Department of Economics and Finance, Chair of Advance Econometrics

DECOMPOSING THE CHANGE IN WAGE INEQUALITY: A COUNTERFACTUAL ANALYSIS ON ITALIAN DATA

Supervisor

Prof. Giuseppe Ragusa

Candidate

Siria Angino

ID 641841

Co-supervisor

Prof. Giovanna Vallanti

Academic year 2012/2013

To all my teachers,

in the broadest sense.

Abstract

Economic inequality is a relevant issue under many aspects: growth concerns, policy orientation, social dynamics. The aim of this paper is, first, to establish if there has been a change in wage inequality for Italian employees between 2000 and 2010, and second, to decompose this eventual variation into a wage structure and a composition component. The decomposition method used is the one proposed in Chernozhukov, Fernandez-Val and Melly (2009). We find that inequality remained nearly constant, except when considering the lower end of the distribution: in this case inequality strongly decreased. We also perform the detailed decomposition for the wage structure effect.

Keywords: wage inequality, counterfactual distribution, decomposition analysis.

Contents

1 Introduction

> For every one that had shall be given, and he shall have abundance: but from him that had not shall be taken even that which he had.
>
> Matthew 25:29

Inequality starts being a complex issue from the very moment of its definition. This word has the capacity to trigger different ideas and reactions in the reader's mind. Little by little, we are going to clarify the issue at stake and what we are interested in through this paper.

First, we deal with economic inequality. It concerns the uneven allocation of resources among the participants of an economy - considered as individuals or groups - or among economies themselves. The focus of this paper is "intra-country" inequality, that is, within countries, and completely neglects the inter-country one.

It is appropriate to immediately clarify that economic inequality and poverty are distinct concepts, although they are deeply intertwined and it is often very hard to disentangle their effects. Inequality is a broader concept, that takes into account the whole distribution of resources, while poverty considers only participants endowed with a level of wealth below a certain threshold, the poverty line. Stated in different terms, inequality is a relative measure while poverty is an absolute one.

Similarly, 'inequality' is not a synonym for 'unfairness'. The first term concerns resource allocation in a given state of the world, the latter is a judgment about the morality of such state.

Of course, economic inequality is a fertile ground for ethical discussions. Opinions differ about the morality of inequality but, especially when disproportionate, it is generally considered undesirable and unfair. Inequity aversion studies[1] show that people have a preference for fairness and resistance to incidental inequalities. Humans are sensitive to inequities in favor[2] as well as against[3] them, tend to reject unequal allocations and are

[1] E. Fehr, K. M. Schmidt, A Theory of Fairness, Competition and Cooperation, 1999, and M. I. Norton, D. Ariely, Building a better America, 2011

[2] In the Dictator game, one person chooses how to split the reward between himself and another participant to the experiment. Under the standard economic assumptions, the "dictator", the first to move, should decide to keep everything, but he/she often allocates money to the other participant, reducing the amount of money they receive. The results seem to demonstrate that people are sensitive to inequalities even if they are in their favor. Behavioral economic approach thus suggests to include in one's utility function also others' utilities.

[3] In the Ultimatum game, one person chooses again how to split the reward, but here the second

willing to sacrifice potential gains to stop another individual from receiving a superior reward[4].

Two extreme views on inequality can be identified. On one side, there is the idea that people should be given resources according to their needs in order to ensure that everyone receives the same level of utility. If we consider people to be quite homogeneous with respect to basic needs, this statement roughly implies an equal distribution of resources (with more resources given, for example, to sick people, so that they can achieve the same satisfaction level of the others).

The opposite principle claims that people should be given resources according only to their merits. Under this view, economic inequality is simply the reflection of the fact that "men are born unequal".

Some scholars[5] in the "new egalitarianism" school emphasize the necessity of equal opportunity renovated at each generation: in this way transfers from the old generation to the new one do not threat the benefits of meritocracy in the medium/long run. According to this view, thus, equality does no longer concern removing class distinctions but rather equalizing life chances across generations.

A third way is J. Rawls's one. In 'Theory of Justice'[6], he states that inequalities in the distribution of wealth are only justified when they improve the status of society as a whole, including the poorest members.

We will see soon why inequality matters and what its consequences are. The core focus of this paper, however, is on its determinants, more specifically those that can be attributed to the worker's personal characteristics such as gender, education level, experience and so on. The economic inequality analysed is wage inequality, as we will explain hereinafter.

This paper addresses two questions:

1. *Did the wage inequality increase or decrease for Italian employees from 2000 to 2010?* We use data from the Survey on Household Income and Wealth (SHIW) by Bank of Italy for these two years and check for changes in inequality.

2. *How can we decompose such change, if any?* We use the technique proposed by Chernozhukov, Fernandez-Val and Melly (2009) to perform a counterfactual

participant can put a veto: if he does, both individuals receive nothing. Where traditional economic approach suggests that the second individual should accept any distribution (with the limit case of receiving 0), offers of less than 20% of the reward are often rejected (Henrich et Al, 2004).

[4] Such apparently self-destructive behavior facilitates in the end stable cooperation and trade.

[5] P. Diamond, A. Giddens, The New Egalitarianism, 2005

[6] Rawls, John. A theory of justice. Belknap Press, 1999.

analysis and to see how the eventual variation can be decomposed into the effect of a change in the covariates distribution and the effect of a change of the reward of the same covariates between the two years.

First, we are going to look at reasons to care about inequality in “Relevance”. In “Measures” we are casting a quick glance on inequality metrics, in “Definition” we are clarifying the issue at stake. The “Literature” section contains a brief review concerning how the decomposition challenge has been tackled by different scholars. In “Analysis” we will review the data and check for the fit of the model, in “Decomposition” we show and comment the results of the counterfactual exercise. “Expansion” broadens the descriptive analysis and consider self-employed too. Main findings are summarized in the “Conclusions”.

2 Relevance

We deem not useless to provide preliminarily the reader with some good reasons to care about economic inequality and go through the rest of this paper.

Plenty of studies have been undertaken with the aim of identifying and quantifying the social consequences of economic inequality. The task is not simple. Take the controversial study conducted by Wilkinson and Pickett in 2009[7]. They found a strong correlation between measures such as crime, obesity, mental illness, drug use rates and economic inequality. Two critics were moved to such results and notably to their interpretation:

a. The effect found may be due to poverty and have little to do with social status differences.

b. A third factor may cause both inequality and social problems (cultural reasons, for example), so that no causality link truly exists.

Common sense and empirical observation suggest that people enjoy their possessions not according to an absolute measure, but relatively to what people surrounding them possess. They struggle to keep up with their friends', neighborhoods' and acquaintances' lifestyle: this is known as the "peer effect".

Inequality thus makes people potentially unhappier, but there is something good in it: it provides incentives. The heavier the economic stratification, the heavier the social one, the greater the competition for status, for that "regard" that A. Smith conceived as one of the major driving forces behind economic activity:

> To be observed, to be attended to, to be taken notice of with sympathy, complacency, and approbation, are all the advantages which we can propose to derive from it. It is the vanity, not the ease, or the pleasure, which interests us[8].

Therefore, in its "optimal", meritocratic form, inequality rewards those who work harder, who innovate, who take risks, contributing to development. On the other, dark side of the moon, this rat race to "keep up with the Joneses" causes stress and dissatisfaction, reduces social cohesion and increases social discontent. Society gets weaker.

Moreover, if inequality gets too wide, the gap between the poor and the rich may be or may appear too big to close, resulting in the same lack of incentives as in a "complete

[7] Wilkinson, Richard G., and Kate Pickett. The spirit level. Bloomsbury Press, 2011.

[8] A. Smith, A Theory of Moral Sentiment, 1759

equity" scenario. As evidence shows, high inequality is usually accompanied by low intergenerational social mobility. The 'Great Gatsby Curve' presented in 2012 by A. Krueger plots this relationship. On the x-axis there is the Gini Coefficient, a measure of inequality that will be explained in next section, while on the y-axis we find the "generational income elasticity", i.e. the percentage increase in one's income when the parents' income increase of 1%. This measures the persistence of the advantages across generations: the higher it is, the lower social mobility is.

What is the link between social immobility and inequality? One example is the lack of access for poor children to better (more expensive) schools, which are crucial to find high-paying jobs, and the lack of health care - more common among the "have-nots" - which may lead to sickness, limiting education and employment.

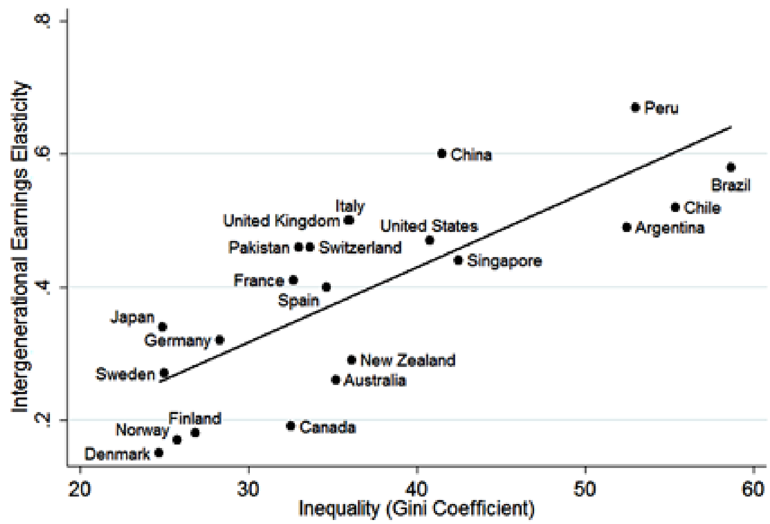

Figure 2.1:
The Great Gatsby Curve
Source: Corak and Granghro (2011) [4]
(Italy and UK share the same point.)

An additional point that deserves our attention is that economic inequality translates into political inequality and vice versa, in a vicious spiral. Political power generated by wealth can indeed shape government policies to be financially beneficial to the rich, in a process called "rent-seeking". Income is not attributed to the wealthy because of their ability to create value, but because of their "weight". This raises inequality even more, and common people are increasingly unable to participate to the democratic processes in the proper way.

These are some of the many so-called "Matthew effects" that can be met in social sciences: initial advantages tend to carry further advantages, and disadvantages further disadvantages, both among individuals and groups through time, creating wider and wider gaps between the "haves" and the "have nots". The name is due to the line that opens this paper.

Another big debate concerns the link between economic growth and inequality. The main idea is that the way the pie is divided affects the size of the pie itself.

In 1975 A. Okun wrote a book, "Equality and Efficiency: the Big Tradeoff"[9]. The title says it all. Okun sustained that pursuing equality reduces incentive to work, save and invest, and thus decreases efficiency. He condemned redistributive measures because some output simply disappears in the transit due to transaction costs and the "have-nots" do not receive all that is taken from the "haves". Moreover, if rich people are rich, it is because resources in their hands were transformed into something more valuable than in the hands of the poor. Taking resources from the "able", the rich, and giving them to "less able" results in an even lower aggregate level of outcome.

J. Stiglitz, Nobel laureate, challenged this view and listed the channels through which inequality has an adverse effect on economic growth. His analysis[10] concerns USA, but we see no reason why the same logic should not apply to Europe and Italy too.

The starting consideration is that more inequality usually implies the thinning of the middle class, which produces a number of consequences. First, the middle class generally sustains job creation and economic growth through its consumption spending. The upper class instead tends to save most of its income.

Second, if the middle class is unable to invest in its future through investments in education and in business, it further decreases its own potential in the medium and long run. O. Galor and J. Zeira[11] demonstrated how inequality in the presence of credit market imperfections has a long-lasting effect on human capital formation and consequently on economic development.

Third, a poorer middle class provides less tax revenues to the State. Besides, Stiglitz argues, people at the top of the distribution are possibly experts in avoiding taxes, in gaining tax-breaks and other favorable treatments by their governments (again, the "rent-seeking" process). Lower tax receipts translate into less fundamental investments in education, research, infrastructure and health; all interventions that would foster long-term economic growth.

[9] Okun, Arthur M. "Equality and efficiency: The big tradeoff" Brookings Institution Press, 1975.

[10] [?]

[11] Galor, Oded, and Joseph Zeira. "Income distribution and macroeconomics." The review of economic studies 60.1 (1993): 35-52.

Therefore, Stiglitz and Okun agree on the fact that the division of the pie affects its size, but with the crucial difference that, according to Stiglitz, the more equal the shares, the bigger the pie will be. There's no trade-off, rather a direct proportionality; no need to ask "how much growth are we willing to sacrifice for a little more equality?". Societies are actually sacrificing growth by allowing inequality.

Was then Okun, back in 1975, so wrong? Probably not. B. Milanovic[12] clearly summarizes this point: both views were correct, times have changed. Before, physical capital had the central role in sustaining growth, so savings and investments were the key. Having more rich people who could save a larger proportion of their income was crucial. Nowadays, value is in people, in human capital, so widespread education is the secret to growth, and widespread education is achieved with more equality.

But how much equality do we want? As in any economic issue, what we look for is not an extreme, rather an optimal point: in Horace's words "est modus in rebus" – "there is measure in all things".

[12] Milanovic, Branko. "Global inequality recalculated and updated: the effect of new PPP estimates on global inequality and 2005 estimates." The Journal of Economic Inequality 10.1 (2012): 1-18.

3 Measures

In order to measure the distribution of income and determine its dispersion within a given population, social scientists use income inequality metrics. Such measures are of course "pocket-tools" that try to condensate as much information as possible in a synthetic form: a scalar number, a single summary statistic.

Fields (1987) identifies four characteristics in the form of axioms that inequality metrics should possess:

- Scale Irrelevance: if each income is multiplied by the same constant, the inequality measure shall not change. It shall be independent of the aggregate level of income.

- Independence from the Population Size: the inequality measure shall remain constant when the size of the population changes, if this change does not affect the income shares of the corresponding percentile groups. This implies that a world composed by two persons, in which one has all the income, shall have the same inequality level of a world with four persons in which two persons have nothing and two share income equally - a characteristic that may not be always desirable.

- Pigou-Dalton Condition, or Transfer Principle: if some income is transferred from a rich person to a poor one, while still preserving their respective income ranks, the inequality metric shall not increase (weak form of the principle) or shall decrease (strong form).

- Anonymity, or Symmetry: if two individuals swap their incomes, the inequality measure shall remain the same. In other words, it is not relevant who possesses what; the metric does not take into account merit considerations.

Now, what have researchers in their toolbox?

A very basic, yet effective metric used is the Decile Dispersion Ratio, or 90/10 ratio. It is obtained dividing the average income of the richest 10% of the population by the average income of the bottom 10%. This metric thus expresses the income of the rich as a multiple of the income of the poor, neglecting the rest of the distribution. Of course the percentages may change (80/20, 70/30, 50/10), allowing a sensitivity analysis or a focus on the section of the income distribution which is more relevant for the researcher's scopes.

Other simple metrics include the coefficient of variation and the proportion of total income earned. The coefficient of variation is measured as the standard deviation of income over its mean and it gets smaller as inequality decreases. Since it relies on two easy statistical concepts, which even people with a limited statistical background are familiar with, it is easy to be explained and understood by a non-technical audit. However, the fact that it does not have an upper bound makes interpretation and comparison more difficult.

The proportion of total income earned measures the proportion of income earned by the poorest x% of the population. Such index however is unable to provide information about how this proportion is shared, and similarly it does not say anything about the other part of the distribution.

The Gini Index is definitely the most common index. It is based on the Lorenz Curve, the graphical representation of the cumulative distribution function of income in the population compared to the perfectly equal distribution of income, the 45° line. In a perfectly equal society, in fact, the poorest x% of the population would earn x% of the total income and the Lorenz Curve would correspond to the 45° line. As inequality increases, the Lorenz Curve deviates from it (the poorest 25% of the population may earn 10% of the total income, for example).

The Gini Index is the ratio of the area between the 45° line and the Lorenz Curve (the striped area) over the total area under the 45-degree line. Consequently, it spans from 0, perfect equality, to 1, maximum inequality, when one person corners all the income.

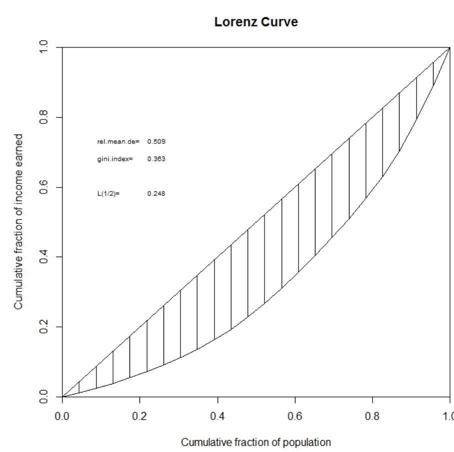

Figure 3.1:
An example of the Lorenz Curve.

Such simplicity comes at a price. The Gini Index is weak in discriminating among different types of inequality. Two Lorenz Curves may cross and different income distributions can result in very similar Gini Indexes, so that comparisons are hard. As a limit example, consider two economies with the same level of aggregate wealth (four in our example, as four are the people in the economy):

- economy *I*, in which half of the individuals or families have no income, and the other half share equally the total income - {0,0,2,2} -; and

- economy *II*, in which the lowest 75% of the population shares equally the 25% of resources, and the top 25% has 75% of resources - $\{\frac{1}{3},\frac{1}{3},\frac{1}{3},3\}$.

These economies have the same Gini Index, 0.5. Figure 3.2 plots the Lorenz Curves for this particular case.

A second weak point is the lack of decomposability, which requires the existence of a coherent relationship between inequality in the society as a whole and inequality in its parts. A decomposable index can be written as a function of inequality within subgroups and inequality between subgroups.

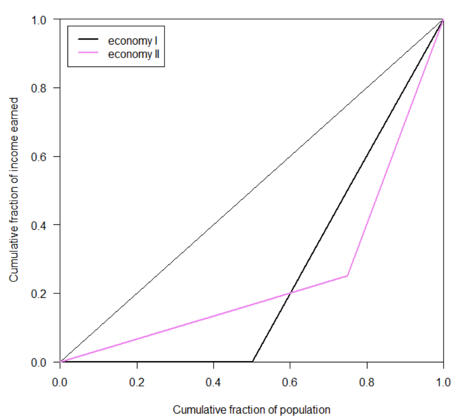

Figure 3.2:
Example of crossing Lorenz Curves

Finally, the Gini Index places an implicit relative value on changes that occur in different parts of the distribution: an equal income transfer from a person to a poorer one reduces the Gini Index more if they are close to the middle of the distribution than if they are both at one end. For example, transferring 1€ from a person with an income of 10.000€ to one with 9000€ lowers the index more than if the transfer occurs between the amounts 1000€ and 900€ or 100.000€ and 90.000€.

The Gini Index is defined as a "positive" measure of inequality, in the sense that it is derived by statistical concepts, it provides a description of the income distribution and it does not use any concept of social welfare, at least explicitly. As Atkinson noted, in fact, the Gini index is not purely statistical, not neutral to social judgment, because it implicitly contains one about the weight to be attached to each point on the income distribution, as previously noted.

The Atkinson family of indexes was conceived with the intention to correct the inability of other measures to attribute different importance to inequalities in the various parts of the income distribution. In this sense, the Atkinson Index is a normative index because it incorporates Rawls' concept of social justice using a sensitivity parameter, ε. This allows placing greater weight on given points of the distribution: the larger ε, the more sensitive the index is to changes in the lower end of the distribution. As the Gini coefficient, the Atkinson Index spans from 0 to 1.

Atkinson Indexes are not additive decomposable, but rather multiplicative, as showed by Lasso de la Vega and Urrutia[13].

Another relevant family of indexes is the General Entropy one. A sensitivity parameter, this time α, is incorporated in the indexes. Usually, the values of α used are -1, 0, 1, 2, with the more positive values making the index more sensitive to inequalities at the high end of the income distribution. The theoretical range of the indexes goes from 0 (perfect equality) to infinity. The most important characteristic of the GE indexes is decomposability: they can be expressed as a weighted average of inequality within groups plus inequality among them. The Theil Index is the most famous in the family of GE, with parameter α=2.

This parade of measures reminds us that a complex issue like inequality cannot be perfectly summarized - reduced - in a single number.

To conclude this section, we present the Italian situation in the European dimension. Figure 3.3 provides an overview of the Gini coefficient for income inequality. Data cover the period 2008-2012. Southwestern countries are generally more unequal than the northeastern ones. Italy is at the same level as Portugal and UK.

Focusing again on Italy, two sources are considered. The first is the OECD, which has data from mid 80s to late 2000s. The second is a paper by L. Pistaferri and T. Jappelli (2009), based on the SHIW, with data from 1975 to 2006, provided discontinuously but periodically (every year until 1987, every two years afterwards). Numbers are not identical, of course, but the shape of the two curves is similar. A decreasing inequality in the 80s diverted his path, went up and stabilized. The problem is that data for more recent years are not available, exactly when the crisis struck and it would have been more relevant to check for inequality changes.

[13] M.A.C. Lasso de la Vega and A. M. Urrutia, A new factorial decomposition for the Atkinson measure, 2003, Economic Bulletin

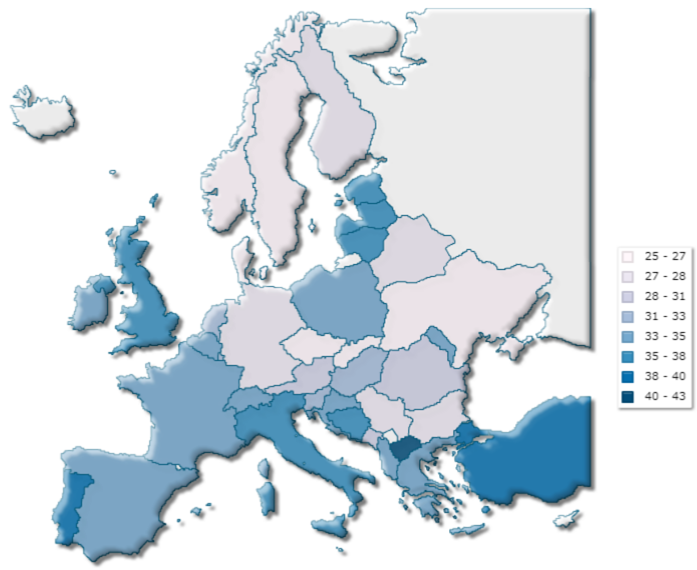

Figure 3.3:
Gini Index for European countries (2008-2012)
Source: World Bank, Development Research Group

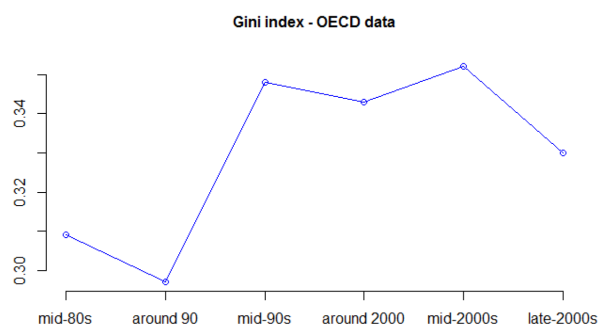

Figure 3.4:
Evolution of the Gini Index in Italy.
Source of data: OECD website

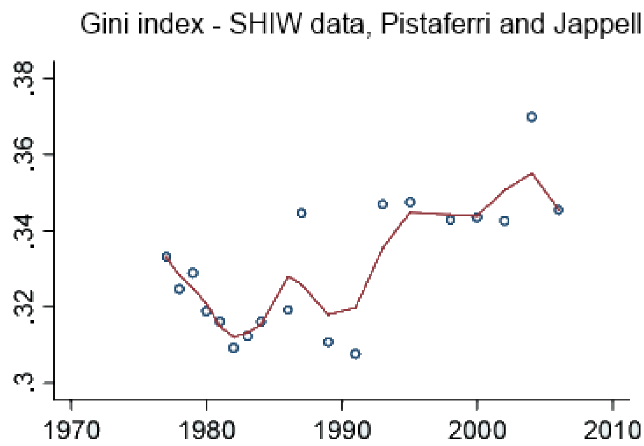

Figure 3.5:
Evolution of the Gini Index in Italy
Source: L. Pistaferri and T. Jappelli (2009)

4 Definition

We have not yet defined for which economic quantity we are measuring dispersion. Candidates are many: wealth, income, wage... What should we use to represent a person's well-being in society?

The most comprehensive measure is wealth, a context-dependent term. One definition is "a person's immediate command over resources": money, assets, body. A number of problems immediately arise. Monetization and aggregation of such disparate possessions are extremely complex because the market price may not reflect value appropriately, provided that a market exists. Moreover, wealth should include also less tangible assets, even harder to price. Education, for example, usually promises a higher future income, or having a job entitles its owner to enjoy pension's rights. We should shift to aggregate earnings over the entire life of individuals, relying either on calculations when he or she is deceased, either on estimates, with all the difficulties linked to such forecasting exercise.

A more limited definition of wealth is the stock of real and financial assets a person has accumulated until a given moment in time. This includes property, savings, ownership of land, rights to private pensions, financial instruments, etc.

Shifting from a "stock" to a "flow" point of view, income is defined as the sum of all the wages, salaries, profits, interests' payments, rents, gifts and other forms of earnings received in a given period of time. It refers to an arbitrary time unit, so it does not take into consideration past accumulation of wealth, neither social wage such as the benefits received from social goods (public services, security), once again hard to quantify. However, information on income is generally more widely available than for wealth, and it is comprehensive enough to give a proper picture.

In this paper the analysis is restricted to wage and wage inequality. Wage is the remuneration that a person receives for her work, the part of income strictly related to labor supply. The other components of income, instead, may be earned whether the person worked or not.

Wage inequality does not concern what people have, but what they get out of their job. We can consider wage inequality as a "conditional inequality": conditional on the fact that the person has found a job. Therefore, we do not account for unemployment or non-employment. On the other hand, these factors should be carefully considered in an analysis that spans labor market and society inequality more widely. Among the main causes of income inequality in Italy, for example, Brandolini (2008)[14] indicates a low participation to the job market, as measured by the number of labor income earners in a household. Another inequality determinant he identifies is the inefficient resource redistribution policy that, for the same nature of our analysis, focused on the labor market only, we do not take into consideration.

Wages are measured in both monthly and hourly time spans for individual employee. The choice between households and individuals is not neutral, as the first has a varying number of members and income pooling and intra-family transfers are not taken into account.

After having dropped observations for students, unemployed and non-employed, the sample is further restricted, in line with the international literature: self-employed are excluded, leaving only employees.

Wage determinants are possibly different for the two groups. The way individual characteristics are linked to wages is different and probably not comparable. Self-employed are more directly affected by economic cycles, they can more readily adjust to shocks by changing their labor supply and so on. This may produce poor quality results in the decomposition analysis. Finally, self-employed tend to report incomes lower than the actual values[15], and this would affects results in an unpredictable way.

Nonetheless, a brief descriptive analysis for wage inequality including also self-employed is provided afterwards.

[14] Brandolini, Andrea. "Income Inequality in Italy: Facts and Measurement.", 2008, Bank of Italy

[15] Cannari, Luigi, and Giovanni D'Alessio. "The opinion of Italians on tax evasion."Bank of Italy Economic Research Paper 618 (2007).

5 Literature

Decomposition methods break down the difference in a given distributional statistic between two groups or years into explanatory factors. They allow answering different questions of different nature: explaining to which factors the gender gap or the race gap can be attributed; quantifying contributions of labor, capital and productivity in the growth of an economy and so on. However, the will of understanding the factors behind inequality growth is the force that mainly stimulated research, especially after the increase of observed wage disparities in the USA and other developed countries since the late 70s.

The first attempts to decompose inequality in labor economics date back to Oaxaca and Blinder, who wrote their fundamental papers in 1973. The distributional statistic of interest was in both cases the mean of the outcome variable, but soon enough the analysis spread to other parameters such as the variance (that I will not cover in this paper) and the quantiles.

5.1 Mean

The Oaxaca-Blinder decomposition (OB) is widely used to study the difference in mean wages between two groups or, as in our case, two periods. The wage equation is assumed to be linear and separable in observable and unobservable terms. On this basis we can write two equations in matrix notation:

$$\begin{aligned} &Y^{00} = X^{00}\beta^{00} + \varepsilon^{00} && \text{for year 2000} \\ &Y^{10} = X^{10}\beta^{10} + \varepsilon^{10} && \text{for year 2010} \end{aligned}$$

where X is a matrix of covariates in a given year, both vectors β contain also the intercept and $\mathbb{E}[\varepsilon^{year}|X^{year}] = 0$. The estimated gap between the average incomes in the two years is:

$$\bar{Y}^{10} - \bar{Y}^{00} = \bar{X}^{10}\beta^{10} - \bar{X}^{00}\beta^{00}$$

where $\bar{X}^{10}$ and $\bar{X}^{00}$ are now vectors containing the average value for each variable. Conditional expectations of the error terms for both years are zero by the zero conditional mean assumption. Adding and subtracting the same term[16]:

[16] Exchanging the reference group does not involve any specific estimation issue, just a different interpretation.

$$\bar{Y}^{10} - \bar{Y}^{00} = \bar{X}^{10}\hat{\beta}^{10} - \bar{X}^{10}\hat{\beta}^{00} + \bar{X}^{10}\hat{\beta}^{00} - \bar{X}^{00}\hat{\beta}^{00}$$
$$\bar{Y}^{10} - \bar{Y}^{00} = \bar{X}^{10}(\hat{\beta}^{10} - \hat{\beta}^{00}) + (\bar{X}^{10} - \bar{X}^{00})\hat{\beta}^{00}$$

where $\hat{\beta}^{10}$ and $\hat{\beta}^{00}$ are estimated intercepts and slope coefficients for the two years. The term added and subtracted represents the counterfactual wage that would have been paid in 2000 to a representative 2010 worker (with 2010 average characteristics). The overall gap can be rewritten as:

$$\hat{\Delta}_o^{\mu} = \bar{Y}^{10} - \bar{Y}^{00} = \bar{X}^{10}\hat{\Delta\beta} + \Delta\bar{X}\hat{\beta}^{00} = \hat{\Delta}_s^{\mu} + \hat{\Delta}_x^{\mu}$$

where $\Delta\bar{X} = \bar{X}^{10} - \bar{X}^{00}$, $\hat{\Delta\beta} = \hat{\beta}^{10} - \hat{\beta}^{00}$.

The gap $\hat{\Delta}_o^{\mu}$ is decomposed into:

- $\hat{\Delta}_s^{\mu}$, the wage structure effect, the effect of a change in the relationship linking the covariates X to the Y ($\bar{X}^{10}\hat{\Delta\beta}$).
- $\hat{\Delta}_x^{\mu}$, the composition effect, the effect of the change in the distribution of the set of covariates X ($\Delta\bar{X}\hat{\beta}^{00}$).

In cases where the group membership is linked to the immutable characteristics of the worker (gender, race), the expressions "discrimination part" and "unexplained part" are used respectively.

The decomposition between wage structure and composition effect (called the "aggregate decomposition") can be pushed even further to obtain a detailed decomposition, that is, to subdivide both $\hat{\Delta}_s^{\mu}$ and $\hat{\Delta}_x^{\mu}$ into the respective contributions of the covariates: $\Delta_{s,k}^{\hat{\mu}}$ and $\Delta_{x,k}^{\hat{\mu}}$, for k=1, 2, ..., K where K is the total number of covariates considered. While the detailed decomposition for the composition effect is always clearly interpretable, the wage structure effect detailed decomposition is not, since it arbitrarily depends on the choice of the omitted group (more on this under "A word of caution").

It is interesting to notice that the wage structure effect can be seen as a treatment effect, the causal effect of a binary variable on the wages. This is not so apparent in our case with time differences (unless there have been some changes in policies in the meantime), but in an unionization framework, for example, the unionized workers can be considered as the treated group and the wage structure effect as an Average Treatment Effect on the Treated (ATT).

Differently from the program evaluation framework, where the composition effect is a confounding factor to be controlled for, inequality studies are usually interested in it. However, it may be convenient to represent the wage structure effect as a treatment effect because the zero conditional mean assumption can be replaced by a weaker assumption: ignorability. Ignorability does not require errors (or unobservables, ε) to be mean independent of X as long as their conditional distribution given X is the same in both groups or years.

In other words, while a linear regression is affected by the ability bias when there are no variables accounting for it, resulting in inconsistent estimates, the aggregate decomposition remains valid if the structure of dependence between ability and education and/or other variables is the same in both groups. This is a precious result given that our regression could potentially feel such bias, and there are no reasons to believe that the relationship between ability and education has changed significantly during the years.

Unfortunately, the same result does not apply for the detailed decomposition.

A word of caution

The OB decomposition and the other methods described below are subject to some limitations. While they are useful to quantify the contribution of different factors on the difference in the outcome in an "accounting" sense, to provide an explanation in the statistical sense, they do not shed a direct light on the mechanisms underlying such relationships. They may or may not provide causality evidence unless some stringent assumptions are met (as the OLS regression, for example). However, decomposition methods point out which factors are quantitatively relevant, indicating a direction for further analysis.

Furthermore, it is important to underline that all decomposition methods implicitly follow a partial equilibrium approach. When the counterfactual treatment "$\bar{X}^{10}\hat{\beta}^{00}$" is written, we posit that workers in 2010 are paid according to the wage structure of 2000 (and the same assumption applies to the counterfactual distributions hereinafter). This is of course a fake scenario, a "what if", and there is no guarantee that people would not have responded to the new wage structure changing their labor supply or in some other way. The economy could have reached a new equilibrium.

An effective example provided by Lewis (1963) concerns unionization analysis. Here, the counterfactual distribution is the wage distribution that would arise if union workers were paid accordingly to a non-union schedule, like if unions did not exist. Yet, there are strong reasons to believe that eliminating unions would have an effect not only on the union workers's wages, but also on those of non-unions workers.

In cases in which the simple counterfactual treatment or distribution is suspected not to be the appropriate one, it may be better to consider a new wage structure. For the OB decomposition, some scholars[17] proposed alternative counterfactuals. Mainly, their idea is to create a new wage structure by weighting the existent ones. The methods differ in the choice of the weighting matrix Ω in the expression $\beta* = \Omega\beta_{00} + (I - \Omega)\beta_{10}$, where in the normal OB decomposition Ω=I.

The reference group (or omitted group) problem is another shortcoming, as showed by Oaxaca and Ransom (1999). In case of categorical covariates, results in the detailed decomposition for the wage structure effect are affected by the choice of the omitted group.

5.2 Distribution

In case one is interested to wage inequality as we are, a decomposition based on the average wage or variance is not enough. We want to evaluate the contribution of the covariates and of the wage structure at different points of the distribution, since the relevance of different factors may vary widely. Take for example the effect of minimum wage, that clearly affects the bottom end of the wage distributions, or unionization, that tends to affect its middle part. Thus it is important to go beyond too simplistic summary measures.

Unfortunately, the OB decomposition cannot be used for distributional statistics other than the mean. Let's consider the quantile decomposition case.

A value y is said to be the q-th quantile of the probability distribution of the statistic Y if $P(Y \leq y) = q$, while the percentage q is called the "quantile rank" of y. In the quantile regression, the conditional interpretation for β is similar to the one of the OLS regression: $\mathbb{E}(Y|X) = X\beta$ and analogously $Q_q(Y|X) = X\beta_q$.

However, the Law of Iterated Expectations which is used in the OB decomposition is not respected. In fact, for the mean, $\mathbb{E}(Y) = \mathbb{E}[(Y|X)|X] = \mathbb{E}(X)\beta$ while, for the quantile regression $Q_q(Y) \neq \mathbb{E}(Q_q(Y|X)|X) = \mathbb{E}(X)\beta_q$. Because of this, one has to look somewhere else.

Retrieving the cumulative distribution is the main pillar of all solutions proposed, given that any standard distributional statistic can be derived from that. In the following part

[17] Reimers (1983), Cotton (1988), Neumark (1988).

of the paper, $F_{Y_{00}|X_{00}}(y|X)$ and $F_{Y_{10}|X_{10}}(y|X)$ represent the conditional distributions describing the stochastic assignment of wages to workers with characteristics X for 2000 and 2010 respectively; $F_{Y(00|00)}$ and $F_{Y(10|10)}$ represent the observed wage distribution functions; finally $F_{Y_{00}|X_{10}}(y|X)$ or $F_{Y(00|10)}$ is the counterfactual wage distribution that would have prevailed if people of 2010 were paid according to the 2000 wage schedule. This latter distribution is not observable.

Similarly to the OB decomposition for the average wage, the difference in the observed wage distributions can be decomposed as:

$$F_{Y(10|10)} - F_{Y(00|00)} = [F_{Y(10|10)} - F_{Y(00|10)}] + [F_{Y(00|10)} - F_{Y(00|00)}]$$

The first term on the right-hand side is the wage structure effect $\hat{\Delta}_s^d$, the second the composition effect $\hat{\Delta}_x^d$.

The challenge is exactly to build the counterfactual distribution:

$$F^c = F_{Y(00|10)}(y) = \int F_{Y_{00}|X_{00}}(y|X)dF_{X_{10}}(X) \tag{5.1}$$

We can divide the approaches that have been proposed over time in three groups. The first set of methods suggests replacing each value of Y_{10} with a counterfactual value $Y_{00}^C = g(Y_{10}, X)$.

Juhn, Murphy and Pierce (1993, JMP) propose to obtain the counterfactual wage for 2010 replacing both the return to observables and unobservables with those of 2000. One of the main drawbacks of this method (as well as of DFL later on) is that results are not immune to the choice of the reference group whose wage structure is kept fixed (2000 in our case).

We start by expressing differently the unobserved term in the wage equation for each year:

$$Y_{year} = X\beta_{year} + \varepsilon_{year} = X_{year}\beta_{year} + F_{year}^{-1}(Q(X_{year})|X_{year})$$

where $F_{year}^{-1}(Q(X_{year})|X_{year})$ is the inverse cumulative distribution of residuals conditional on X and $Q(q_{year}|X)$ the conditional rank (percentage of observations that are the same or lower than it) in the residual distribution.

First, the counterfactual unobservables distribution is considered:

$\varepsilon_{00}^{C} = F_{00}^{-1}(Q(X_{10})|X_{10})$

and it is used to obtain two counterfactual distributions:

$Y_{00}^{C,I} = X_{10}\beta_{10} + \varepsilon_{00}^{C}$

$Y_{00}^{C,II} = X_{10}\beta_{00} + \varepsilon_{00}^{C}$

Changes in inequality over time may be attributed to:

- changes in the distribution of the observed characteristics, X;
- to changes in return of those observable characteristics, β;
- to changes in the distribution of residuals, $F(\tau|X)$.

Depending on the effect we want to estimate, we use the first or the second counterfactual.

Notice that this method is based on the assumption that the rank of a given worker in the distribution of ε_{10} is the same as in the distribution of ε_{00}. What does it mean? A helpful example is provided by Fortin, Lemieux and Firpo. Consider the scenario in which the residual vector ε contains two (unobserved) ability measures, cognitive and manual ability. If in year 2000 cognitive ability was more valued that in 2010, it is likely that the ranking of a worker would change. This assumption would thus be violated.

Such assumption is much stronger that the ignorability one described above, for which it is sufficient that the conditional distribution of cognitive and manual ability given X is the same in 2000 and 2010.

The JMP method cannot be extended straightly to the detailed decomposition. Moreover, in practice it is not clear how to input residuals depending on X in the counterfactual residual distribution. Some assume that residuals are independent of X, so that we are left with $F_{00}^{-1}(\tau_{10})$. It is then sufficient to compute the rank of the residual ε_{10} for worker i over the whole 2010 sample, then pick the corresponding residual in 2000. If the rank for a given worker in 2010 is 30% ($\tau_{10} = 0.3$), in 2000 we take the 30% percentile of the 2000 residual distribution. This simplification is however unrealistic.

Machado and Mata (2005, MM) use the conditional quantile regression to overcome such obstacle.

The idea is again to transform each observation of Y_{10} into a counterfactual value Y_{00}^C by means of the quantile regression. There are two important differences with respect to the JMP method. First, MM provide a way to estimate in an explicit way the (inverse) conditional distribution. Second, instead of transforming actually each observation into a new counterfactual one, a simulation approach is used.

The suggestion by MM is to estimate the quantile regression for all q in [0,1] to characterize a full conditional distribution of Y|X, using the relationship $Q_q = F^{-1}(q)$.

As we said, $Y_{00}^C = g(Y_{10}, X)$. In MM we have

$$g(Y, X) = F^{-1}_{Y_{00}|X_{00}}(F_{Y_{10}|X_{10}}(Y|X)|X)$$

A simulated value q is drawn from a uniform distribution. The linear quantile regression is estimated for each q and used to obtain simulated values for both Y_{10} and Y_{00}^C. Finally, these values are compared to obtain the wage structure effect. The composition effect is obtained as the difference of the overall and the wage structure effect.

MM suggest a linear conditional quantile regression model for $F^{-1}_{Y|X}(q|X)$;

$$F^{-1}_{Y_{year}|X_{year}}(q, X) = Q_{year,q}(Y|X) = X\beta_{year}(q)$$

The wage structure effect can be easily decomposed in detail, replacing one after the other the β_qs of 2010 by those of 2000. The problem is that this method is path dependent, meaning that the decomposition results depend on the order in which the decomposition is performed. Unluckily, there is no way to obtain the detailed decomposition for the composition effect, the one that, as we already pointed out, is always clearly interpretable.

Another limitation of the MM approach lies in the reliance on simulation methods which are computational intensive and may become problematic, especially for big datasets.

The second approach is to use a reweighting function to estimate the counterfactual distribution of interest.

Di Nardo, Fortin and Lemieux (1996, DFL) propose a generalization of the Oaxaca-Blinder decomposition.

The idea is that the counterfactual distribution can be expressed as:

$$F_{Y(00|10)}(y) = \int F_{Y_{00}|X_{00}}(y|X)dF_{X_{10}}(X) = \int F_{Y_{00}|X_{00}}(y|X)\Psi(X)dF_{X_{00}}(X)$$

where $\Psi(X) = \frac{dF_{X_{10}}(X)}{dF_{X_{00}}(X)}$.

$\Psi(X)$is a reweighting function, the ratio of probability mass at each point of the set of covariates in 2010 relative to 2000. The $\Psi(X)$ thus reweights the 2000 density so that observations with a higher frequency in 2010 than in 2000 are made "heavier" and vice versa.

While in theory this reweighting function is a simple expression, problems arise in its empirical estimation. It is likely that, for some values of X, the probability mass at the numerator or at the denominator is 0. The DFL method overcomes this obstacle by cleverly applying the Bayes' rule:

$$P(A|B) = \frac{P(B|A)P(A)}{\sum P(B|X) \cdot P(X)}$$

Remember that $F_{X_{10}} = F(X|t = 2010)$, so

$$P(X|t = 2010) = \frac{P(t=2010|X) \cdot dF(X)}{\int P(t=2010|X) \cdot dF(X)}$$

and similarly for $F_{X_{00}} = F(X|t = 2000)$. Therefore, we can rewrite $\Psi(X)$ as

$$\Psi(X) = \frac{P(t=2010|X) \cdot \int P(t=2010|X) \cdot dF(X)}{P(t=2000|X) \cdot \int P(t=2000|X) \cdot dF(X)} = \frac{P(t=2010|X) \cdot P(t=2000)}{P(t=2000|X) \cdot P(t=2010)}.$$

$P(t = 2010|X)$ and $P(t = 2000|X)$ can be evaluated by a probit, logit or similar models in a parametric fashion[18], while cumulative distributions are estimated in a non-parametric way. The entire DFL method is thus defined as "semi-parametric".

The DFL model is recommended for the aggregate decomposition given its simplicity, and also in case a simple mean decomposition has to be performed, since the OB method results in biased estimates when the conditional expectation of Y given X is non-linear. However, reweighting has some undesirable properties in small samples, and the DFL method cannot be extended easily to the detailed decomposition case.

A third way relies on the estimation of the conditional distribution of the wage outcome in year 2000 ($F_{Y|X}(Y|X)$) and its subsequent manipulation.

[18]Also non-parametric logit models have been proposed (Hirano, Imbens and Ridder, 2003).

Chernozhukov, Fernandez-Val and Melly (2009, CVM) start from the direct estimation of $F_{Y_{00}|X_{00}}(Y|X)$. This is the method we are going to apply to the Italian dataset, as it is very flexible and applicable to the detailed decomposition, even if results are path dependent.

A separate "distribution regression model" is estimated for each value of y using $F(y|X) = \Lambda(P(X) \cdot \beta(y))$, where P(X) is a vector of transformations of X and Λ is a link function. CVM propose the complementary log-log link function, $\Lambda(z) = 1 - e^{-e^{z}}$, and we follow this suggestion. First the conditional distribution function is estimated:

$$\hat{F}_{Y_{00}|X_{00}}(y|x) = \Lambda(P(x)\hat{\beta}_{00}(y)) \qquad (y, x) \in \chi_{00}, \gamma_{00}$$

where

$$\hat{\beta_{00}(y)} = \underset{b \in R^p}{arg\hat{m}ax} \sum_{i=1}^{n00} [1\{Y_{00,i} \leq y\} \, ln[\Lambda(P(X_{00,i})b)] + 1\{Y_{00,i} \geq y\} \, ln[1 - \Lambda(P(X_{00,i})b)]$$

and p is the dimension of P(X). Link functions other than the complementary log-log can be used as well: the logit, probit, linear, log-log functions. It is worth noting that, if P(X) is rich enough, the choice of the link function is not relevant.

An estimate of the covariates distribution $F_{X_{10}}(x)$ is obtained using the empirical distribution function:

$$\hat{F_{X_{10}}}(y) = \int_{\chi_{10}} \hat{F}_{Y_{10}|X_{10}}(y|x) d\hat{F}_{X_{10}}(x)$$

Once we have $\hat{F}_{Y_{00}|X_{00}}(y|x)$, we integrate over the distribution of X_{10} in order to obtain the counterfactual distribution:

$$\hat{F}_{Y(00|10)}(y) = \int_{\chi_{10}} \hat{F}_{Y_{00}|X_{00}}(y|X) d\hat{F}_{X_{10}}(X)$$

$$\hat{F}_{Y(00|10)}(y) = \frac{1}{n_{10}} \sum \hat{F}_{Y_{00}|X_{00}}(y|X) \qquad \text{for } x \in \chi_{10},\ n_{10} \text{ number of obs. for 2000}$$

The conditional distribution of wages given the set of covariates X is estimated using regression estimators, while the covariate distribution is evaluated in a nonparametric fashion. The estimator for the functionals of the observed and of the counterfactual marginal distributions of wages are proved to be uniformly consistent and asymptotically Gaussian. The contribution of CVM includes the provision of limit distribution theory and inference tools for counterfactual estimators based on the distribution regression.

Firpo, Fortin and Lemieux (2009, FFL) propose a path independent decomposition using the recentered influence function (RIF). The RIF regression is a standard regression except for the dependent variable, which is replaced by the RIF of the statistic of interest. Influence functions are a tool used to assess the effect (or "influence") of removing one observation on the value of a statistic without having to recalculate it. FFL provide many different RIF regressions depending of the statistic of interest. Once the RIF is computed, a simple OLS regression is performed.

While in the CVM method proportions are inverted globally in the space of quantiles, in the FFL method the inversion is locally performed. So, in case of a linear relationship between counterfactual quantiles and proportions they give the same results, but the local approximation of FFL may be poor for extreme quantiles. Still the detailed decomposition obtained is path independent.

The following table summarizes the main assumptions and characteristics of the methods described:

Method	Assumptions	Features
JMP	Cond. rank pres. / Linearity of $E(Y\|X)$	No detailed decomposition
MM	Cond. rank pres. / Linearity of $Q_\tau(Y\|X)$	No detailed decomposition
DFL	Invariance of Conditional Distr.	Path dependent
CVM	Cond. rank pres. / Invariance of Cond. Distr.	Path dependent
FFL(RIF)	Invariance of Conditional Distr.	Path independent

6 Analysis

6.1 Dataset description

The dataset used is the Survey of Household Income and Wealth (*Indagine sui bilanci delle famiglie italiane*, SHIW) by Bank of Italy. The years considered are, as repeatedly said, 2000 and 2010.

The sample has been restricted to employees aging between 18 and 65 years. We measure wages using both the hourly and the monthly log-wage to see if results are affected by the unit of measure. Wages are expressed in nominal terms.

Unfortunately, wages are net rather than gross, and similarly in all Italian datasets. The information on taxes paid is impossible to directly retrieve and very difficult to calculate given the complexity of the Italian tax system, so we cannot take into account the effect of potential tax policy changes in our analysis.

The set of regressors includes a gender variable, years of potential experience, dummies for the highest level of education achieved, for the area of residence (north, center or south) and for the type of contract owned.

Potential experience is defined in the "Mincerian" way, that is, as the number of years E a worker of age A could have worked assuming she started school at 6 and studied for S years:

$$E = A - S - 6$$

A quadratic term is also included to capture nonlinear effects.

Education is measured using dummies for each level achieved[19]. The use of dummies instead of years of education[20], as in the USA literature, is motivated by social and cultural differences. In the Italian system, there is a premium for finishing the final years of a school level and getting a degree or a diploma. Years not capped by a certification are usually disregarded by potential employers. This effect is called "sheepskin effect".

There exist three possible types of contract: open-ended contract (*contratto a tempo indeterminato*), fixed term contract (*contratto a tempo determinato*) and temporary contract (*contratto interinale/somministrazione di lavoro* after 2007). This latter type of

[19] Elementary school, middle school, vocational school (*istituto professionale*), high school diploma, bachelor degree, master degree and PhD

[20] In the SHIW years of education achieved are not asked; the highest diploma/degree achieved is. From that it is possible to calculate "potential years of education" but, for lack of additional data, this estimate cannot take into account drop-outs or repeated school years.

employment consists of a relationship among three entities: the employee, an agency and an employer, a public or private firm that needs the worker. Two contracts are stipulated: one "consultancy contract" (*contratto di somministrazione*) between the agency and the employer, one "concluded work contract" (*contratto di lavoro concluso*) between the agency and the worker. The main job contract is between the worker and the agency, which has to pay him or her in an adequate manner with respect to the kind of performed job.

The following table provides the summary statistics of the covariates. Standard errors are in parentheses.

Variable	2000	2010
N obs	6278	5538
Hourly log wage	1.895 (0.46)	2.152 (0.44)
Montlhy log wage	6.952 (0.48)	7.179 (0.46)
Primary S.	0.09	0.04
Middle S.	0.31	0.29
Vocational S.	0.09	0.10
High S.	0.37	0.38
Bachelor C.	0.01	0.02
Master C.	0.12	0.15
PhD	0.00	0.01
Female (%)	0.41	0.45
Experience	22.30 (11.7)	25.07 (11.7)
Experience^2	634.31 (565.7)	776.06 (584.5)
Open-ended contr.	0.89	0.85
Fixed contr.	0.10	0.14
Temporary contr.	0.01	0.01
North	0.48	0.47
Center	0.22	0.22
South	0.30	0.31
Age group		
18-25	0.096	0.057
25-40	0.397	0.298
40-65	0.506	0.644

We start from a preliminary analysis concerning the change in covariates over time, i.e., the composition effect.

The percentage of female workers slightly rose between 2000 and 2010. Many different factors could have caused this phenomenon. Women may have decided to enter the workforce because of a social and cultural change. They may also have started to accept lower wages for leaving their "housewives status" (minimum monthly and hourly wages actually decreased). However, we do not have the tools to say more, or to prove any of these guesses.

The percentages describing the workers' area of residence remained constant, and similarly those for types of contracts, where the open-ended contracts slightly decreased and the fixed contracts increased by the same amount.

Average education rose, but again just a little: in the dataset there are more workers who achieved at least a vocational school diploma. High school and Master Degree are the levels that grew the most.

Years of experience increased in the whole distribution, but especially in the central part, with the median changing of 4 years while the10th and 90th quantile of 2. The 2000 distribution is skewed towards the left, while the 2010 one is skewed towards the right. Related to this, it is worth noting that the workers' population grew older. The percentage of young workers diminished: this may be due to a higher fraction of young people that choose to continue their studies, to an increase in the NEET category (Not in Education, Employment or Training), but also to the progressive ageing of the active population.

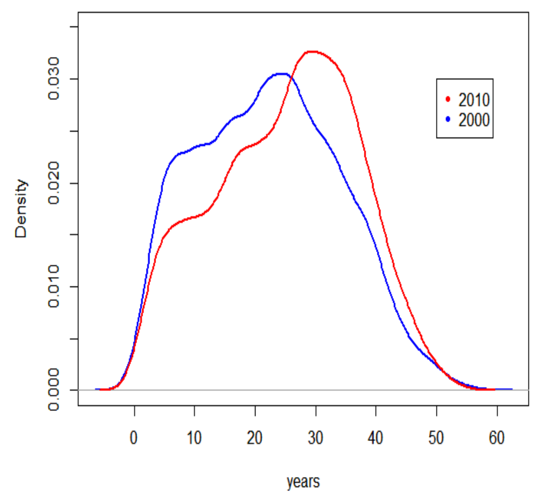

Figure 6.1:
Experience density functions

6.2 Residual analysis

It is interesting to study whether the ability of the covariates we include in our regression to explain wages changed over time. Put in another way, we are going to check whether the unobserved characteristics grew in relevance in the wage equation. This task is performed using a basic residual analysis.

First, we run a WLS regression in both years on the set of covariates for hourly and monthly wages. The coefficients represent the percentage change in wages with respect to a baseline individual, a man who possesses neither education nor experience, who is from the North of Italy and is endowed with an open-ended contract.

There are few non-significant variables, notably primary school coefficients and contracts dummies in the hourly wage case for 2000.

Hourly Wages	2000			2010		
	coeff.	rob se	t-test	coeff.	rob se	t-test
Primary S.	0.090	0.060	1.490	0.039	0.123	0.320
Middle S.	0.239	0.060	3.960	0.189	0.122	1.550
Vocational S.	0.348	0.062	5.620	0.282	0.123	2.290
High S.	0.519	0.061	8.520	0.434	0.123	3.540
Bachelor C.	0.725	0.074	9.730	0.693	0.128	5.420
Master C.	0.851	0.063	13.500	0.771	0.123	6.250
PhD	0.932	0.155	6.010	0.995	0.131	7.580
Female	-0.093	0.010	-8.980	-0.134	0.010	-12.970
Experience	0.037	0.002	20.550	0.029	0.002	15.790
Experience^2	0.000	0.000	-12.320	0.000	0.000	-8.810
Fixed contr.	-0.015	0.025	-0.610	-0.123	0.018	-7.000
Temporary contr.	-0.066	0.095	-0.700	-0.204	0.059	-3.460
Center	-0.073	0.011	-6.690	-0.029	0.013	-2.200
South	-0.120	0.013	-9.450	-0.060	0.012	-4.910
Intercept	1.040	0.062	16.830	1.395	0.123	11.310

Monthly Wages	2000			2010		
	coeff.	rob se	t-test	coeff.	rob se	t-test
Primary S.	-0.007	0.058	-0.120	0.072	0.081	0.890
Middle S.	0.137	0.058	2.360	0.193	0.077	2.490
Vocational S.	0.244	0.059	4.110	0.307	0.079	3.900
High S.	0.393	0.058	6.750	0.422	0.078	5.410
Bachelor C.	0.498	0.071	6.990	0.560	0.085	6.570
Master C.	0.648	0.061	10.670	0.704	0.079	8.900
PhD	0.865	0.153	5.650	0.951	0.097	9.760
Female	-0.275	0.011	-25.820	-0.293	0.011	-27.180
Experience	0.035	0.002	20.060	0.028	0.002	15.520
Experience^2	0.000	0.000	-12.660	0.000	0.000	-9.570
Fixed contr.	-0.177	0.023	-7.730	-0.231	0.018	-12.970
Temporary contr.	-0.218	0.074	-2.970	-0.407	0.083	-4.930
Center	-0.095	0.012	-8.160	-0.055	0.014	-3.850
South	-0.165	0.013	-12.840	-0.114	0.011	-10.030
Intercept	6.371	0.059	107.840	6.579	0.079	82.940

In order to check for the fit of the model, we are interested in the residual sum of squares (SSR) over the total sum of squares (TSS). This is equivalent to $1 - R^2$. R^2 is, in linear regressions, the square of the correlation between the dependent variable, wages, and its predicted values. For example, a R^2 of 0.3 means a correlation of 0.55. Results are displayed below, in the first table.

The goodness of the model stays nearly constant: unobserved characteristics such as ability seem to be no more important than in the past.

Nevertheless, unobserved characteristics could have different importance across the distribution. For example, if ability is more relevant for highly-paid positions, the fit of the model for these observations should turn out to be worse than for the rest of the workers. To check for this eventuality, we run ten quantile regressions that span the entire distribution and consider the pseudo-R^2.

Changes in these measures are very small. In each year, the hourly pseudo-R^2 increases moving toward the top of the distribution, while the monthly pseudo-R^2 follow an unstable pattern. From 2000 to 2010 there is generally an improvement, but again it is faint and not really relevant. We can conclude that the fit of the model stays constant both across years and across the distribution.

R^2	2000	2010
Hourly reg.	0.30	0.31
Monthly reg.	0.32	0.35

Hourly pseudo-R^2			Monthly pseudo-R^2		
Quantile	2000	2010	Quantile	2000	2010
10th	0.22	0.19	10th	0.24	0.23
25th	0.20	0.19	25th	0.22	0.22
50th	0.20	0.21	50th	0.19	0.21
75th	0.22	0.23	75th	0.19	0.22
90th	0.22	0.23	90th	0.23	0.25

7 Decomposition

In this section we finally analyse the evolution of wages and wage inequality.

Wages experienced an increase, as it is evident from the density functions in Figure 7.1 that shifted to the right under both considered measures. We stress once again that these are nominal wages and we do not account for any inflation effect[21].

Since we are going to deal with cumulative distributions, it is more convenient to refer to Figure 7.2.

The 2010 distribution lines are always below the 2000 ones. Deducing an increase in wages under these terms may seem counterintuitive, but call to mind the meaning of $F(y)$. This is equivalent to $P(Y \leq y)$, the probability that wages are below a given threshold or, differently stated, the percentage of people earning less than y. In order to have more people enjoying higher wages, $F(y)$ has to be the lowest possible, therefore higher wages correspond to the red lines in our graph.

We apply the Chenozhukov-Val-Melly method to estimate the counterfactual distribution, represented by the black lines, and to decompose the change in the wage distribution. The R code can be found in Appendix A.

$$F_{Y(10|10)} - F_{Y(00|00)} = [F_{Y(10|10)} - F_{Y(00|10)}] + [F_{Y(00|10)} - F_{Y(00|00)}] \tag{7.1}$$

For any given wage, the vertical distance between the red and the blue points is the overall difference we would like to explain, the left-hand side of the equation, $\hat{\Delta}_o^d$. On the right-hand side we find the distance between the red and the black points, the wage structure effect $\hat{\Delta}_s^d$, and the distance between the black and the blue points, the composition effect $\hat{\Delta}_x^d$, respectively.

[21]Inflation trended down in a stable fashion over the period, oscillating between 2-3% in 2000 and ending up at 1% in 2010.

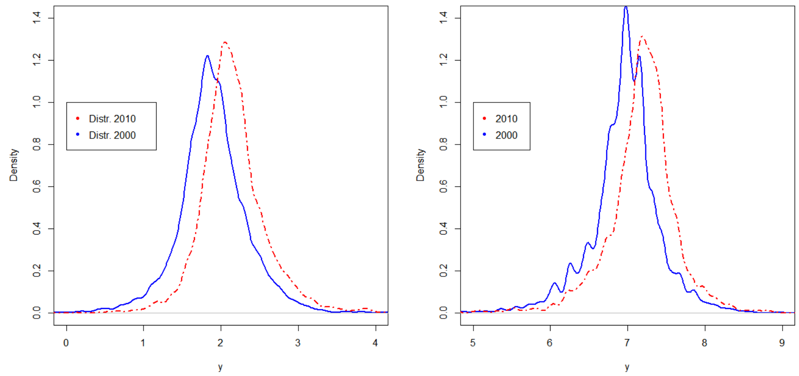

Figure 7.1:
Wage density functions, hourly and monthly

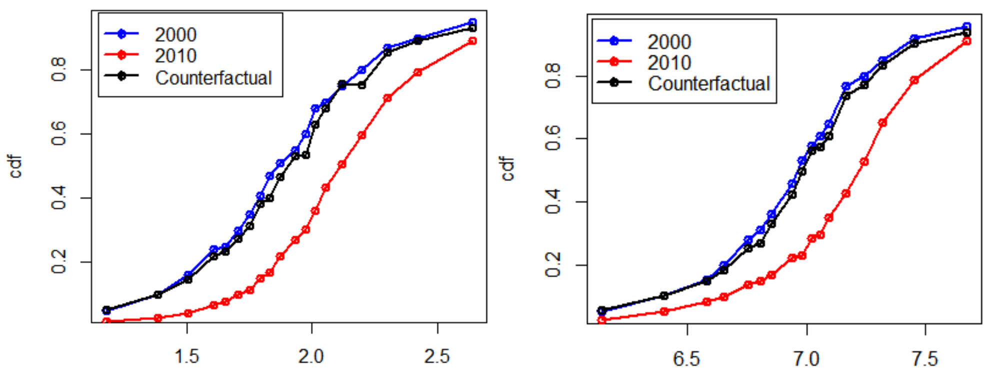

Figure 7.2:
ECDF and counterfactual distribution plotted on log-wages, hourly and monthly.

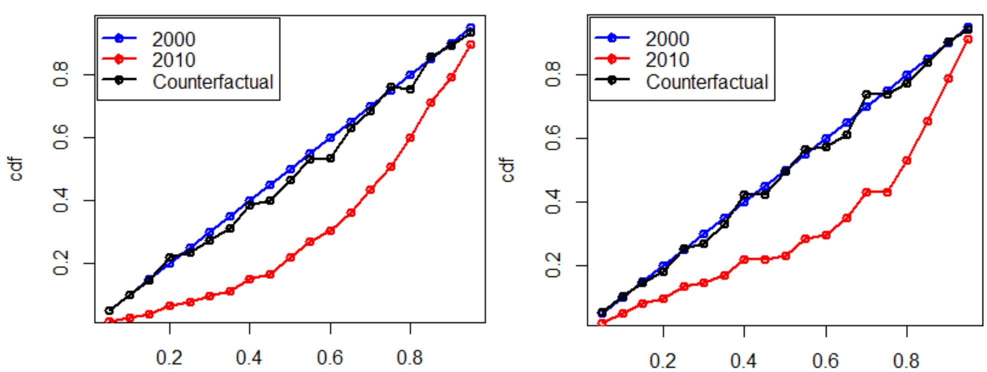

Figure 7.3:
ECDF and counterfactual distribution plotted on 2000 quantiles, hourly and monthly.

Figure 7.3 plots the same results in a slightly different way. On the x-axis there are the 2000 quantile ranks instead of actual log-wages. This is the reason why in Figure 7.2 points are not equally spaced while they are in Figure 7.3, with a regular distance of 0.05.

The 45° blue line is the 2000 reference distribution. The more the distance between the blue and the red points, the more the growth has been significant. To make things clearer, consider Figure 7.3a. The fourth from last blue point represents the 80th 2000 quantile (the associated cumulative probability is 0.8 of course, since 2000 is out reference distribution). The corresponding red point is close to 0.60 instead. The same wage that in 2000 corresponds to the 80th quantile is nearly the 60th in 2010. We have already explained why this is a signal of growth.

Figure 7.4 shows the overall difference on quantile ranks and its decomposition into wage structure and composition effect. A way to interpret this graph is the following: it tells the amount (negative in this case) to add to each 2000 quantile rank to know which quantile rank in 2010 corresponded to that q-quantile. For example, the overall difference at the 40th 2000 quantile is about 0.25; this means that that 40th quantile became the 15th[22] in 2010.

The same applies to the wage structure and to the composition effect. The former is the main explanation for the overall difference: the reward for workers' personal characteristics increased significantly on the whole 2000 distribution, but more relevantly in its central part. The composition effect moves in the same direction, but with a more modest contribution, which stays nearly constant around 0.03 under both measures.

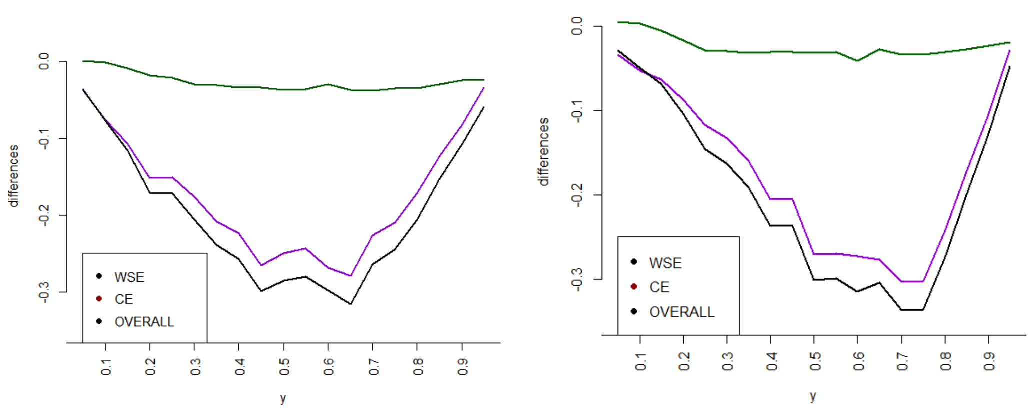

Figure 7.4:
Aggregate decomposition, hourly and monthly wages

[22]0.40-0.25=0.15

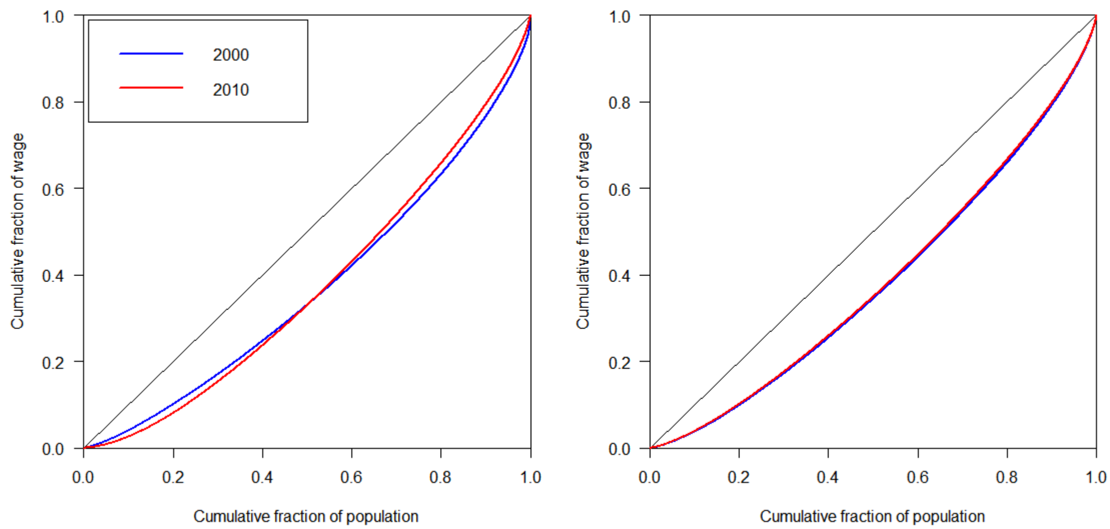

Figure 7.5:
Lorenz Curves, hourly and monthly

However, our focus is not on wage growth but on wage inequality. How can we use the decomposition in this sense?

As a preliminary note, we consider a general indicator as the Gini Index to have a broad picture of what happened to inequality in the time span of our interest. Figure 7.5 plots the Lorenz Curves in hourly and monthly terms. The monthly curves are almost non-identifiable, but the red one is slightly higher than the other, so we can conclude that inequality barely decreased as the Gini Index in the table below also suggests.

The hourly curves are more interesting because they cross. The left side of the 2000 curve is above the 2010 one (closer to the 45° line), suggesting that the 2000 distribution is more equal, but the contrary occurs at the other end. The Gini Indexes tell that inequality decreased, but we cannot rank the states of the world in a dominant manner. As we said in the "Measures" section, the Gini Index is weak in discriminating among different types of inequality.

	Hourly	Monthly
2000	0.262	0.237
2010	0.256	0.228
Δ	-2.3%	-3.8%

In order to measure inequality and decompose its change, we drop the Gini Index and rely solely on percentile ratios, a modified version of the decile dispersion ratios: instead of taking, for example, the average wage of the richest 10% and the poorest 10%, we consider simply the 10th and the 90th quantiles. The use of four ratios (90/10, 50/10, 90/50, 75/25) is functional to check explicitly for changes in inequality in particular points of the distribution. Moreover, ratios allow to neglect the effect of inflation since it affects in the same way the whole distribution.

Notice that, while regressions and graphs consider log wages, percentile ratios are computed using linear wages in order to ease the interpretation of the ratios themselves. A 90/10 ratio of 2, for example, simply means that the 90th quantile is the double of the 10th quantile.

First, in Figure 7.6 the bootstrapped quantile growth rates are plotted. They agree with the Gini Index results. The trend in both lines is decreasing: growth declined moving toward top wages, thus inequality decreased.

The hourly trend line is convex, the monthly one is concave. Actually, the hourly growth rates are nearly flat between the 50th and 85th quantiles.

A second remark is that the points in Figure 7.6a are closer to the trend line, while the one in Figure 7.6b are more scattered and difficult to interpret. Finally, the hourly wages increased more than the monthly ones. This is due to a general decrease in the number of hours worked (the average fell from 42.7 to 40.7). The clear divergence at the extremes is due to the inverse-U-shaped, negative growth rate trend of hours worked: workers in the middle of the distribution worked slightly less, and those at the extremes worked much less.

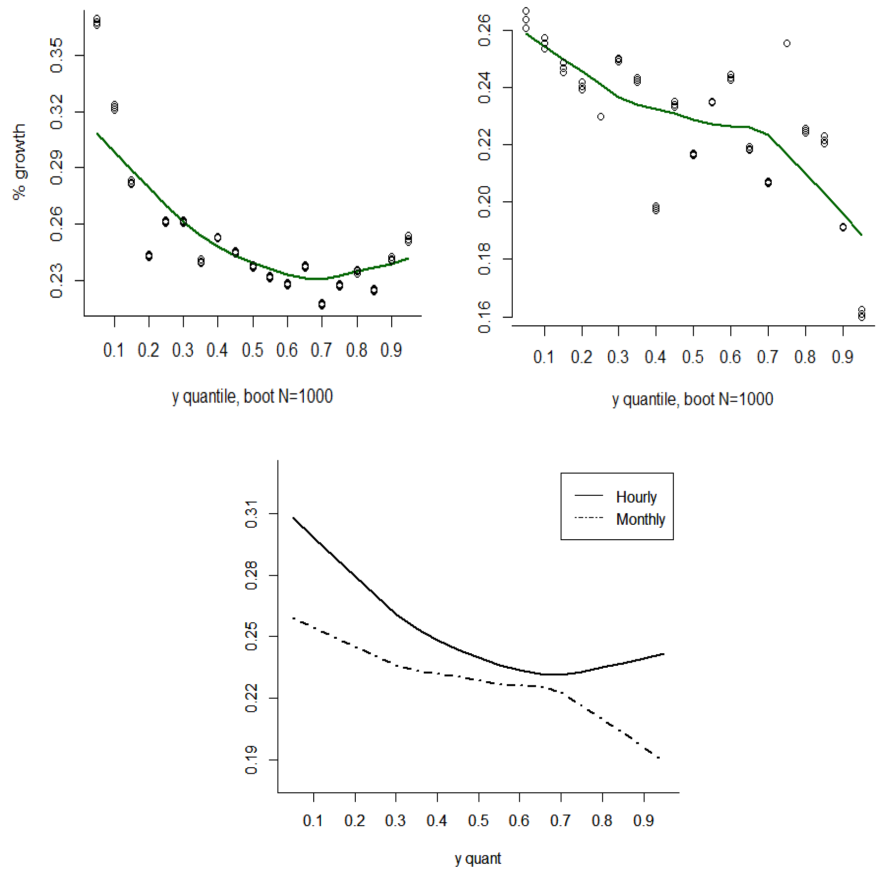

Figure 7.6:
Wage growth by quantiles with confidence intervals, hourly and monthly, and wage growth trends compared.

Going to the out-and-out analysis, we decompose the percentile ratios similarly to the cumulative distribution case:

$$PRatio_{10|10} - PRatio_{00|00} = [PRatio_{10|10} - PRatio_{00|10}] + [PRatio_{00|10} - PRatio_{00|00}]$$

In order to estimate the formula above, we need quantiles for the counterfactual distribution. They can be retrieved inverting the formula $F_{(00|10)}(y) = q$ into $F^{-1}_{(00|10)}(q) = y$ using a minimizing algorithm:

$$y* = \underset{y}{argmin} |F_{(00|10)}(y) - q)|$$

where q is the quantile rank required.

Bootstrapped results (N=100, as in CVM) are shown in the tables below. We report 2000 and 2010 ratios, the counterfactual ratios, the overall differences both in absolute and in percentage terms (relative to 2000 ratios), finally the wage structure and the composition effects.

Generally inequality diminished, with the exception of the hourly 90/50 and the monthly 75/25. However, if we split the ratios in two, the 90/10 and 50/10 on one side, the 90/50 and 75/25 on the other, we notice that, while the first couple experienced a more substantive decrease, the direction of the other's change is more uncertain and the overall difference is mild. The strong growth of the 10th quantile has a crucial role, as previously underlined. Therefore we can say that inequality stayed the same between most parts of the distribution, unless we consider the lowest quantiles.

When we look at the decomposition itself, the wage structure and the composition effect move in opposite directions and, in almost all cases, the wage structure is the prevailing one. We also find that the wage structure and the composition effect are statistically significant for the hourly ratios, while monthly ones often are not. Moreover the latter are always lower, in absolute terms, than the hourly ones. Non-significant measures at the 95% level are identified with an asterisk.

Aggregate decomposition, hourly ratios

	2000	Count.	2010	Δ	% 2000 ratio	Wage str.	Composition
90/10	2.82	4.82	2.61	-0.21	-8.0%	-2.21	2.00
50/10	1.63	2.57	1.49	-0.13	-8.7%	-1.08	0.94
90/50	1.74	1.88	1.75	0.01	0.6%	-0.13	0.14
75/25	1.60	2.41	1.55	-0.05	-3.2%	-0.86	0.81

Aggregate decomposition, monthly ratios

	2000	Count.	2010	Δ	% 2000 ratio	Wage str.	Composition
90/10	2.83	2.93	2.65	-0.18	-6.8%	-0.28*	0.10*
50/10	1.77	1.81	1.70	-0.07	-4.1%	-0.11*	0.04*
90/50	1.60	1.52	1.56	-0.04	-2.6%	0.04*	-0.08
75/25	1.50	1.41	1.54	0.04	2.6%	0.13	-0.09

As previously noted, the dataset composition did not change much from 2000 to 2010. This observation, coupled with an observed generally low composition effect, suggests that the variation in the distribution of each covariate had a minimal impact on the change in inequality. Therefore, we are not further investigating on this side.

We proceed with a detailed decomposition for the wage structure effect, instead. We consider in particular a subset of variables of interest: gender, three education dummies and the fixed-term contract.

The previously made observations are valid also in this case. Once again, in fact, all effects for hourly ratios are statistically significant while some are not in the monthly case.

Differences between the 90/10-50/10 couple and the 90/50-75/25 one remain: the latter experiences a weak decrease in inequality or even an increase, while the former shows a clear, strong inequality decrease (except for the monthly 90/10 effects).

The "negative reward" from being a woman[23] decreases inequality. This is a common finding in the inequality decomposition literature[24]; the gender coefficient usually results increasingly negative moving towards top quantiles and this makes the distribution less dispersed.

On the contrary, education contribution to wages is higher for top quantiles, boosting inequality as one can infer from our table for the middle-top part of the distribution in the monthly case. The same reasoning possibly applies for the fixed-term contract: its contribution is less negative moving towards top quantiles and the wage dispersion rises.

Wage structure Detailed Decomposition, hourly ratios

	Women	Middle School	High School	Master Degree	Fixed-term
90/10	-1.94	-2.16	-1.93	-2.03	-2.08
50/10	-0.92	-1.05	-0.93	-1.00	-1.03
90/50	-0.14	-0.13	-0.12	-0.11	-0.12
75/25	-0.62	-0.66	-0.62	-0.63	-0.55

Wage structure Detailed Decomposition, monthly ratios

	Women	Middle School	High School	Master Degree	Fixed-term
90/10	-0.62	-0.27*	-0.34	-0.28*	-0.13*
50/10	-0.35	-0.32	-0.33	-0.36	-0.24
90/50	-0.01*	0.27	0.22	0.30	0.29
75/25	-0.16	0.14	0.10	0.15	0.16

[23] In the "Residual analysis" subsection, the female coefficient is significantly negative.
[24] [12]

8 Expansion

In this section a quick glance is cast on employees and self-employed together.

First, the percentage of self-employed in 2000 is 18.5 and 16.7 in 2010. Historically, Italy has always had a self-employment rate higher than other developed countries[25]. Scholars found reasons for this in the well-developed Italian entrepreneurial spirit, but also in the advantages of fiscal evasion, much easier than for self-employed to achieve and prosecuted ineffectively, and in the fragmentation of activities in small firms, due to a legislation that protects them from large-size competitors (until recent times, for example, market regulation has discouraged the spread of chain stores.) Moreover, some contractual arrangements created lately, such as the "continuous and coordinated contractual relationships" (*Contratto di collaborazione coordinata e continuativa*), have fostered fictitious self-employment. Under these contracts, employees are sometimes hired as self-employed in order to benefit from the fiscal reduction, to bypass employment protection legislation and national contract provisions.

As done previously, we start by looking at wage distributions and growth. The first striking feature is the flat shape of the self-employed density function. The standard deviation is nearly the double of the employees' one, wages are more dispersed and average income is slightly higher.

The black lines are the density functions for the whole dataset. Since the self-employed are relatively a small percentage, these lines are closer to the employees' ones.

[25] OECD Factbook 2010 – Economic, Environmental and Social Statistics, OECD, 2010

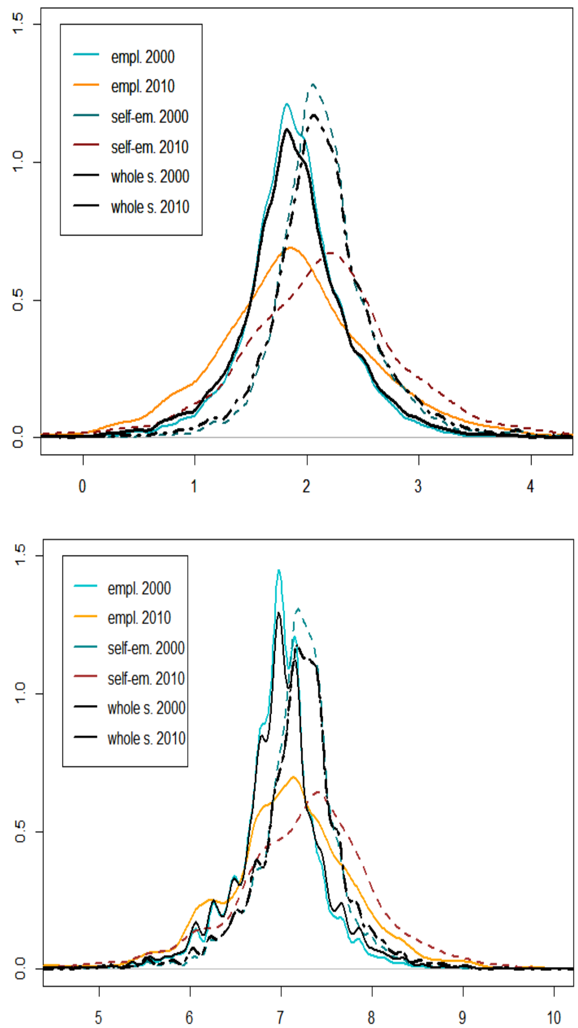

Figure 8.1:
Wage density functions for employees, self-employed and merged, hourly and monthly log wages.

The ratios confirm that inequality is critically higher in the self-employed group. Furthermore, it is always significantly increasing for the monthly case and often for the hourly one, especially when it concerns the lowest part of the distribution - exactly the opposite with respect to the employees' case. Actually, changes in wage inequality for self-employed and employees move in opposite directions in all cases except for the 75/25 ratios.

The Gini coefficient reports a slight decrease in inequality - inequality that is again substantially higher for self-employed.

Taken into account all those who earn a wage, inequality increases when considering the poor extreme of the distribution while it decreases when considering the central and upper part of it. The general picture provided by the Gini coefficient of an overall decrease in inequality hides a much more complicated reality.

The analysis of this section and of the previous one exactly points out that inequality has many faces and that one must carefully choose which one to look at. A general statement such as "inequality has increased", but also "inequality has decreased", is misleading because it does not define the measure we are considering (income, wage...) and it can mask severe differences in the subgroups disparities. A focus on within and between inequality is thus always necessary to provide a clear and truthful picture of the whole scenario.

Hourly		2000	2010	Δ	Monthly		2000	2010	Δ
90/10		3.21	3.46	7.7%	90/10		2.97	3.11	4.7%
	Empl.	2.82	2.61	-7.4%		Empl.	2.83	2.65	-6.4%
	Self-empl.	5.83	5.91	1.3%		Self-empl.	5.17	6.67	28.9%
50/10		1.73	1.92	10.9%	50/10		1.61	1.79	11.4%
	Empl.	1.63	1.49	-8.6%		Empl.	1.77	1.70	-4.0%
	Self-empl.	2.29	2.42	5.8%		Self-empl.	2.59	3.00	16.0%
90/50		1.85	1.80	-2.9%	90/50		1.85	1.74	-6.0%
	Empl.	1.74	1.75	0.6%		Empl.	1.60	1.56	-2.5%
	Self-empl.	2.55	2.44	-4.3%		Self-empl.	2.00	2.22	11.1%
75/25		1.72	1.60	-7.1%	75/25		1.62	1.54	-5.2%
	Empl.	1.60	1.55	-3.1%		Empl.	1.50	1.54	2.7%
	Self-empl.	2.33	2.27	-2.9%		Self-empl.	2.00	2.49	24.7%

Hourly		2000	2010	Δ	Monthly		2000	2010	Δ
		0.32	0.30	-6.3%			0.28	0.27	-3.6%
	Empl.	0.26	0.25	-2.3%		Empl.	0.24	0.23	-3.8%
	Self-empl.	0.50	0.46	-8.0%		Self-empl.	0.40	0.41	2.5%

9 Conclusions

In the introduction, we asked two questions: did the wage inequality increase or decrease for Italian employees between 2000 and 2010? How can we decompose such change, if any?

The answer to the first question is that, overall, inequality decreased. The level of nominal net wages actually increased, but at different speeds for different parts of the distribution. Hourly wages rose more than monthly ones, a difference that is due to a drop in the number of hours worked.

The quantile growth rates trended down as approaching top wages. Digging deeper, we see that the ratios considering the 10th quantiles are the ones experiencing the more substantive growth while inequality stayed practically unchanged for rest of the distribution.

In order to answer to the second question, we perform a decomposition analysis following the Chernozhukov, Fernandez-Val and Melly's approach: we build a semiparametric conditional distribution for 2000 and integrate it over the 2010 covariates. We use the resulting counterfactual distribution to estimate the wage structure effect (the effect of a change in the reward of the covariates) and the composition effect (the effect of a change in the distribution of the covariates themselves).

The rise in wages is mainly explained by the wage structure effect, while the composition effect is almost negligible. When analysing percentile ratios instead, the two effects partially offset each other (in most cases the wage structure effect favors a drop in inequality and the composition effect acts in the opposite direction), but the wage structure prevails in the end.

We also further decompose the wage structure effect, being the composition one nearly irrelevant, as the analysis of the covariates reveals. Female workers made the distribution more equal over time, while the education and the fixed-term dummies have two different effects: increasing inequality for most part of the distribution, decreasing it when the lowest quantiles are considered.

A quick glance has also been cast on self-employed inequality, which turns to be critically higher than for employees. Accounting for self-employment too shows that inequality increases when considering the 10th quantile and decreases for the rest of the distribution, overturning previously reported results.

It would be interesting to use more recent data in order to see how the crisis impacted inequality: has it squeezed all wages, leading to a fall, or has it hit only a part of them? In 2010 the devastating effects of the crisis on real economy were not yet fully unchained.

Appendix

A Code

In this appendix you find the code to perform the decomposition described in Chernozhukov, Fernanzed-Val and Melly (2009) and the one used to obtain the results disclosed in this paper.

A.1 Function.R

```
inner <- function (b, y0, Y, X)
  {
  # b is a vector (1xK) where K is the number of variables/coloumns in X
  # y0 is a scalar
  # Y is a vector (1xN), year 0
  # X is a matrix (NxK), year 0
  Lambda <- 1-exp(-exp(X%*%b))
  L0 <- log(Lambda)
  L1 <- log(1-Lambda)
  ind0 <- Y<=y0
  ind1 <- 1-ind0
  return(-sum(ind0*L0+ind1*L1))
  }

  fycondx <- function(y0, x0, Y, X, b00)
  {
  # Prob(Y<=y0 | X=x0) - equation 3.5 in CVM
  # y0 is a scalar
  # x0 is a vector (1xK)
  # Y is a vector (1xN), year 0
  # X is a matrix (NxK), year 0
  b0 <- b00
  return(1-exp(-exp(x0%*%b0)))
  }

  fycounter <- function (y0, X0, Y, X)
```

```
{
## calculate unconditional counterfactual distribution
## \int Pr(Y<=y0|X0) dF(X0) - equation 3.2 in CVM
# y0 is a scalar
# x0 is a matrix (NxK), year 1
# Y is a vector (1xN), year 0
# X is a matrix (NxK)), year 0
b00 <- optim(par=c(rep(0, ncol(X))), fn=inner,
y0=y0, Y = Y, X = X)$par
out <- apply(X0, 1, function(u) fycondx(y0, u, Y=Y, X=X, b00))
mean(out)
}

fy <- function (y,Y)
{
# empirical distribution function for the ys
# Pr(Y<=y)
# y is a scalar
# Y is a vector (1xN)
return (sum(as.numeric(Y<=y))/length(Y))
}

fnew<- function (y0, X0, Y, X,qi)
{
# difference between the counterfactual cdf for y0 and a given quantile rank
# y0 is a scalar
# qi is a scalar
#X0,Y,X as in 'fycounter'
abs(fycounter(y0,X0,Y,X)-qi)
}
```

A.2 Results code

```
# download functions from an external source
source('functions.R')
# data downloading
```

```
library(foreign)

# ask for the dataset of group/year 0
dataset=read.dta ("m00.dta")
attach(dataset)
# store all indep. variables of interest of year/group 0 in Px0
Px0=cbind(1,female ,mincerexp, mincerexp2, sud, north, elem, medie, prof, sup, trien,
uni, dott, det, inter)

# store the potential dep. variables of interest of year/group 0
ylm0 <-ylmmese
ylh0 <-ylmh
lnym0<-lnym
lnyh0<-lnyh
detach(dataset)

# ask for the dataset of group/year 1
dataset=read.dta("m10.dta")
attach(dataset)

# put all variables of interest of year/group 1 in Px1
Px1=cbind(1,female, mincerexp, mincerexp2, sud, north, elem, medie, prof, sup, trien,
uni, dott, det, inter)

# store the potential dep. variables of interest of year/group 1
lnym1<-lnym
lnyh1<-lnyh
detach(dataset)
rm(dataset)

# define the X matrix and the y vector for group/year 0 and 1
Y0<-lnyh0
Y0<-lnyh0
# create a vector of quantiles for which CDFs have to be evaluated
p<-seq(.05, .95, 0.05)
```

```
yrange <- quantile(Y0, p, names=FALSE)

# I - EVOLUTION OF WAGES

#create the counterfactual distribution for a range of values indentified by yrange
counter=length(yrange)
for (i in 1:length(yrange))
{
counter[i]=fycounter(y0=yrange[i], X0=Px1, Y=Y0, X=Px0)
print(counter[i])
}

#use the counterfactual vector to estimate the wage str. and the composition effect
wse=length(yrange)
cse=length(yrange)
for (i in 1:length(yrange))
{
wse[i]=fy(yrange[i],Y1)-(counter[i])
cse[i]=counter[i]-fy(yrange[i],Y0)
}

## Bootstrapped results
n <-100
set0=cbind(Y0,Px0)
coumatrix1<-matrix(0,length(yrange), n)
count<-1

for (i in 1:n)
{
# counterfactual distribution bootstrapped
set0b=set0[sample(nrow(set0), replace=T),]
Px1b= Px1[sample(nrow(Px1),replace=T),]
# resampling for data of year 0 are made together, indep and dep variables
for (i in 1:length(yrange))
{
coumatrix1[i,count]=fycounter(yrange[i],X0=Px1b,Y=set0b[,1],X=set0b[,2:16])
```

```
print (cbind(count,i))
}
count=count+1
}

coumean=apply(coumatrix1, 1, mean)
#standard error of ws
standcou=apply(coumatrix1, 1, sd)
# t-test for ws
tcou=coumean*sqrt(n)/standcou

# bootrapped wage structure and composition effect
wsmatrix<-matrix(0,length(yrange), n)
cematrix<-matrix(0,length(yrange), n)
count<-1
for (i in 1:n)
{for (i in 1:length(yrange))
{
wsmatrix[i,count]=fy(yrange[i],Y1)-coumatrix1[i,count]
cematrix[i,count]=coumatrix1[i,count]-fy(yrange[i],Y0)
}
count=count+1
}

#mean wage structure(ws) effect
matmeanws=apply(wsmatrix, 1, mean)
#standard error of ws
standws=apply(wsmatrix, 1, sd)
# t-test for ws
tws=matmeanws*sqrt(n)/standws
#mean composition effect (ce)
matmeance=apply(cematrix, 1, mean)
#standard error for the ce
standce=apply(cematrix, 1, sd)
#t-test for ce
```

```
tce=matmeance*sqrt(n)/standce

# II - Inequality

yq<-1:length(p)
for (i in 1:length(p))
{
# find the counterfactual quantiles minimizing fnew
yq[i]=optimize(fnew, interval=c(0,3),Px1,Y0,Px0,p[i], tol=0.00001)$minimum
} #interval is c(3,8) when considering monthly log-wages

#quantile ratios decomposition
set0=cbind(Y0,Px0)
set1=cbind(Y1,Px1)
pr=c(0.10,0.25,0.50,0.75,0.90)
ypr=c(1:length(pr))

#initializing matrices
ratios1<-matrix(NA,4,n)
ratios0<-matrix(NA,4,n)
ratios01<-matrix(NA,4,n)
ws<-matrix(NA,4,n)
ce<-matrix(NA,4,n)
over<-matrix(NA,4,n)

for (i in 1:n)
{
set0b=set0[sample(nrow(set0), replace=T),]
set1b=set1[sample(nrow(set1),replace=T),]
for (j in 1:length(ypr))
{
ypr[j]=optimize(fnew, interval=c(0,3),X0=set1b[,2:16],
Y=set0b[,1],X=set0b[,2:16],qi=pr[j], tol=0.0001)$minimum
print(cbind(i,j))
}
#2010 ratios matrix
```

```
ratios1[1,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.1))
ratios1[2,i]<-exp(quantile(set1b[,1],0.5))/exp(quantile(set1b[,1],0.1))
ratios1[3,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.5))
ratios1[4,i]<-exp(quantile(set1b[,1],0.75))/exp(quantile(set1b[,1],0.25))
#2000 ratios matrix
ratios0[1,i]<-exp(quantile(set0b[,1],0.9))/exp(quantile(set0b[,1],0.1))
ratios0[2,i]<-exp(quantile(set0b[,1],0.5))/exp(quantile(set0b[,1],0.1))
ratios0[3,i]<-exp(quantile(set0b[,1],0.9))/exp(quantile(set0b[,1],0.5))
ratios0[4,i]<-exp(quantile(set0b[,1],0.75))/exp(quantile(set0b[,1],0.25))
#counterfactual ratios matrix
ratios01[1,i]<-exp(ypr[5])/exp(ypr[1])
ratios01[2,i]<-exp(ypr[3])/exp(ypr[1])
ratios01[3,i]<-exp(ypr[5])/exp(ypr[3])
ratios01[4,i]<-exp(ypr[4])/exp(ypr[2])
#overall difference matrix
over[1,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.1))-
    exp(quantile(set0b[,1],0.9))/exp(quantile(set0b[,1],0.1))
over[2,i]<-exp(quantile(set1b[,1],0.5))/exp(quantile(set1b[,1],0.1))-
    exp(quantile(set0b[,1],0.5))/exp(quantile(set0b[,1],0.1))
over[3,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.5))-
    exp(quantile(set0b[,1],0.9))/exp(quantile(set0b[,1],0.5))
over[4,i]<-exp(quantile(set1b[,1],0.75))/exp(quantile(set1b[,1],0.25))-
    exp(quantile(set0b[,1],0.75))/exp(quantile(set0b[,1],0.25))
#wage structure effect matrix
ws[1,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.1))-exp(ypr[5])/exp(ypr[1])
ws[2,i]<-exp(quantile(set1b[,1],0.5))/exp(quantile(set1b[,1],0.1))-exp(ypr[3])/exp(ypr[1])
ws[3,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.5))-exp(ypr[5])/exp(ypr[3])
ws[4,i]<-exp(quantile(set1b[,1],0.75))/exp(quantile(set1b[,1],0.25))-exp(ypr[4])/exp(ypr[2])
#composition effect matrix
ce[1,i]<-exp(ypr[5])/exp(ypr[1])-exp(quantile(set0b[,1],0.9))/exp(quantile(set0b[,1],0.1))
ce[2,i]<-exp(ypr[3])/exp(ypr[1])-exp(quantile(set0b[,1],0.5))/exp(quantile(set0b[,1],0.1))
ce[3,i]<-exp(ypr[5])/exp(ypr[3])-exp(quantile(set0b[,1],0.9))/exp(quantile(set0b[,1],0.5))
ce[4,i]<-exp(ypr[4])/exp(ypr[2])-exp(quantile(set0b[,1],0.75))/exp(quantile(set0b[,1],0.25))
}

# 2000 bootstrapped ratios
```

```
avrat0<-apply(ratios0,1,mean)
# 2010 bootstrapped ratios
avrat1<-apply(ratios1,1,mean)
# counterfactual bootstrapped ratios
avrat01<-apply(ratios01,1,mean)
# overall difference bootstrapped
avover<-apply(over,1,mean)
# wage structure bootstrapped
avws<-apply(ws,1,mean)
# composition effect bootstrapped
avce<-apply(ce,1,mean)
# boostrapped standard deviation for oevrall difference, wage structure and composition effect
stover<-apply(over,1,sd)
stws<-apply(ws,1,sd)
stce<-apply(ce,1,sd)

# III - Detailed decomposition
set0=cbind(Y0,Px0)
#initializing matrices for 2010 and counterfactual ratios
ratios01md<-matrix(NA,4,n)
ratios1md<-matrix(NA,4,n)

for (i in 1:n)
{
set1b=set1[sample(nrow(Px1), replace=T),]
set0b=set0[sample(nrow(Px0), replace=T),]
Px1mod=Px1[sample(nrow(Px1),replace=T),]
# set[,X] gives the variable to replace, thus to consider in the detailed decomposition
middle=sample(set0[,3],nrow(Px1),replace=T)
Px1mod[,2]=middle
for (j in 1:length(ypr))
{
ypr[j]=optimize(fnew, interval=c(0,3),X0=Px1mod,
Y=set0b[,1],X=set0b[,2:16],qi=pr[j], tol=0.0001)$minimum
print(cbind(i,j))
```

```
}
#bootstrapped counterfactual ratios for the detailed decomposition
ratios01md[1,i]<-exp(ypr[5])/exp(ypr[1])
ratios01md[2,i]<-exp(ypr[3])/exp(ypr[1])
ratios01md[3,i]<-exp(ypr[5])/exp(ypr[3])
ratios01md[4,i]<-exp(ypr[4])/exp(ypr[2])
#bootstrapped2010 ratios
ratios1md[1,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.1))
ratios1md[2,i]<-exp(quantile(set1b[,1],0.5))/exp(quantile(set1b[,1],0.1))
ratios1md[3,i]<-exp(quantile(set1b[,1],0.9))/exp(quantile(set1b[,1],0.5))
ratios1md[4,i]<-exp(quantile(set1b[,1],0.75))/exp(quantile(set1b[,1],0.25))
}

avrat1md<-apply(ratios1md,1,mean)
avrat01md<-apply(ratios01md,1,mean)
dif=ratios1md-ratios01md
# wage structure effect for the detailed decomposition and standard deviation
mdif=apply(dif,1,mean)
sdif=apply(dif,1,sd)
```

Software used:

RStudio. 2012. RStudio: Integrated development environment for R (Version 0.96.122). Boston, MA

StataCorp. 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.

B Additional results

This section contains the 90/10 and 50/10 ratios and their evolution for each subgroup (only employees are considered).

By age group

Hourly	2000		2010			
	50/10	90/10	50/10	90/10	Δ50/10	Δ90/10
18-25	1.90	2.78	1.72	2.34	-9.5%	-15.8%
26-40	1.56	2.5	1.49	2.25	-4.5%	-10.0%
41-65	1.54	2.67	1.46	2.57	-5.2%	-3.7%

Monthly	2000		2010			
	50/10	90/10	50/10	90/10	Δ50/10	Δ90/10
18-25	1.88	2.60	2.00	2.72	6.4%	4.6%
26-40	1.71	2.57	1.76	2.50	2.9%	-2.7%
41-65	1.56	1.5	1.69	2.57	8.3%	2.8%

By gender

Hourly	2000		2010			
	50/10	90/10	50/10	90/10	Δ50/10	Δ90/10
Men	1.60	2.76	1.48	2.50	-7.5%	-9.4%
Women	1.56	2.8	1.5	2.66	-3.8%	-5.0%

Monthly	2000		2010			
	50/10	90/10	50/10	90/10	Δ50/10	Δ90/10
Men	1.49	2.58	1.46	2.30	-2.0%	-10.9%
Women	1.92	2.80	1.92	2.82	0.0%	0.7%

By education group

Hourly	2000		2010			
	50/10	90/10	50/10	90/10	Δ50/10	Δ90/10
Middle S.	1.60	2.33	1.43	2.00	-10.6%	-14.2%
High S.	1.62	2.59	1.48	2.38	-8.6%	-8.1%
Master D.	1.78	2.99	1.62	3.00	-9.0%	0.3%

Monthly	2000		2010			
	50/10	90/10	50/10	90/10	Δ50/10	Δ90/10
Middle S.	1.83	2.59	1.71	2.38	-6.6%	-8.1%
High S.	1.73	2.67	1.76	2.60	1.7%	-2.6%
Master D.	1.50	3.00	1.67	3.00	11.3%	0.0%

C Acknowledgments

First, I want to thank Prof. Giuseppe Ragusa, my supervisor, for giving me two among the most precious things a human can give to another: time and knowledge.

Thanks to my parents, who gave me - among other countless things - the love and the serenity to get to this point, and who, by removing many obstacles on my way, made learning nothing but a pleasure.

Thanks to my brother Michele, for being always ready for a videogame break when I'm tired with studying, and for being a young man I am proud of.

Thanks to all my family as well; their affection is distance-proof.

Thanks to Francesca, my soul mate, *dearest creature.*

Thanks to Jacopo for his kindness and dedication to all things he undertakes - and our friendship is not an exception. Thanks to Luca for sharing passions and quality time.

Thanks to the girls who came along in the undergraduate course and made a reluctant economist as I was happy to come to classes: Lucrezia, so gentle and yet so strong; Marta, whose cleverness is reached only by her kindness of heart; Federica, always wise and lovely; Sara, a ray of light no matter what; Martina, with her outstanding wit; Liron, sweet and caring.

Thanks to Dasha for all our interesting talks, and Francesca, who has always a good advice.

Thanks to Sebastián, a brilliant mind, a good friend, who shared with me unforgettable adventures in Paris - not least a 5-hour queue for the *Palais de l'Élysée.*

Thanks to Martina, Raluca, Ilaria - *Klimt, Sherlock and the Devil* -, Elena and Piero, who love important discussions as I do. Our friendship has not yet been proven by the passing of time, but why shouldn't we trust our present?

Finally, thanks to Fabrizio for a programming hint that is worth hours of computations.

Yes, they are a lot a people. I've always known I'm a lucky girl.

References

[1] Anthony B Atkinson. On the measurement of inequality. *Journal of economic theory*, 2(3):244–263, 1970.

[2] Alan S Blinder. Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, pages 436–455, 1973.

[3] Victor Chernozhukov, Ivan Fernandez-Val, and Blaise Melly. Inference on counterfactual distributions. 2009.

[4] Miles Corak. Inequality from generation to generation: the united states in comparison. *The Economics of Inequality, Poverty, and Discrimination in the 21st Century, ABC-CLIO, forthcoming*, 2012.

[5] Frank Alan Cowell. *Measuring inequality*, volume 2. Prentice Hall/Harvester Wheatsheaf London, 1995.

[6] Fernando G De Maio. Income inequality measures. *Journal of epidemiology and community health*, 61(10):849–852, 2007.

[7] John DiNardo, Nicole M Fortin, and Thomas Lemieux. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. Technical report, National Bureau of Economic Research, 1995.

[8] Gary S Fields. Measuring inequality change in an economy with income growth. *Journal of Development Economics*, 26(2):357–374, 1987.

[9] Sergio Firpo, Nicole M Fortin, and Thomas Lemieux. Unconditional quantile regressions. *Econometrica*, 77(3):953–973, 2009.

[10] Nicole Fortin, Thomas Lemieux, and Sergio Firpo. Decomposition methods in economics. *Handbook of labor economics*, 4:1–102, 2011.

[11] Tullio Jappelli and Luigi Pistaferri. Does consumption inequality track income inequality in italy? *Review of Economic Dynamics*, 13(1):133–153, 2010.

[12] José AF Machado and José Mata. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, 20(4):445–465, 2005.

[13] Ronald Oaxaca. Male-female wage differentials in urban labor markets. *International economic review*, 14(3):693–709, 1973.

[14] Owen A O'Donnell, Eddy KA van Doorslaer, Adam Wagstaff, and Magnus Lindelöw. *Analyzing health equity using household survey data: a guide to techniques and their implementation*. World Bank Publications, 2008.

[15] Joseph E. Stiglitz. The price of inequality. *New York Times*, 20-01-2013.

[16] Roberto Torrini. Cross-country differences in self-employment rates: The role of institutions. *Labour Economics*, 12(5):661–683, 2005.