19.5 and 21.4 per cent among workers who lived in households with three or more taxpayers and two taxpayers, respectively, and 39 per cent among households with only one taxpayer. Basically, as the number of income-earners increases, also the economic conditions of the household improve. However, it has to be stressed that the working poverty is particularly high in the Municipality of Bologna even among those who live in households with a high number of workers, if compared to the average level of in-work poverty in Italy computed with a national threshold.

Moreover, consistent with a picture of territorial income difference within the Municipality of Bologna, which affects annual taxable incomes, evidence also shows that the in-work poverty rate varies across the neighbourhoods of the Municipality: Santo Stefano, which was the richest area during all the years examined, had the lowest in-work poverty incidence in 2017 (25.9 per cent), whereas San Donato, the most disadvantage neighbourhood, had the highest (33.1 per cent in 2017). However, the situation at the top of the ranking slightly changed since 2007, when Savena held the lowest in-work poverty incidence (16.2 per cent), followed by Borgo Panigale (17.2 per cent) and Reno (17.5 per cent).

Hence, the evidence summarized in this Section suggests that the relatively high in-work poverty characterising the Municipality of Bologna mainly depends on the interplay between the low wages received by a high proportion of workers and the low number of labour income recipients in many households.

4.5   Longitudinal analysis

Henceforth, the static approach of the previous Section is set aside in order to focus on the longitudinal perspective of in-work poverty in the Municipality of Bologna. Indeed, this Section is intended to investigate in-work poverty dynamics.

An interesting result of this research is the proposal of a new dimension in the measurement of in-work poverty at the local level which refers to the time individuals spend below the poverty threshold. Moreover, in the longitudinal analysis of in-work poverty, it is of great interest to be able to characterise the complete low-income pattern of individuals along time, i.e. the transitions to and from in-work poverty and the recurrence of the poverty spells. In fact, some individuals experience only a single, short spell of poverty, during an observation window, while others are caught in a poverty trap, which implies low income for an extended period. As we will see, the shares of in-work poverty persistence may vary substantially across individuals and family characteristics.

Hence, this Section analyses the dynamics of in-work poverty, including the duration, the frequency and recurrence of poverty spells, as well as the types of movements into and out of in-work poverty within the Municipality of Bologna during an observation window of eleven years. Differences in working poverty across population subgroups are also analysed.

4.5.1   The persistence of in-work poverty

This sub-Section focuses on the short-term versus long-term poverty persistence and how it is distributed across different subgroups of the

working population in the Municipality of Bologna. The empirical analysis here is based on the eleven-years panel data (2007-2017), which settle the length of the window for which working individuals can be followed over time.

The first interesting result presented here, which is useful to understand the phenomenon of poverty persistence, regards the duration of in-work poverty experiences during the period 2007-2017. It came out that the number of low-income workers in a given year is substantially different from the number of taxpayers touched by in-work poverty over several years: in fact, the latter is significantly higher than the cross-sectional incidence calculated over each year of the period analysed.

Some preliminary results on the number of years spent in poverty by each observation in our panel, pooling the data for all the poverty and non-poverty periods without considering the order of each occurrence, are presented in Table 4.2. The first result that attracts attention is the fraction of the population that had low income in at least one year within the 2007-2017 period, which is much higher than the average cross-sectional in-work poverty of the period measured in the previous Section. Results on spell duration, in fact, indicate that nearly 45.28 per cent of the sample is touched at least once by in-work poverty. Among those who were found to be poor at least once (columns 5 and 6 in Table 4.2), 19.94 per cent may be considered temporary poor, having stayed below the threshold for only one year; 27.62 per cent of those who are affected by poverty have incomes below the threshold for six or more years out of eleven, thus may be considered permanent poor. It should also be noted that relatively few people (1.24 per cent) were continuously poor in all the

Table 4.2: Number of years in the working poverty condition for individuals in the Municipality of Bologna in the eleven-years observation window (2007-2017).

		All individuals		*Poor at least once*	
No. of years in poverty (x)	Abs no. of poor	Percentage of poor	Proportion of poor at the end of x years	Percentage of poor	Proportion of poor at the end of x years
0	113,944	54.72	100.00	-	-
1	18,797	9.03	45.28	19.94	100.00
2	13,324	6.40	36.25	14.13	80.06
3	11,290	5.42	29.85	11.98	65.93
4	14,510	6.97	24.43	15.39	53.95
5	10,220	4.91	17.46	10.84	38.56
6	7,797	3.74	12.55	8.27	27.72
7	5,768	2.77	8.81	6.12	19.45
8	4,272	2.05	6.04	4.53	13.33
9	3,173	1.52	3.98	3.37	8.80
10	2,530	1.22	2.46	2.68	5.43
11	2,591	1.24	1.24	2.75	2.75
Total	208,216	100.0	0.0	100.0	0.0

Source: Data processing on MEF data.

eleven waves of the observed period, and thus configures a sort of hard core of long-term poverty. However, for most individuals, the risk of poverty seems distant, and, in fact, as many as 52.9 per cent are never touched by in-work poverty within the 2007-2017 period. In order to give a more straightforward comprehension of the results and aware of the potential difficulties of deriving this type of measure, we now set up a new variable with the aim to distinguish individuals by the number of years spent below the poverty threshold (see S. Jenkins, 2011), as:

1. Not poor in any of the eleven-years observation window;

2. Temporary poor: individuals who were poor at one wave;

3. Short-term persistent poor: those who experienced in-work poverty from 2 to 6 years over the eleven years period (not necessarily consecutive);

4. Long-term persistent poor: those who have experienced in-work poverty for more than 6 years over the period examined (not necessarily consecutive).

Moreover, we also consider an alternative longitudinal indicator of poverty, addressed as "chronic poverty"[6]. This additional measure is derived by computing an individual's longitudinal averaged income observed at some fixed number of consecutive years and comparing this average with the poverty threshold. An individual is described as chronically poor if his or her averaged income is below the standard poverty line, which is computed as the 60 per cent of the median local household income in the first wave of the observation window 2007-2017 for each individual. This approach has the advantage of accounting for changes in income that occur in the observation window considered by individuals experiencing transient fluctuations of income above and below the standard poverty threshold.

Table 4.3 summarizes the distribution of the two in-work poverty measures over the period 2007-2017 for the working individuals in the Municipality of Bologna. The first indicator is reported in columns 2-5, whereas the chronic poverty results is in column 6. The estimates for the population are reported in the first row, whereas the other number present the rates of in-work poverty persistence for a range of population sub-groups, where the group characteristics are defined in the first year of the observation window. More than

---

[6]For more details, see Rodgers and J.L., 1993; Chaudhuri and Ravallion, 1994; S. Jenkins et al., 2001.

Table 4.3: In-work poverty persistence measures for individuals in the Municipality of Bologna by gender, age, citizenship, household type, number of income earners in the household and area of residence over the eleven-years unbalanced panel observation window (2007-2017).

Group	Never	One-off	Short-term	Long-term	*Chronic poverty*
All individuals	54.72	9.03	23.70	12.55	25.62
Gender					
Women	54.48	9.06	23.73	12.72	26.00
Men	54.96	9.00	23.66	12.38	25.26
Age group					
Less than 25	36.59	12.05	37.61	13.76	36.98
25-54	52.52	9.22	24.21	14.05	26.82
Greater than 55	77.19	6.23	12.72	3.86	12.50
Citizenship					
Italian	62.77	9.44	19.43	8.36	16.88
Foreign	8.77	6.65	48.11	36.46	75.58
Household type					
Couples with children	64.40	8.35	17.65	9.61	16.61
Couples without children	72.49	7.34	14.40	5.77	12.39
Single-member households	44.68	9.81	30.32	15.18	35.46
Single-parent households	48.48	9.90	25.87	15.75	28.88
Others	46.91	9.36	27.96	15.76	31.59
No. of taxpayers in the household					
One	41.60	9.36	30.73	18.31	38.77
Two	62.48	8.65	19.26	9.62	18.19
More than two	65.31	9.17	18.67	6.86	14.29
Area of residence					
Borgo Panigale	56.53	8.82	22.87	11.78	23.68
Navile	52.13	9.14	24.65	14.08	28.56
Porto	54.00	9.69	24.36	11.95	25.67
Reno	56.50	8.66	22.45	12.39	24.48
San Donato	51.20	8.80	24.94	15.07	29.75
San Vitale	53.87	9.01	24.56	12.56	26.25
Santo Stefano	56.04	8.95	23.39	11.62	23.81
Saragozza	54.05	9.46	23.97	12.52	25.34
Savena	58.58	8.72	21.80	10.89	22.38

Source: Data processing on MEF data.

half of the individuals did not experience in-work poverty during the observed period, of the remaining half, 9 per cent experienced only one years of in-work poverty, 23.70 percent experienced between 2 and 6 years in poverty and 12.55 per cent suffered from poverty more than 6 years.

Non-Italian workers experienced the highest rate of persistent in-work poverty, both short-term and long-term (48.11 per cent and 36.46 per cent, respectively). At the second place for gravity of the situation, we find the younger taxpayers, i.e. those aged less than 25 years old, who reported values equal to 37.61 per cent for short term in-work poverty and 13.76 per cent for long term poverty. Another subgroup which reported dramatic values is that of households with only one income earners: 18.31 per cent of whom experienced working poverty for six or more years and more than 30 per cent between two and six years. Other groups with a pronounced rate of in-work poverty over the observed period were the single-member households, couples with children and other cohabiting members, single-parent households, those living in San Donato and Navile, and, to a lesser extent, workers aged 25-54 years old.

The estimates of chronic poverty (column 6) show that one-fourth (25.62 per cent) of the all the working individuals in the Municipality of Bologna registered a longitudinal averaged income over the period 2007-2017 below the poverty threshold. The subgroups that revealed the most dramatic situation, basically, are the same found with the alternative in-work poverty persistent measure.

The most disadvantaged group is that of non-Italian taxpayers, whose rate of chronic poverty reached the 75.58 per cent among the entirety of non-Italian taxpayers. Then, 38.77 per cent of the households

with only one income earners registered averaged low-income values below the threshold over the examined period. Besides, the other groups with a high valued of chronic poverty are younger taxpayers, single-member households and couples with children and other cohabiting members and, to a lesser extent, those living in San Donato.

On the other hand, individuals who appears to have low values of chronic poverty with respect to the entirety of the population are couples without children (12.39 per cent), those aged more than 55 years old (12.5 per cent) and households with three or more income earners (14.29 per cent).

These results should be useful for policy design by identifying the composition of the poverty population, including significant differences in individual and household characteristics between the short-term and the long-term working poor individuals.

4.5.2   Recurrent versus transitory in-work poverty

At this point of the analysis, it is important to investigate the frequency, i.e. the relevance of poverty spells recurrence. Indeed, this sub-Section focuses on the repetition of poverty spells for the overall population and some subgroups.

Longitudinal analyses of poverty dynamics showed that most poverty spells are short and transitory (see OECD, 2001; Arranz and Cantó, 2012). However, these short periods of poverty are usually recurrent. Since, apparently, most individuals who exit poverty in a given year will re-enter it within a short time frame, the analysis of repeated poverty spells help us to give a more inclusive overview of in work-poverty persistence in our sample. Our findings highlight the rele-

vance of the phenomenon of repeated spells of poverty, i.e. the fact that, in many cases, individuals report incomes below the poverty line in non-consecutive years, determining, in other words, exits and returns to the in-work poverty status.

Table 4.4 focuses on the frequency distribution of elapsed poverty spells for the population and across several subgroups, where the group characteristics are defined in the first wave in which the individual recorded the poverty status. Evidence shows the importance of considering multiple spells in the analysis of poverty dynamics: out of the 94,272 individuals who experienced at least one poverty spell, 27.65 per cent re-entered in the working poverty twice or more along the respective observation window, of which 23.04 percent have two occurrences and 4.62 per cent have three or more occurrences. No deep differences were observed across the population subgroups. However, those that experienced the highest rates of spells recurrence are the couples with children (31.00 per cent), households with more than two income earners (31.56 per cent for three or more taxpayers and 30.64 for two taxpayers), younger workers (30.74 per cent) and, to a lesser extent, couples with children and other cohabiting members (29.47 per cent). On the other hand, the groups who registered the lowest values for spells recurrence are those aged greater than 55 years old (15.79 percent), couples with children (25.13 per cent), families with only one income earner (25.17 percent) and single-member households (25.23 percent).

These results suggest that the in-work poverty spells repetition was high both for the overall population and for the subgroups without distinction. However, it appeared that the great majority of individuals experienced only a single poverty spell over the period

Table 4.4: Recurrent and transitory in-work poverty spells in the Municipality of Bologna by gender, age, citizenship, household type, number of income earners in the household and area of residence over the eleven-years unbalanced panel (2007-2017).

	No. of recurrence of poverty spells		
Group	**One**	**Two**	**Three or more**
All individuals	72.35	23.04	4.62
Gender			
Women	73.66	22.10	4.24
Men	71.06	23.96	4.98
Age group			
Less than 25	69.26	26.21	4.53
25-54	71.05	23.90	5.05
Greater than 55	84.21	14.16	1.62
Citizenship			
Italian	72.19	23.03	4.78
Foreign	72.71	23.06	4.23
Household type			
Couple with children	69.00	25.13	5.87
Couple without children	74.87	21.05	4.08
Single-member households	74.77	21.50	3.74
Single-parent households	72.35	22.94	4.71
Others	70.52	24.47	5.00
No. of taxpayers in the household			
One	74.83	21.22	3.94
Two	69.36	25.15	5.49
More than two	68.43	26.11	5.45
Area of residence			
Borgo Panigale	71.01	23.52	5.47
Navile	72.49	23.02	4.49
Porto	72.21	23.14	4.65
Reno	71.10	24.03	4.87
San Donato	72.59	22.63	4.77
San Vitale	72.18	23.08	4.74
Santo Stefano	73.35	22.51	4.15
Saragozza	72.71	23.01	4.29
Savena	72.47	22.86	4.67

Source: Data processing on MEF data.

2007-2017.

4.5.3 Poverty exit and re-entry rates

We now move to the analysis of the exit and entry rates from in-work poverty using non-parametric estimates depending on the duration of the spells spent in and out of in-work poverty. The aim of this sub-Section is, basically, to measure the relevance of spell duration on the probability of leaving or re-entering poverty for working individuals. First, we define the number of people who start a poverty (non-poverty) spell in year t as the number of individuals not poor (poor) at time $t-1$, who are poor (not poor) at time t. Given that the exit rates are those that relate to people who have just fallen into in-work poverty, so they are at risk not to escape from it and, analogously, the re-entry rates refer to people who have just finished an experience of in-work poverty, so they are at risk of re-entering it, the panel used for this analysis will include only individuals who begin a spell of poverty or non-poverty in his or her second wave within the observed period (2007-2017).

Following other works (Devicienti, 2002; Devicienti and Gualtieri, 2004), the approach we have employed for the following analysis exclude left-censored spells but, on the other hand, include right-censored spells, i.e. the spells whose end date (transition out of the current state) had not yet occurred at the end of the observed period and so, the spell end date is unknown and the total length of time between entry to and exit from the state is unknown. This implies that the individuals who were experiencing in-work poverty in the last wave of the panel contribute to the estimated exit rates. As a result, in our data set, having eleven waves available, the exclusion

of left-censored spells implies that only individuals starting in the second wave are to be considered. In this way, since exits from poverty or non-poverty status can only occur in one of the waves following the one in which the individual first enters the spell, so each spells of poverty or non-poverty can be observed from one to at most ten years.

The method used to examine the probability of entering and exiting

Table 4.5: Survival function and exit rates from in-work poverty by duration for all individuals beginning a poverty spell.

Number of years since start of the spell	At risk of exit at the start of the period	Survival Function	Std Error	Exit Rate	Std Error
1	107742	1.00	-	0.138	0.001
2	65058	0.863	0.001	0.085	0.001
3	41082	0.789	0.001	0.065	0.001
4	25735	0.738	0.002	0.056	0.001
5	15789	0.697	0.002	0.049	0.002
6	9406	0.663	0.002	0.045	0.002
7	5243	0.633	0.003	0.039	0.003
8	2689	0.608	0.003	0.045	0.004
9	1125	0.581	0.004	0.030	0.005
10	304	0.563	0.005	-	-

poverty at a point in time is based on the idea that the number of people entering and exiting poverty is obtained by calculating changes in each individual poverty status across two years.

Suppose $t_1 < t_2 < \cdots < t_j < \cdots < t_k$ represent the survival times observed in the dataset. We can determine the following quantities:

- a_j: the number of individuals observed to fail, i.e. to make a transition out of the state, at time t_j;

- b_j: the number of individuals whose observed duration is cen-

Table 4.6: Survival function and re-entry rates to in-work poverty by duration for all individuals beginning a non-poverty spell.

Number of years since start of spell	At risk of enter at the start of the period	Survival Function	Std Error	Re-entry Rate	Std Error
1	149081	1	-	0.075	0.001
2	103246	0.925	0.001	0.046	0.001
3	72409	0.883	0.001	0.033	0.001
4	50077	0.854	0.001	0.028	0.001
5	33795	0.830	0.001	0.026	0.001
6	21853	0.808	0.001	0.024	0.001
7	13187	0.788	0.002	0.019	0.001
8	7244	0.773	0.002	0.020	0.002
9	3383	0.758	0.002	0.017	0.002
10	1084	0.745	0.003	-	-

Source: Data processing on MEF data.

sored in the interval $[t_j, t_{j+1})$, i.e. they are still in state at time t_j but not at time t_{j+1};

- c_j: the number of individuals at risk of ending their spell immediately prior to t_j, which is made up of those who a have a censored or completed spell of length t_j or longer:

$$c_j = (a_j + b_j) + (a_{j+1} + b_{j+1}) + \cdots + (a_k + b_k).$$

The exit rate is computed as the number of individuals who make a transition out of the state divided by the number of individuals at risk of exit, i.e. the exit rate at time t_j are computed as $e_j = a_j/n_j$. On the other hand, the survival probability, which is the proportion of those entering a state who survive to the previous observed survival time, is computed simply as the product of one minus the proportion who made a transition out of the state by that

time, i.e. one minus the exit rates at each of the survival times:

$$\hat{S}(t_j) = \prod_{j|t_j<t} \left(1 - \frac{a_j}{n_j}\right).$$

According to Table 4.5 and Table 4.6, which show the the estimates of survival probabilities and hazard rates, both types of spells show a decline of the transition hazards as duration evolves. There is evidence of negative duration dependence for both the type of spells: the longer an individual stays in poverty (out of poverty) the less likely it is that he or she will leave that state in the next period.

However, some differences can be observed between the exit and re-entry hazard rates. First, the probability of re-entering in-work poverty is significantly lower than the probability of exiting from the in-work poverty status. This should imply that the length of non-poverty spells is of longer duration than that of in-work poverty. Secondly, both the exit and the re-entry hazards registered a rapid decline during the first years and, then, it slowly declined on wards as the duration of spell increased.

In particular, Table 4.5 shows for the group of individuals just starting a poverty spell that about one in ten (13.8 percent) managed to exit after the first year; after five years, the probability of exiting poverty is nearly 5 percent.

On the other hand, Table 4.6 presents the survival function and the re-entry rate for those who just finished a poverty spell and so at risk of re-entering the in-work poverty status. The re-entry rates are nearly halved than exit rate: 7.5 per cent of those just starting a non-poverty spell will re-enter after the first years, after five years the probability of re-entering is 2.6 per cent and after nine years the

probability of re-entry poverty is less than 2 per cent.

If we compare these results to those of Devicienti and Gualtieri (2004) for Italy in the period 1994-2001, we find negative duration both in exit and re-entry rates as well, but the hazard rate values are of greater size: the chance of leaving poverty after one year has been estimated at 0.58 for Italy, and the chance or re-entering poverty after one year is 0.25. However, comparing these results is difficult for several reasons, apart from the territorial coverage: we exploited an administrative dataset instead of survey data, we employed a different poverty definition and a local threshold and, also, the time periods covered by the two analysis are not overlapping.

As a general pattern, as observed in OECD (2001), the exit rate is lower if the annual poverty rates are higher, as well as lower cross-sectional poverty is associated with less poverty persistence. As a matter of fact, lower exit rate should generate long duration for poverty spell: however, the actual average of waves in poverty is 3.5 years estimated for an individual just entered in the samples and followed for the next waves both in our sample and in Devicienti and Gualtieri (2004) for Italy.

4.5.4 Stayers and changers

This sub-Section focuses on the analysis of the factors that may cause the entrance or escape from the in-work poverty status. We attempt to examine the link between household related events and individual poverty transitions.

A major issue that the researchers face when doing subgroup breakdowns of longitudinal analysis of in-work poverty poverty is that the characteristics of an individual used to classify him or her into a

subgroup may themselves change over time. Therefore, an individual's subgroup membership may differ over the period of time over which we observe his or her income. The following analysis is based on the idea that the risk of a poverty transition between two years may be associated with "trigger events", i.e. events that are a type of change in the characteristics of the individual that can vary over time (see Bane and Ellwood, 1986; S. Jenkins, 2000). Therefore, we would like to assess which events, based on the disposable time varying variables of our dataset, could have an impact on the route to or from in-work poverty. Previous analyses have shown that some of the factors that may influence the most the poverty status are the household type and the number of taxpayers in the family.

Following S. Jenkins et al. (2001), we may define "stayers" to be those individuals who had the same in-work poverty status in both the first and last year of the eleven-years observation window, and "changers" to be those individuals who had different levels of income in the first and last wave. In an analogous fashion, we identify stayers and changers using the household type and the number of income earners in the family.

One interesting aspect of this analysis is reported in the transition matrix between the poverty and non-poverty condition from the first to the last year of the observed period (2007 and 2017). For the sake of homogeneity, reference is made only to individuals who declared income in all the years in between 2002 and 2017. Therefore, for this analysis, 81,203 observation-years are considered[7].

---

[7]The use of this balanced panel implies an important caveat for the analysis results: the statistics reported may be subject to attrition bias since they are calculated from a sample consisting of individuals present in eleven consecutive waves of the corresponding panel dataset. However, what it is important for us is to show the possible tie between a change in the household composition and the poverty status.

In Table 4.7, we use the four-way in-work poverty classification

Table 4.7: Transition matrix between poverty and non poverty condition for working individuals in the Municipality of Bologna between the tax years 2007 and 2017 (percentage values).

	IWP status in 2017			
IWP status in 2007	*Definitely poor*	*Barely poor*	*Almost poor*	*Definitely not poor*
Definitely poor	41.01	13.22	11.46	34.31
Barely poor	24.42	14.84	15.57	45.17
Almost poor	17.30	10.57	14.54	57.59
Definitely not poor	5.38	3.06	4.51	87.05

Source: Data processing on MEF data.

reflecting the potential trajectories of each individual. On the main diagonal of the matrix, we find the stayers[8]. Evidence shows that, except for the "definitely not poor" who presented a very stable situation (87.05 per cent after eleven years had not change their status), the other categories registered fluctuations over time, demonstrating a rather transitory nature of the classes of individuals with income below or close to the standard poverty line. For "almost poor" category, i.e. those individuals whose income is slightly above the standard poverty line, only 14.54 per cent remained in the same class after eleven years, nearly 30 per cent worsened their economic condition by slipping into the classes below the standard poverty line, and the remaining improved their status by joining the "definitely not poor" (57.59 percent). As far as "barely poor" households are concerned, 14.84 per cent remained in the same class, more than 60 per cent managed to place themselves in the class of a higher order and the remainder (24.42 percent) worsened their status, ending up in the "definitely poor" category. Finally, the situation of the most

---

[8]These individuals may have experienced different transitions over time and then, in the last wave returned to the initial year status.

disadvantage individuals in 2007 is remarkable since 41.01 percent of the individuals still had income below the 80 per cent poverty threshold in 2017. It is surprising, however, that of the remaining 59 per cent, 34.31 per cent managed to make the leap and go into the class of "definitely not poor".

We now provide some evidence about the rate of individuals who

Table 4.8: Transition matrix between poverty and non-poverty condition for working individuals in the Municipality of Bologna by household type between the tax years 2007 and 2017 (percentage values).

Household type in 2007	Household type in 2017				
	Couple with children	*Couple without children*	*Single-member household*	*Single-parent household*	*Others*
Status improved					
Couple with children	71.76	5.65	5.97	8.53	8.08
Couple without children	32.42	44.92	13.35	1.91	7.42
Single-member household	14.43	6.56	57.98	4.74	16.29
Single-parent household	12.04	1.95	16.67	54.89	14.45
Others	21.59	6.92	13.15	8.86	49.48
Status worsened					
Couple with children	52.73	6.97	13.25	16.78	10.27
Couple without children	31.80	42.40	11.31	5.89	8.60
Single-member household	20.28	4.57	51.22	8.79	15.14
Single-parent household	9.10	2.02	29.31	44.34	15.23
Others	20.17	5.94	15.09	16.40	42.40

Source: Data processing on MEF data.

changed household type and number of taxpayers in the family from the first and the least year of the observation window, together to information about how these rates differed between individuals for whom a positive change occurs and those who experienced a negative change in the economic condition. Evidence shows that nearly 46 per cent changed the structure of the family they belonged to. This fact

is remarkable since it could be related to an improvement or a worsening in the income condition of the individuals who experienced it. From Table 4.8, it is possible to grasp the existing dynamics for the type of households and their temporal transformations for individuals who improved their status and for those who did not improve it, separately. Interesting evidence is that the percentage of people who changed their household type was greater for the group of individuals whose in-work poverty status worsened between 2007 and 2017 than for those who improved it (51.30 versus 40.09 per cent). Some interesting patterns also emerged. Individuals who worsened their in-work poverty status reported higher percentage of changes from household types that should have more than one income earner to households type in which there should be a single income earner.

The same reasoning can be repeated analysing changes over time

Table 4.9: Transition matrix between in-work poverty condition and not in-work poverty condition for individuals followed by all the period in the Municipality of Bologna by number of taxpayers in the household between the tax years 2007 and 2017 (percentage values).

	No. of taxpayers in 2017		
No. of taxpayers in 2007	*One*	*Two*	*Three or more*
Status improved			
One	53.33	38.88	7.79
Two	17.98	58.82	23.21
Three or more	14.73	43.89	41.38
Status worsened			
One	69.46	27.79	2.74
Two	34.18	54.32	11.50
Three or more	40.62	39.98	19.40

Source: Data processing on MEF data.

in the number of income earners within the household. Evidence shows that 46.5 per cent of the individuals have changed the num-

ber of taxpayers within the household to which they belong to (see Table 4.9). In this case, the number of individuals who changed their income earners situation did not differ between the two categories: the rate is equal to 45.8 per cent for those who improved the in-work poverty status versus 47.1 per cent for those who did not improve it. For individuals who improved the economic condition between 2007 and 2017, the percentage of those who belonged to households who reduced the number of income earners were widely fewer than those who increased this number (28.58 per cent versus 71.42 per cent). On the other hand, the share of households that reduced the number of taxpayers among those whose status worsened between 2007 and 2017 is the 66.44 per cent of families who experienced a changed in the income earners composition.

Given these results, it turns out that clear patterns are hard to discern and draw clear conclusions about what may cause the change in the in-work poverty status is not straightforward. However, these results may reflect the positive effect on income of an increase in the number of workers in the family and, on the other hand, the negative effect on income of a decrease in the number of taxpayers. This also seems to confirm the difficulty that one-earner households face in avoiding the in-work poverty in a society where two-earner households have increasingly become the norm. Moreover, data may reflect the positive effects on income associated with partnership formation.

Chapter 5

Final Remarks

In this work we defined a socio-economic profile of the citizens residing in the Municipality of Bologna, exploring the relationship between income distribution and some important socio-demographic variables. To this purpose, we exploited newly available administrative data on individual income tax returns linked to data from the municipal Registry Office.

The main results of this project can be summarised as follow. In Chapter 2, we compared administrative with surveys data, highlighting both the strengths and weaknesses of the two sources. It emerged that our dataset enabled a complete coverage of the population, providing for the lack of official survey-based statistical information at the municipal, as well as sub-municipal, level.

Moreover, it appeared that our dataset was much better at capturing very high earners in comparison with data from the IT-SILC sample survey. The fact that through administrative data is possible to examine the appropriate share of income of the richest is a very important aspect, since it enhances the analysis of inequality, allowing to observe the trends relating to the high tail of the distribution with a much greater precision than that ensured by sample surveys.

The latter, in fact, due to their limited number of observations and the high rate of non-response of the rich to the interviews, run the risk of not adequately representing high-income individuals in the sample.

However, our archive also displayed some limitations, mainly related to the income definition. A potential drawback of working with administrative data is that the target population is based on the legislation requirements needed for the registration of a specific program. In fact, our database did not include individuals exempted to submit the tax return, i.e. those who collected only non-taxable income or whose amount was under an exemption threshold. The lack of information about non-taxable incomes, such as social, maternity, attendance, and pension allowances, as well as contribution for rentals and household expenses, affects the estimation of disposable income. Indeed, these components has a relatively important weight, especially in the left tail of the income distribution. Integrating information about non-taxable incomes from other administrative sources available would be useful to improve income estimate.

Moreover, another drawback displayed by our dataset concerns the continuity. Changes in the legislation in force over the years influenced taxable income definition, making the cross-sectional comparison complicated. We attempted to solve this problem considering income from municipal additional tax, for which empirical evidence suggested a greater continuity.

The analyses of Chapter 3 confirmed the widespread idea that the national average data can hide very different situations at the local level. Two very evident aspects in the data were the progressive, albeit still partial, recovery of the relative positions of women com-

pared to those of men and the strong re-composition of the income distribution from young and adult individuals to the advantage of older ones. In our opinion, attention should be paid to this type of evidence: the re-composition in progress, in fact, foreshadow the emergence over time of new demographic and economic figures who appear to be more at risk of poverty and, therefore, who deserve more attention. This aspect, in a context of scarce financial resources and very modest growth in taxable income, poses to the policy maker questions of complex solution that the recent Covid-19 crisis can only have exacerbated.

At the territorial level, the gap in the average and median level of taxable income between the northern and the southern areas of the Municipality has not smoothed out. San Donato appeared to have the most dramatic income condition with respect to the other neighbourhoods.

The condition of non-Italian taxpayers was even more dramatic. In real terms, their income modestly decreased since the beginning of the period: this led to a slight increase in the already existing gap with the Italian citizens. In 2017, the foreigners declared an average income around 50 per cent lower than that of the Italian counterpart. From an age point of view, the presence of foreign citizens was numerically greater in the younger classes. In addition, in 2017, under 25 surprisingly gained, albeit slight, higher median income than the Italian young taxpayers.

From the analysis of household income, it was possible to better qualify gainers and losers. Couples without children were those who have seen their income increase the most. On the other hand, households with children and other cohabiting members were subject to

a reduction in equivalent income. Given that the first group include mainly retired individuals, whereas the second mainly consist of younger age individuals, this split allowed us to better qualify the generational issue already evident in the analysis of individual data. The degree of incomes dispersion, calculated with the Gini index on equivalent household income, modestly decreased if computed both before and after the personal income tax. On the other hand, the income gap between rich and poor slightly increased. In 2017, the disposable income of the richest 10 per cent was around four times higher than that of the poorest 10 per cent. Splitting the analysis of income disparities between the upper end and the middle, and between the middle and the lower end of the income distribution, evidence showed that there is more inequality among the poor than among the rich. Moreover, the analysis of high-income earners showed that the share of women in the top groups rose. Nevertheless, they still were under-represented and the degree of under-representation increased as one approaches the top of the distribution. On the contrary, younger taxpayers worsened their already disadvantaged situation at the top of the income distribution. Their degree of representation in the top decile nearly halved between 2002 and 2017.

Finally, the examined period brought changes in tax regulations, which allowed the local administration to increase the taxation and to introduce the tax exemption threshold for low-income taxpayers. This implied a considerable increase in the total revenue from the municipal surtax, when compared to national accounts. The analysis of the redistributive and progressivity effects of the tax system showed different impacts of the reforms. However, overall, both the

effects registered an albeit little increased.

Our results of Chapter 4 showed an high incidence of in-work poverty in the period 2007-2017, measured at cross-sectional level with a local threshold. Specifically, it appeared to be related to the interplay between low-income workers, mainly younger and non-Italian citizens, and the limited number of income recipients within the household. This suggests the need of measures to improve both individual earnings and conditions in the labour market, but also measures to increase the number of income earners within the family.

From the longitudinal analysis of in-work poverty, it came out that the groups with the most dramatic experience of persistent in-work poverty were non-Italian workers, younger taxpayers, and households with only one income earner. On the other hand, the subgroups that appeared to ward off poverty were couples with children, older taxpayers, and households with three or more income earners. This upheld the need, already evident from the previous analysis, to implement measures, or to strengthen the existing ones, to support young workers and non-Italian citizens.

Finally, the analysis of transitions to and from poverty confirmed the importance of partnership formation on income positive transitions. Moreover, it proved once again the difficulty that one-income earner households faced in avoiding poverty with respect to families with two or more taxpayers.

In conclusion, some points may deserve future investigation. First, the overall income estimates could be improved if it were possible to enhance the information about non-taxable incomes, who escape from our dataset. Similarly, it would be interesting to integrate data with other existing forms of taxation, which account for a decrease

in the disposable income.

Secondly, it would be desirable to continually update the archive year by year as tax data become available, with the involvement of producers and suppliers of the data.

Third, it would be interesting to extend the longitudinal analysis of in-work poverty exploiting the discrete-time duration models to investigate which households and individual variables are associated with transitions into and out of poverty at the local level (see Devicienti, 2002; Devicienti and Gualtieri, 2004).

Finally, it would be extremely interesting to reproduce the dataset and the analysis throughout the Country with the aim of comparing income and welfare conditions across the municipalities.

Appendix A

The Italian personal income tax

The personal income tax plays a central role in the tax systems of all advanced economies, both as the share on total tax revenues and for its impact on the redistribution of income.

It was born and developed in the second part of the 19th century in the most important European and North American countries. It flanked and gradually replaced previous systems where the various income sources (capital, labour, etc.) were taxed separately, becoming the primary source of public revenues.

In the Italian tax system, the Italian personal income tax (abbreviated as Irpef) is a personal and progressive income tax, whose aim is to make the tax system progressive and, at the same time, to achieve a consistent redistribution of income[1].

Any form of income of resident and non-resident individuals produced in Italy is subject to the tax. The tax base is constituted by the sum of all incomes owned by the taxpayer. Once the single income components have been defined (see Table A.1), they must be added algebraically to determine the total income[2].

---

[1]For further details about the milestones of Irpef from its birth to present-day, see Pellegrino and Panteghini, 2020.

[2]For a detailed description of how the total taxable income for Irpef is built and the characteristics of each income component that contribute to its formation, see Appendix B.

Table A.1: Summary of income components and criteria for the definition of the tax base for the Italian personal income tax.

Income category	Definition of the tax base
Real Estate Income	
Land income	
- Dominical income	Revalued cadastral appraisal rate
- Agricultural income	Revalued cadastral appraisal rate
Buildings income	
- Main residence	Revalued cadastral income (deduction from total income equal to the annuity)
- Buildings not leased in the Municipality of the main residence	50% of the cadastral income increased by 1/3
- Others buildings not leased	Excluded from the tax base
- Buildings leased	95% of actual income
Capital income	Actual income
Income from dependent work	Actual income
Income from self-employment	Actual income net of production costs (identified analytically or on a flat-rate basis depending on the type of income)
Business income	Effective income: economic profit with positive or negative variations due to differences between civil and fiscal criteria (e.g. capital gains, depreciation, surplus, interest payable)
Other incomes	
Financial capital gains and income from derivative products	Income net of capital losses
Other incomes	Mixed criteria

Source: Bosi and Guerra, 2018.

Deductions are then subtracted from the total income to obtain the taxable income (see Table A.2). The most important deductions are the compulsory social security and welfare contribution of self-employed, and the cadastral income from the property used as main residence. Other forms of deduction include for example the periodic allowance to the spouse and the contribution paid to collective supplementary pension funds.

Table A.2: Operating scheme of the personal income tax in the Italian tax system.

Total income
- Deductions
= Taxable income
↓
Tax rate scale application
↓
= Gross tax
- Allowances
= Net tax
- Tax credits
- Advance payment of tax
= Owed tax

Source: Bosi and Guerra, 2018.

A tax rate scale is then applied to the tax base according to a progressive scale scheme in order to obtain the gross tax. Since 1998, regions and municipalities are allowed to apply a flat or progressive surtax to Irpef. They may vary according to general rules defined by the central government.

In the next step, allowances are subtracted from the gross tax so, the net tax is obtained. The allowances can be grouped into five categories: allowances by source of income, which include the "80 Euro" bonus, for family loads, for personal expenses, for rents and first home loans, and other allowances. They help to define the progressiveness of the tax and allow distinctions of the tax burden according to the source of income. They treat differently individuals who belong to families with different composition and, therefore, dissimilar contributory capacity. Moreover, they personalize the tax and encourage honest behaviour. However, the possibility of using them presupposes the presence of a gross tax debt that is at least as large as the amount of deductions to which the taxpayer is entitled.

Otherwise, the phenomenon of insufficiency of deductions occurs[3].

Once the net tax has been calculated, the taxpayer can deduct a set of tax credits from the tax owed to the State. Generally, for these tax credits the problem of insufficiency is not present as the unused annual quota can be carried forward in the declarations of subsequent tax periods without any time limit.

Subsequently, the taxpayer deducts the withholding taxes and obtains the tax to be paid. In case that the withholding incurred are greater than the tax due, the taxpayer can request a refund of the amount he/she has paid in excess, or he/she can use this credit towards the tax authorities to offset the payment of other taxes.

The relationship between the taxpayer and the tax authority begins with a declaration which is followed by a payment made either by the withholding agent or by the taxpayer himself. The Irpef declaration takes place annually and the deadline is standardized as of 30 September of the subsequent year. The submission of declarations is possible in three different ways. The 770 Form must be used by withholding agents and by the State Administrations for the declaration of the withholding made, as well as the other required contribution and insurance data. The 730 Form is mainly used by employees and retirees who, compared to what is declared in the 770 Form by their employer, must change their position with respect to the tax authority (i.e. they have more than one income, they have non-employment income and/or they have expenses which give them the right to deductions and allowances). Finally, the remaining part of taxpayers, mainly self-employed, is required to submit the tax return using the Individual Income Form (UPF)[4].

---

[3]For further details about deductions and allowances, see Bosi and Guerra, 2018
[4]For further details about the procedures to declare the taxable income, see Appendix C.

Appendix B

The definition of the taxable income

In the Italian tax system, the identification of the total taxable income is illustrated in the law income tax - "Testo Unico delle Imposte sui Redditi" (abbreviated TUIR) - which was introduced into the legislative system by the D.P.R. 22 December 1986, n. 917.

The tax base is defined as the sum of all incomes owned by the taxpayers for resident individuals, and only incomes produced in the territory of the State for non-residents (art. 3, *TUIR*, 1986).

Incomes making up the tax base are classified into six categories (art. 6, *TUIR*, 1986):

1. real estate income;

2. capital income;

3. income from dependent work;

4. income from self-employment;

5. business income;

6. other incomes.

Real estate incomes (art. 25, *TUIR*, 1986) are land and buildings incomes located in the territory of the State. They are in turn distinguished in dominical incomes, agricultural incomes, and buildings incomes.

Dominical incomes (art. 27-28, *TUIR*, 1986) are constituted by the dominical part of the average ordinary income which can be drawn from the land through the exercise of agricultural activities. Adjacent lots of urban buildings, lands rented for non-agricultural uses, as well as those producing business income are not considered as an integral part of the dominical income.

Agricultural incomes (art. 32-34, *TUIR*, 1986) correspond to the part of the average ordinary income of the land attributable to the working capital and to the organization work employed in the exercise of agricultural activities, within the limits of the potentiality of the land.

Both incomes are determined on the basis of cadastral appraisal rates. The cadastral rates refer to the cadastral income attributed to each parcel into which the agricultural land has been divided. The cadastral income represents the average, ordinary and continuous yield obtainable from a parcel.

Buildings incomes (art. 36-37, *TUIR*, 1986) refer to the average ordinary income that can be drawn from urban real estate units. It is determined by applying the appraisal rates, established according to the rules of the cadastral law for each category, or by direct estimation for buildings with special use.

Income from buildings includes both figurative and actual income. The first group include the owner's house and/or those kept available. If the building is the main residence, the taxable income is the

cadastral income revalued to 5 per cent. Income from main residence contributes to the determination of the taxpayer's total income but is then deductible from this one. Since 2013, the real estate income relating to units kept available (second home or home given for free use) are excluded from the taxable income, unless they are in the same municipality where the property used as main residence is located. In the latter case they are increased by a third and included in personal income tax for 5 per cent of the resulting amount. Actual incomes include income from buildings leased to third parties. They are reduced by 5 per cent to take the maintenance and management costs of the property into account. Incomes from rented flats are included in the definition of total income if the taxpayer opts for the ordinary regime. Alternatively, since 2011, if the taxpayer opts for a flat rate tax regime[1], they are excluded. However, income subject to the flat rate tax are included in the definition of income used for the calculation of effective deductions for workloads and families.

Capital incomes (art. 44-45, *TUIR*, 1986) are defined as any income from a relationship which implies the use of a capital. Specifically, they include:

- interest and other profits from mortgages, deposits and current accounts;

- interest and other profits from bonds and similar securities, and from securities other than shares and similar securities, as well as mass certificates;

---

[1]The flat rate tax ("cedolare secca") is a tax system introduced by art.3, L. D. n. 23/2011. It allows individuals to submit the income from rents to a substitute tax which has a rate equal to 21 per cent, or lower for certain type of contracts. For example, for those arranged in municipalities with a high population density was 15 per cent in 2013 and, then, 10 per cent since 2014.

- never-ending annuities and annual benefits;

- remuneration for performance of sureties or other guarantees;

- profits from the participation in companies and entities subject to the corporate income tax;

- profits from partnership agreements;

- income from the management, in the collective interest of a plurality of subjects, of assets made up of sum of money and assets entrusted by third parties or from related investments;

- interest and other incomes from other relationships having as their object the use of the capital, excluding relationships through which positive and negative differentials can be realized depending on an uncertain event.

However, it should be emphasized that capital incomes have a moderate significance in the context of personal income tax because they are generally subject to a substitute tax system or, in the case of individual savings plans, are exempt from taxation.

Incomes from dependent work (art. 49-50, *TUIR*, 1986) are those from a relationship with object the job performance, with any qualification, employed by and under the direction of others. They include home business when considered as dependent work according to the rules of job legislation.

The following categories are those defined to be "assimilated" into income from dependent work:

- the remuneration received by member workers of production, service, small fisheries, agricultural and primary processing of

agricultural products cooperatives, within the limits of the current salaries increased of 20 per cent;

- benefits and payments received by third parties for employed job, with the exception of those that by contractual clause must be paid to the employer and those that must be paid to the State;

- amount paid as a scholarship or allowance, award or subsidy for study or professional training purposes, if the beneficiary is not linked by dependent work relationships towards the financier;

- amount received during the tax period for whatever reason, also in the form of donations, in relation to the offices of manager, mayor or auditor; associations and other entities, with or without personality legal status; collaborations with newspapers, magazines or encyclopaedia; participation in committee, as well as those received in relation to other collaborative relationships without subordination constraints, within the regulatory framework of a unitary and continuous relationship without the use of organized means and with established periodic remuneration;

- the remuneration of priests, as well as congrua and their supplements;

- the reimbursement for free professional intramural activities of the employees of the National Health Service, of the personnel referred to in art. 102, D.P.R. 11 July 1980, n. 382 and of the personnel referred to in art. 6, paragraph 5, D.L. n. 502 and subsequent amendments;

- allowances, attendance fees and other remuneration paid by the

State, regions, provinces and municipalities for the exercise of public functions, as well as the remuneration paid to the members of tax commissions, the justices of peace and the experts of monitoring tribunal, excluding those that must be paid to the State;

- the benefit received by members of the National and European Parliament and the allowances, however denominated, received for the elected offices and functions referred to in artt. 114 and 135 of the Constitution and the L. 27 December 1985, n. 816, as well as the consequent life allowances received depending on the termination of the aforementioned elective offices and the allowance of the President of the Republic;

- annuities and fixed-term annuities, other than those having a social security function;

- pension benefits indicate in D. L. 21 April 1993, n. 124, however disbursed;

- other social allowances, however denominated, whose production does include neither capital nor job;

- the remuneration received by subjects engaged in community service works in compliance with specific regulatory provisions.

Incomes from self-employment work (art. 53-54, *TUIR*, 1986) are those from the exercise of arts and professions, i.e. the exercise by usual profession, even if not exclusive, of self-employment activities other than those included into business incomes. Self-employment incomes also include:

- income from the economic use by the author or inventor of an intellectual property, industrial patent and process, formula or information related to the experience acquired in the industrial, commercial or scientific field, if not obtained in the exercise of commercial enterprises;

- profit sharing, when the contribution is constituted exclusively by the work performance;

- profit sharing due to the promoters and founding partners of joint stock company, in a partnership company with limited responsibility;

- benefit for the ending of agency relationships;

- income from the activity of raising protests exercised by the municipal secretaries.

The income from the exercise of arts and professions is constituted by the difference between the amount of payment in cash or in kind received in the tax period, also in the form of profit sharing, and the expenses incurred in the same period for the exercise of the art or profession. The payments are calculated net of the social security and welfare contributions established by law to be paid by the taxpayers.

Business incomes (art. 55-56, *TUIR*, 1986) are those within the scope of commercial companies. Business income includes the following categories:

- income from the exercise of an activity organized in the form of a business aimed at providing services that are not part of the art. 2195 of the Civil Code;

- income from the exploitation of mines, quarries, peat bogs, salt mines, lakes, ponds and other inland waters;

- income from land, when derived from the exercise of agricultural activities referred to in art. 32, albeit within the limits established therein, where they belong to a general or limited partnership, as well as to an organization of non-resident individuals carrying out a business activity.

Business income tax base is defined as the net profit (or loss) resulting from the profit and loss account.

Other incomes (art. 67, *TUIR*, 1986) include types of income, mainly capital gains, not attributable to the categories previously examined. In this regard, it is important to distinguish between capital gains from the use of financial capital and incomes from derivative products, which are generally taxed, and other capital gains, which are taxed with homogeneous criteria only in cases identified in detail by the legislator. The first category, complementary to capital incomes, includes:

- capital gains related to shareholdings;

- capital gains related to non-equity securities, currencies and precious metals;

- incomes earned through derivative contracts (swaps, options, futures, etc.) and other fixed-term contracts.

In the second category, capital gains are realized on:

- allotment of land and subsequent sale, even partial, of land or buildings here erected;

- transfer upon payment of properties acquired by inheritance and donation, and of those used as main residence by the transfer or his family, as well as capital gains realised from the transfer upon payment of building lands.

Finally, other heterogeneous incomes are included in this category:

- winnings from lotteries, prize contests, games and bets organized for the public, prizes from skill tests or by chance, as well as those awarded for artistic, scientific or social merits;

- real estate income that cannot be determined by land registry, including those of land rented out for non-agricultural uses;

- real estate income from property located abroad;

- income from the economic use of intellectual property, industrial patents and processes, formulas and information related to experience acquired in the industrial, commercial or scientific field, enjoyed by the heirs of the author or inventor;

- income from the usufruct and subletting of real estate units; from the rent of vehicles, machines and other mobile properties; from the rent and usufruct of businesses. The lease and the usufruct of the sole company by the entrepreneur are not considered to be made in the exercise of the business, but in the event of a subsequent total or partial sale, the capital gains realized form the total income as different income;

- income from commercial or self-employment activities, not normally exercised;

- travel allowances and reimbursements received by individuals who carry out amateur sports activities.

Appendix C

The procedures for declaring the taxable income

We report the general rules for the submission of the tax returns to the Revenue Agency updated to 2018 and referred to 2017 fiscal year, given the complexity of the Italian tax system and the considerable amount of changes incurred to the tax laws in the period 2002-2017.

Model 770

Modello 770 (2018) is a document produced by withholding agents, which must present it in order to declare the remuneration paid to individuals and/or companies with whom they had employment relationships. It must be submitted electronically by enterprises, public and private institutions and self-employed, when they assume the role of withholding agents.

The requirement to fill in this form depends on incomes subject to taxes: the types included are mainly income from dependent work and "assimilated", and retirement benefits.

The model contains data related to the certifications issued to the subjects who received income from employment and "assimilated", severances, benefits in the form of capital provided by pension funds, commissions, and other incomes. It also contains contribution, social security and insurance and tax assistance data.

This form has to be submitted also by the subjects who paid a sum exposed to withholding tax on capital gains, compensation for commercial goodwill, contributions to public and private entities, release from life insurance contracts, premiums, winnings and other financial income, including those from participation in undertakings for collective investment in transferable securities under foreign laws, profits and other income from investments in corporations, atypical securities, and other income, as well as those who have paid sums or values subject to withholding tax.

The subjects are:

- corporations resident in the territory of the State;

- commercial entities equivalent to joint stock companies resident in the territory of the State, i.e. public and private entities having as exclusive or main goal the exercise of commercial activities;

- non-commercial entities resident in the territory of the State, i.e. public entities, which also include regions, provinces, municipalities, and private entities not having as exclusive or main goal the exercise of commercial activities;

- not recognized associations, consortia, special companies established in accordance with the artt. 22 and 23, L. 8 June 1990, n. 142;

- companies and entities of any kind, with or without legal personality, not resident in the territory of the State;

- trusts;

- condominiums;

- partnerships (simple, general and limited partnerships) resident in the territory of the State;

- armament companies resident in the territory of the State;

- de facto or irregular companies resident in the territory of the State;

- companies or associations, without legal personality, resident in the territory of the State, set up between individuals for the exercise in associated form of arts and professions;

- marital companies resident in the territory of the State, when the activity is carried out in a partnership between spouses;

- European economic interest groups (GEIE);

- natural persons who practice commercial or agricultural businesses;

- natural persons who exercise arts and professions;

- state administrations, including those with an autonomous system, which operate the withholding in accordance with the art. 29, D.P.R. n. 600/73;

- heirs not carrying on the deceased withholding agent's business.

Model 730

The *Modello 730* ([2018](#)) has to be submitted by taxpayers which received income from lands and buildings, income from dependent work and similar (e.g. continuous and coordinated contractual relationships and project work contracts), some category of income from self-employed work for which the VAT number is not required and other type of income (e.g. income from land and buildings located abroad and income subject to separate taxation). This form is used by taxable persons who, compared to what was declared in 770 by their employer, must change, for some reason, their position with respect to the tax authorities. Examples are the taxpayers who own more than one income or they want to declare any expenses incurred or benefit from deductions and/or to request reimbursements relating to credits or over payment resulting from the tax returns presented in the previous years or from deposits already paid for the year in question.

The taxpayer must check whether he/she is obliged to submit the tax return or whether he/she is exempt. Taxpayers exonerated are those who owned the following incomes:

- income from main residence, related adjacent lots (for which IMU[1] is not owed), and other not rented buildings, when not located in the same municipality as the main residence;

- income from dependent work or retirement benefits, possibly together with main residence, related adjacent lots (for which IMU is not owed) and other not rented buildings, when not located in the same municipality as the main residence;

---

[1] "Imposta Municipale Unica" is an Italian direct tax applied to the real estate component of the assets.

- income from continuous and coordinated contractual relationships including project works, with the exception of administrative collaborations of a non-professional nature made in favour of companies and non-professional sports associations;

- exempt income (e.g. annuities paid by Inail[2] for permanent disability or death, some scholarships, war pensions, ordinary pensions for conscripts, attendance allowances to civilian blind, deaf and legally disabled, and social pensions);

- income subject to substitute tax, different from those subject to the flat tax on dividends;

- income from lands and buildings when do not exceed 500 Euro (including income from the main residence and related adjacent lots);

- income from dependent work, "assimilated" or retirements benefits, together with other incomes, when do not exceed 8,000 Euro, and the working or retirement period must not be less than 365 days (the income from the main residence and related adjacent lots must not be taken into account);

- retirement benefits when do not exceed 7,500 Euro (the retirement period must not be less than 365 days), together with income from land when do not exceed 185.92 Euro, main residence and related adjacent lots (for which IMU is not owed);

- periodic allowances paid by the spouse, in case together with other incomes, not exceeding 8,000 Euro (the income from main

---

[2] "Istituto Nazionale per l'Assicurazione contro gli Infortuni sul Lavoro" is an Italian non-economic public entity who manages compulsory insurance against accidents at work and occupational diseases, subject to the supervision of the Italian Ministry of Labor and Social Policies.

residence and related adjacent lots must not be taken into account, and periodic allowances for the child support are excluded);

- incomes "assimilated" to dependent work and other incomes when do not exceed the total amount of 4,800 Euro (e.g. remuneration received for freelance intramural activity of personnel employed by the National Health Service, income from occasional commercial activities and income from occasional self-employment);

- income from amateur sports activities when do not exceed 28,158.28 Euro;

- taxpayers not obliged to keep accounting records, who owns a total income for which the gross tax, net of deductions for the progressive taxation and for dependent family members, do not exceed 10.33 Euro.

Model UPF

Modello UPF (2018) is mainly used by self-employment workers. Some categories of dependent workers and "assimilated" also need to submit this form. Specifically, UPF form must be submitted when the taxpayer:

- has income from self-employment for which a VAT number is required, business income, including income from participation and trusts, and/or other income not included among those declarable with 730 form;

- has income from self-employment work to which art.50, D.P.R.

22 December 1986, n. 917 is applied (e.g. members of artisan cooperatives);

- has income from the production of "agro-energy" over a limit set by the D. L. 24 April 2014, n. 66;

- has capital gains from the sale of qualified or non-qualified shareholdings in companies resident in countries or territories with privileged taxation, whose titles are not traded on regulated markets;

- is not resident in Italy in the previous year and/or in the year in which the return has to be submitted;

- uses tax credits for income produced abroad;

- must also submit other declarations (e.g. VAT, IRAP[3]);

- must file the return on behalf of deceased taxpayers;

- has changed employer, only if the tax exceeds of more than 10.33 Euro the total withholding taxes;

- has received by INPS[4] or other Entities allowances and salary integration, or when the withholding taxes have not been made due to an error, or if the conditions of exemption indicate for the Model 730 do not hold;

- has not been allocated the correct deductions;

- has received wages and/or income from individuals not required to make withholding taxes (e.g. family collaborators or drivers);

---

[3] "Imposta Regionale sulle Attività Produttive" is an Italian tax on industry established with the D. L. 15 December 1997, n. 446.

[4] "Istituto nazionale della previdenza sociale" is the Italian main entity of the public retirement system.

- has income on which the tax is applied separately;

- has not paid the correct amount of the municipal and/or regional surtax, only if the surtax exceeds 10.33 Euro.

Bibliography

Acciari, P., & Mocetti, S. (2013). Una mappa della diseguaglianza del reddito in Italia. *Politica economica, 28*(3), 307–343.

Ahrendt, D., Sándor, E., Revello, A., Jungblut, J.-M., & Anderson, R. (2017). In-work poverty in the EU (Eurofond, Ed.).

Allison, P. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology, 13*, 61–98.

Alvaredo, F., Atkinson, A., Piketty, T., & Saez, E. (2013). The top 1 percent in international and historical perspective. *Journal of Economic Perspectives, 27*(3), 3–20.

Alvaredo, F., & Pisano, E. (2010). Top incomes in Italy, 1974-2004. In Oxford University Press (Ed.), *Top incomes: A global perspective.*

Arranz, J., & Cantó, O. (2012). Measuring the effect of spell recurrence on poverty dynamics - evidence from Spain. *The Journal of Economic Inequality, 10*(2), 191–217.

Atkinson, A., Casarico, A., & Voitchovsky, S. (2016). Top incomes and the gender divide (International Inequalities Institute, London School of Economics and Political Science, Ed.). *5.*

Atkinson, A., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature, 49*(1), 3–71.

Baldini, M. (2021). Redistribution and progressivity of the Italian personal income tax, 40 years later. *Fiscal Studies*, 1–25.

Baldini, M., & Ciani, E. (2011). Inequality and poverty during the recession in Italy. *Politica Economica, 3*, 297–322.

Baldini, M., & Pellegrino, S. (2016). *Chi restituisce il bonus di 80 euro. e perché.* https://www.lavoce.info/archives/41418/chi-restituisce-il-bonus-di-80-euro-e-perche/

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

Bane, M. J., & Ellwood, D. T. (1986). Slipping into and out of poverty: The dynamics of spells. *The Journal of Human Resources, 21*(1), 1–23.

Barbetta, G., Pellegrino, S., & Turati, G. (2018). What Explains the Redistribution Achieved by the Italian Personal Income Tax? Evidence from Administrative Data. *Public Finance Review, 46*(1), 7–28.

Barbieri, P., Cutuli, G., & Scherer, S. (2018). Caratteristiche, dinamiche e persistenza dell'in-work poverty in Italia. *Liw working paper series, 1.*

Beaumont, J.-F. (2020). Are probability surveys bound to disappear for the production of official statistics? *Survey Methodology, 46*(1), 1–28.

Bethlem, J. (2009). The rise of survey sampling. *Dalton Transactions, 15.*

Biewen, M. (2006). Who are the chronic poor? an econometric analysis of chronic poverty in Germany. In *Dynamics of inequality and poverty* (pp. 31–62).

Biggeri, L., & Pratesi, M. (2017). Monetary poverty indicators at local level: Definitions, methods of estimations and comparisons in real term. *International Statistical Institute Congress, Morocco.*

Bishop, J. A., Lee, J. M., & Zeager, L. A. (2017). Incorporating spatial price adjustments in U.S. public policy analysis (ECINEQ - Society for the Study of Economic Inequality, Ed.). *438.*

Bonanno, G., Chies, L., & Podrecca, E. (2019). Poverty and material deprivation dynamics in Italy. *DEAMS Research Paper Series, 2.*

Bonhomme, S., & Hospido, L. (2013). Earnings inequality in Spain: New evidence using tax data. *Applied Economics, 45*(30), 4212–4225.

Bosi, P., & Guerra, M. C. (2018). *I tributi nell'economia italiana* (Il Mulino, Ed.).

Brackstone, G. J. (1987). Issues in the use of administrative records for statistical purposes. *Survey Methodology, 13*, 29–43.

Calzaroni, M. (2011). Le fonti amministrative nei processi e nei prodotti della statistica ufficiale. *ISTAT - Istituto Nazionale di Statistica.*

Cantó, O. (2002). Climbing out of poverty, falling back in: Low income stability in Spain. *Applied Economics, 34*, 1903–1916.

Carrozzi, P., & Rauco, G. (2019). Rapporto statistico sull'area metropolitana romana 2019.

Chaudhuri, S., & Ravallion, M. (1994). How well do static indicators identify the chronically poor? *Journal of Public Economics, 53*(3), 367–394.

Chiappero-Martinetti, E., & Bottiroli Civardi, M. (2002). Povertà nei gruppi e povertà tra i gruppi. In FrancoAngeli (Ed.), *Studi sulla povertà. problemi di misura e analisi comparative* (pp. 65–86).

Chiappero-Martinetti, E., & Pansini, R. V. (2009). Condizione economica, indebitamento e condizione abitativa delle famiglie italiane tra benessere e povertà (Guerini Associati, Ed.). *ORES, L'esclusione sociale in Lombardia. Rapporto 2008*, 71–94.

Ciani, E., & Torrini, R. (2019). The geography of Italian income inequality: Recent trends and the role of employment (Economic Research and International Relations Area, Bank of Italy, Ed.). *Questioni di Economia e Finanza - Occasional Papers, 492*.

Coli, A. (2012). Matching survey and administrative data for studying poverty at the local level. *Rivista Internazionale di Scienze Sociali, 120*(1), 107–115.

Commissione d'Indagine sull'Esclusione Sociale. (2003). Rapporto sulle politiche contro la povertà e l'esclusione sociale.

Comune, M. (2010–2011). *Diseguaglianze dei redditi e povertà delle famiglie attraverso dati amministrativi. un'indagine longitudinale nel Comune di Brescia* (PhD Thesis). Università Cattolica del Sacro Cuore, Milano.

Comune, M. (2014). Un archivio longitudinale amministrativo per la stima della povertà a livello locale. *Rivista di Statistica Ufficiale*, (1-2/2014).

Comune di Milano. (2005). Ricerca sullo stato economico occupazionale delle famiglie milanesi. un nuovo strumento di conoscenza: Il progetto AMeR-IcA. *Milano dati - Serie economia e lavoro*, (6).

Couper, M. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*(4), 464–494.

Couper, M. (2013). Is the sky falling? new technology changing media, and the future of surveys. *Survey Research Methods, 7*(3), 145–156.

Devicienti, F. (2002). Estimating poverty persistence in Britain. *LABORatorio R. Revelli - Centre for Employment Studies, 1*.

Devicienti, F., & Gualtieri, V. (2004). Dinamiche e persistenza della povertà in Italia: Un'analisi con microdati panel di fonte ECHP. *LABORatorio R. Revelli - Centre for Employment Studies*.

Devicienti, F., Gualtieri, V., & Rossi, M. (2014). The persistence of income poverty and lifestyle deprivation: Evidence from Italy. *Bulletin of Economic Research, 66*(3), 246–278.

Devicienti, F., & Poggi, A. (2011). Poverty and social exclusion: Two sides of the same coin or dynamically interrelated processes? *Applied Economics*, *43*(25), 3549–3571.

Di Caro, P. (2018). Redistribution in real-world PIT: Evidence from Italian tax records (Ministero dell'Economia e delle Finanze, Italy, Ed.). *2*.

Di Caro, P. (2019). Decomposing personal income tax redistribution with application to Italy. *Journal of Economic Inequality*, *18*(1), 113–129.

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*, *2*, 125–157.

Di Nicola, F., & Pellegrino, S. (2017). *Bonus 80 euro, i nodi al pettine con i rinnovi dei contratti.* https://www.lavoce.info/archives/49483/bonus-80-euro-nodi-al-pettine-rinnovo-dei-contratti/

Duncan, G., Gustafsson, B., Hauser, R., Schmauss, G., Messinger, H., Muffels, R., Nolan, B., & Ray, J. (1993). Poverty dynamics in eight countries. *Journal of Population Economics*, *6*(3), 215–234.

Elliott, M., & Valliant, R. (2017). Inference for nonprobability samples. *Statistical Science*, *32*(2), 249–264.

European Commission. (2011). European statistics code of practice – revised edition 2011.

European Commission. (2017). *Methodological guidelines and description of EU-SILC target variables.* Eurostat.

Eurostat. (2003). *Definition of quality in statistics.*

Ferrante, M. R., & Fabrizi, E. (2020). *Stima della serie temporale 2008-2017 di paramentri di povertà, deprivazione e diseguaglianza per i distretti socio-sanitari e per le provincie della Regione Emilia-romagna.* https://statistica.regione.emilia-romagna.it/documentazione/pubblicazioni/documenti_catalogati/stima-poverta-2008-2017-distretti-sociosanitari-province-emilia-romagna

Finnie, R., & Sweetman, A. (2003). Poverty dynamics: Empirical evidence for Canada. *Canadian Journal of Economics*, *36*(2), 291–325.

Fransham, M. (2017). Income and population dynamics in deprived neighbourhoods: Measuring the poverty turnover rate using administrative data. *Applied Spatial Analysis and Policy*, *12*(2), 275–300.

Giarda, E., Baldini, M., & Olivieri, A. (2015). A tax-benefit microsimulation model for Italy: A partial evaluation of fiscal consolidation in the period 2011-2014. https://ssrn.com/abstract=2665547

Giarda, E., Baldini, M., Olivieri, A., Pellegrino, S., & Zanardi, A. (2015). Il "Bonus" degli 80 euro: Caratteristiche ed effetti redistributivi. *Rivista di Diritto Finanziario e Scienza delle Finanze, LXXIVS*, 3–22.

Giarda, E., & Moroni, G. (2018). The degree of poverty persistence and the role of regional disparities in Italy in comparison with France, Spain and the UK. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement, 136*(1), 163–202.

Golini, A. (2001). Alcune rapide riflessioni in tema di indagini longitudinali nel demografico e nel sociale (Istat, Ed.). *Rivista di Statistica Ufficiale, 1*, 75–77.

Hand, D. J. (2018). Statistical challenges of administrative and transaction data. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 181*, 555–605.

Hick, R., & Lanau, A. (2018). Moving in and out of in-work poverty in the UK: An analysis of transitions, trajectories and trigger events. *Journal of Social Policy, 47*(4), 661–682.

Iceland, J. (1997). The dynamics of poverty spells and issues of left-censoring. *PSC Research Report Series, 378*.

Jarvis, S., & Jenkins, S. (1997). Low income dynamics in 1990s Britain. *Fiscal Studies, 18*(2), 123–142.

Jenkins, S. (2011). *Changing fortunes: Income mobility and poverty dynamics in Britain* (Oxford University Press, Ed.).

Jenkins, S. P. (1997). Discrete time proportional hazard regression. *Stata Technical Bulletin, 7*(39), 22–23.

Jenkins, S. (2000). Modelling household income dynamics. *Journal of Population Economics, 13*(4), 529–567.

Jenkins, S., Rigg, J., & Devicienti, F. (2001). The dynamics of poverty in Britain (Department for Work and Pensions and University of Essex, Institute for Social and Economic Research, Ed.). *Dwp Research Report Series*.

Johnson, B., & Moore, K. (2005). Consider the source: Differences in estimates of income and wealth from survey and tax data. *Compendium of Federal Estate Tax and Personal Wealth Studies, 2*.

Kakwani, N. C. (1977). Measurement of tax progressivity: An international comparison. *Economic Journal, 87*(345), 71–80.

Kalton, G. (2019). Developments in survey research over the past 60 years: A personal perspective. *International Statistical Review, 87*(S1), S10–S30.

Kapteyn, A., & Ypma, J. Y. (2007). Measurement error and misclassification: A comparison of survey and administrative data. *Journal of Labor Economics, 25*(3), 513–551.

Kim, J., & Tam, S.-M. (2020). Data integration by combining big data and survey sample data for finite population inference. *International Statistical Review, 89*.

Kohut, A., et al. (2012). *Assessing the representativeness of public opinion surveys*. http://www.%20people-press.org/2012/05/15/assessing-the-representativenessof-public-opinion-surveys/

Kopczuk, W., & Saez, E. (2004). Top wealth shares in the United States, 1916–2000: Evidence from estate tax returns. *National Tax Journal, 57*(2), 445–487.

Künn, S. (2015). The challenges of linking survey and administrative data. *IZA World of Labor, 214*.

Larrimore, J., Mortenson, J., & Splinter, D. (2020). Presence and persistence of poverty in U.S. tax data (National Bureau of Economic Research, Inc, Ed.). *Working Paper Series*.

Lee, S., & Valliant, R. (2009). Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods & Research, 37*(3), 319–343.

McKernan, S.-M., & Ratcliffe, C. (2002). Transition events in the dynamics of poverty. *The Urban Institute*. https://ssrn.com/abstract=2205860

Meng, X.-L. (2018). Statistical paradises and paradoxes in big data (i): Law of large population, big data paradox, and the 2016 US presidential election. *The Annals of Applied Statistics, 12*(2), 685–726.

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*(2), 176–204.

Miller, P. (2017). Is there a future for surveys? *Public Opinion Quarterly, 81*(S1), 205–212.

Ministero dell'Economia e delle Finanze. (1979). *Le dichiarazioni dei redditi delle persone fisiche presentate nel 1977.*

Ministero dell'Economia e delle Finanze. (2017). *Dichiarazioni dei redditi persone fisiche (Irpef) e dichiarazioni iva per l'anno di imposta 2017.*

Ministero dell'Economia e delle Finanze. (2018). *Relazione sull'economia non osservata e sull'evasione fiscale e contributiva – anno 2018.*

Modello 730: Istruzioni per la compilazione. (2018). https://www.agenziaentrate.gov.it/portale/documents/20143/355338/link+a+Istruzioni+mod+15022018_730_2018_istruzioni.pdf/5f9828cc-eb98-c687-9b0c-46b812f06a1d

Modello 770: Istruzioni per la compilazione. (2018). https://www.agenziaentrate.gov.it/portale/documents/20143/355744/Link+Istruzioni+770+16012018_770_istr_2018.pdf/aa3a1263-7efc-4a06-68d6-2019e27be84d

Modello UPF: Istruzioni per la compilazione. (2018). https://www.agenziaentrate.gov.it/portale/web/guest/archivio/archivioschedeadempimento/schede-adempimento-2018/dichiarazioni/redditi-persone-fisiche-2018/modello-e-istruzioni-redditi-pf2018

Muffels, R., Fouarge, D., & Dekker, R. (2000). Longitudinal poverty and income inequality a comparative panel study for the Netherlands, Germany and the UK (OSA Institute for Labour Studies, Ed.). *MPRA Paper, 13298.*

Neyman, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society, 97*(4), 558–625.

Oberski, D. L., Kirchner, A., S., E., & Kreuter, F. (2017). Evaluating the quality of survey and administrative data with generalized multitrait-multimethod models. *Journal of the American Statistical Association, 112*(520), 1477–1489.

OECD. (2001). When money is tight: Poverty dynamics in OECD countries. *Employment Outlook, Parigi, Ch.2.*

OECD. (2015). *Income inequality: The gap between rich and poor* (Publishing, Paris, Ed.).

Palamenghi, M., Riva, L., & Trentini, M. (2005). Criteri e metodi di stima del reddito delle famiglie bresciane (Dipartimento Metodi Quantitativi, Università degli Studi di Brescia, Ed.). *247.*

Palma, J. (2011). Homogeneous middles vs. heterogeneous tails, and the end of the "inverted-U": It's all about the share of the rich. *Development and Change, 42*(1), 87–153.

Pattaro, S., Bailey, N., & Dibben, C. (2020). Using linked longitudinal administrative data to identify social disadvantage. *Social Indicators Research, 147*, 865–895.

Pellegrino, S., & Panteghini, P. M. (2020). Le riforme dell'IRPEF: Uno sguardo attraverso 45 anni di storia. *Società Italiana di Economia Pubblica, 85*.

Pellegrino, S., Piacenza, M., & Turati, G. (2012). Assessing the distributional effects of housing taxation in Italy: A microsimulation approach. *CESifo Economic Studies, 58*(3), 495–524.

Penneck, S. (2007). Using administrative data for statistical purposes. *Economic and and Labour Market Review, 1*(10), 19–23.

Piketty, T. (2003). Income inequality in France, 1901-1988. *Journal of Political Economy, 111*(5), 1004–1442.

Piketty, T. (2014). *Capital in the twenty-first century.* Harvard University Press.

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913-1998. *The Quarterly Journal of Economics, 118*(1), 1–39.

Piketty, T., & Saez, E. (2006). The evolution of top incomes: A historical and international perspective (National Bureau of Economic Research, Ed.). *11955*.

Raitano, M., M., J., Pavolini, E., & Natili, M. (2019). ESPN thematic report on in-work poverty (European Social Policy Network, Ed.).

Rao, J. N. K. (2005). Interplay between sample survey theory and practice: An appraisal. *Statistics Canada, 31*(2), 117–138.

Rao, J. (2020). On making valid inferences by integrating data from surveys and other sources. *Sankhyā B: The Indian Journal of Statistics, 83*, 242–272.

Rao, J., & Fuller, W. (2017). Sample survey theory and methods: Past, present and future directions (with discussion). *Survey Methodology, 43*(2), 145–160.

Regione Emilia-Romagna. (2019). *Redditi, povertà ed esclusione sociale in Emilia-Romagna. anno 2017.* https : / / statistica . regione . emilia - romagna . it / documentazione / pubblicazioni / documenti_catalogati / redditi - poverta - esclusione-sociale-2017

Rodgers, J., & J.L., R. (1993). Chronic poverty in the United States. *Journal of Human Resources*, *28*(1), 25–54.

Rubolino, E. (2019). The efficiency and distributive effects of local taxes: Evidence from Italian municipalities. *ISER Working Paper Series*, *02*.

Ruspini, E. (2001). *La ricerca longitudinale. metodologia delle scienze umane* (FrancoAngeli, Ed.).

Signore, M., Brancato, G., Carbini, R., D'Orazio, M., & Simeoni, G. (2012). Quality guidelines for statistical processes - version 1.1. *Italian National Institute of Statistics - Istat*.

Smith, T. (1983). On the validity of inferences from non-randomsamples. *Journal of the Royal Statistical Society. Series A (General)*, *146*(4), 393–403.

Stevens, A. (1999). Climbing out of poverty, falling back in: Measuring the persistence of poverty over multiple spells. *The Journal of Human Resources*, *34*(3), 557–588.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future* (W. W. Norton & Company, Ed.).

Tomarelli, F., & Acciari, P. (2011). Global measures of tax progressivity, tax incidence and the redistribution operated by the Italian Personal Income Tax (2001 – 2009). *CIAT – Tax Administration Review*, *31*, 114–129.

Trapeznikova, I. (2019). Measuring income inequality. *IZA World of Labor*, *462*.

TUIR: D.P.R. 22 december 1986, n. 917 - Testo Unico delle Imposte sui Redditi. (1986).

United Nations Economic Commission for Europe. (2011). *Usign administrative and secondary sources for official statistics - a handbook of principles and practices*.

Verrecchia, F. (2019). *Dati amministrativi, metodi e statistiche per le politiche territoriali* (FrancoAngeli, Ed.).

Yucel, R., & Zaslavsky, A. (2005). Imputation of binary treatment variables with measurement error in administrative data. *Journal of the American Statistical Association*, *100*(472), 1123–1132.