Università degli Studi di Cagliari

DOTTORATO DI RICERCA IN

SCIENZE ECONOMICHE ED AZIENDALI

Ciclo XXIX

Essays on income distribution, economic growth, and taxation

SSD-SECS-P/03

Dottoranda: Valeria Zurru

Coordinatore: Professor Andrea Melis

Tutor: Professor Rinaldo Brau

Esame finale anno accademico 2015 – 2016
Sessione Settembre 2017

Abstract

This dissertation is the result of my experience as a PhD student taking part in the Doctoral Programme at the University of Cagliari. Recent public interest in tax policies which are contemporaneously able to promoting economic growth and reducing income inequality has motivated this research. In this spirit, Chapter 1 gives an overview of the state of art in the relationship between income distribution and economic development, while, the remaining chapters provide an empirical investigation of tax policies that are both growth-enhancing and inequality-reducing.

Chapter 2 investigates the impact of marginal taxation on economic growth, paying particular attention on the role of quality of institutions, in a panel data set of 26 OECD countries for the period 1981-2015. While dealing with omitted variables, measurement errors, and endogeneity issues, our analysis is carried out using fixed effects and generalized method of moment (GMM) techniques. Focusing on the direct effect of marginal tax rates on growth our results show no-significant evidence. Interestingly, fixed effects estimations reveal that the effect of marginal tax rates on growth changes due to the level of institutional quality: higher marginal tax rates appear harmful for growth in countries with institutional quality lower than sample average, as predicted by standard supply-side theory, but this distortionary effect of marginal taxation becomes negligible in right institutions. However, this evidence is not confirmed when GMM estimator is performed.

Chapter 3 examines the effect of joint taxation on female labour behaviour using micro data from the European Union Statistics on Income and Living Conditions (EU-SILC). We look at the impact of joint taxation on the probability of working comparing the female labour responses of a group treated with joint taxation (married women living in Germany) to the female labour responses of a group not treated with (unmarried women living in Germany and married/unmarried women living in Austria). Dealing with potential self-selection bias, this analysis is carried out using instrumental variables methods (two-stage least square and bivariate probit estimators) with cross-sections data in 2012. Our results are consistent with economic theory suggesting a negative effect of joint tax treatment of married couples on female labour decision.

Declaration

I declare that this thesis and the work presented in it are my own and have been written by me as the result of my own research.

I declare that Chapter 2 is a co-authored work with my supervisor, Professor Rinaldo Brau.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at the University of Cagliari.
2. Where I have consulted the published work of others, this is always attributed.
3. Where I have quoted from the work of others, the source is always given.
4. Where the thesis is based on work done by myself jointly with others, I have declared it.
5. None of this work has been published before submission.

Contents

Chapter 2

Chapter 3

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Alla mia Grande famiglia

A voi mamma e papà

Introduction

This dissertation incorporates three academic papers, which have been into three corresponding chapters. In general, all three chapters address a different topic, and they are independent. My primary research interest lies in evaluating the effect of taxation on macro-micro level issues, such as economic growth and female labour behaviour. In particular, recent public interest in tax policies which are contemporaneously able to promoting economic growth and reducing income inequality has motivated this research. In this spirit, the first chapter gives an overview of the state of art in the relationship between income distribution and economic growth, while, the remaining chapters provide an empirical investigation of two *win-win policies*, as the tax policies that are both growth-enhancing and inequality-reducing (Cingano, 2014).

Chapter 1 is titled **"Income distribution and economic growth: theory and empirical evidence".** This chapter provides an overview of the theoretical and empirical literature concerning the relationship between income distribution and economic growth. From a theoretical point of view, fisrt part of this chapter focuses on theories developed by Kaldor (1956), who suggested a positive correlation between growth and inequality, Kuznets (1955), who proposed the U-inverted hypothesis - i.e. economic development influences the income distribution - and last Galor and Zeira (1993) and Alesina and Rodrik (1994), who emphasized the existence of a negative effect of income inequality on growth. In this light, the second part of this chapter offers a picture of the empirical evidence that has been addressed in order to evaluate whether the economic development defines the income inequality level, as Kuznets suggested, or whether the income inequality is a key determinant of economic growth (either positive or negative).

Chapter 2 is titled *"***Marginal tax rate and economic growth: accounting for the quality of institutions".** Dealing with the huge increase in income inequality, it is suggested that income redistribution through taxes and transfers could take to remedy it. In this spirit, most of the recent literature claims rising in income tax progressivity, brought about by higher marginal tax rates (see e.g. Piketty, Saez, and Stantcheva, 2014; Cingano, 2014; Atkinson, 2015). However, there is a disagreement

about how marginal tax rates affect economic performance. On one hand, the traditional literature sustains that lower marginal tax rates are beneficial for economic growth due to efficiency gains in terms of higher labour supply by top earners (Lindsey, 1987; Feldstein, 1995). On the other hand, other scholars argue that a decrease in marginal income taxation is considered no-welfare improving because its impact on income inequality – cuts of top marginal tax rates are indicated as a driving force of the recent surge in upper incomes (Alvaredo, Atkinson, Piketty, and Saez, 2013; Piketty et al., 2014) – largely outweights increases in economic efficiency. Our second chapter builds on this field of study to face how marginal taxation affects economic growth.

In this field of study, one aspect that has not been deeply investigated is whether the quality of institutions influences in some way the impact of marginal taxation on long-term growth. In consideration of this, we try to assess whether some new light can be shed on the relationship between economic growth and marginal taxation by studying the interplay between quality of institutions and marginal taxation on long-term growth. Methodologically, to deal with omitted variables bias, measurement errors, and endogeneity issues, our growth empirical model is carried out using two different estimation strategies – Fixed-Effects and Generalized Method of Moments - for a panel data set of 26 OECD countries during the period 1981-2015.

Chapter 3 is titled **"Female labour decision and joint taxation: evidence from Germany".** In this chapter we focus on tax policies which might operate as a disincentive for female participation in economic life. Several scholars have provided theoretical and empirical evidence that gender inequality in employment affects negatively economic growth (see e.g. Seguino, 2000; Klasen and Wink, 2002; Seguino and Florio, 2003; Blackden, Canagarajah, Klasen and Lawson, 2007; Klasen and Lamanna, 2009). On the other, tax policies which cause gender discrimination in employment by leaving women out of the market generate higher income inequality due to increasing in unemployment rates (Berg and Ostry, 2011). Thus, tax policies which discourage female labour responses might contribute to more income inequality and at the same time might operate as limitation of economic growth.

In the last decades, many efforts have been done into the attempt to stimulating the labour effort participation of women. However, countries show large variations in

the women labour supply. In some countries the female labour supply tends to converge to the men level whereas it has not changed remarkably in other countries. Differences in income tax regimes are proposed as an explanation of these divergent behaviors. Steiner and Wrohlich (2008 p.116) sustains that *"[ ] low labor-force participation rate of married women in Germany is closely related to the negative labor-supply incentives for secondary earners implied by the tax system."* In this spirit, the last chapter of this dissertation is devoted to investigating the effect of joint taxation on female labour decision.

We look at the impact of joint taxation on the probability of working comparing the female labour decision of a group treated with joint taxation (married women living in Germany) to the female labour response of a group not treated with (unmarried women living in Germany and married/unmarried women living in Austria). In fact, Germany applies a joint tax system with income splitting for married couples, whereas Austria uses an individual fiscal schedule for single taxpayers as well as for married couples. Dealing with potential self-selection bias, this analysis is carried out using instrumental variables methods with cross-sections data in 2012.

Chapter 1

Income distribution and economic growth: theory and empirical evidence

1. Introduction

Scientists have discussed extensively the relationship between inequality and economic development. This is an issue that has received attention since ancient times. Jean-Jacque Rousseau, in his second seminal discourse, titled *"The Discourse on the Origin of Inequality"* (1775), stated that

"The first man who, having enclosed a piece of ground, bethought himself of saying "This is mine", and found people simple enough to believe him, was the real founder of civil society. From how many crimes, wars and murders, from how many horrors and misfortunes might not any one have saved mankind, by pulling up the stakes, or filling up the ditch, and crying to his fellows, "Beware of listening to this impostor; you are undone if you once forget that fruits of the earth belong to us all, and the earth itself to nobody." But there is great probability that things had then already come to such a pitch, that they could no longer continue as they were; for the idea of property depends on many prior ideas, which could only be acquired successively, and cannot have been formed all at once in the human mind. Mankind must have transmitted and increased from age to age, before they arrived at this last point of the state of nature (Rousseau, 1775, translated by Cole, G.D.H., p. 23).

Though Rousseau might be not the first that faces the question of inequality, to our knowledge, this is the first work that explains the inequality as a consequence of the economic and social process, instead of a natural status.

In the last century, numerous efforts have been made in order to formalize the relationship between income distribution and economic development, and to explain the causality that links these two variables. However, in the last two decades this issue has received far more attention than has in the past (Atkinson, 1997, 2015). This renewing interest has been motivated by the 2008 financial crisis, which

has put economic growth at the forefront of the public debate, and by the long-lasting increase in income inequality that many OECD countries have experienced in last decades. On this point, Saez (2015) sustains that in the United States the income inequality has reached levels not seen since the late 1920s. Also Alvaredo, Atkinson, Piketty, and Saez (2013) highlight that the upper income share in the United States has risen sharply since 1970; it mostly depends on increasing in the top percentile income share, which has increased from 8.9% in 1976 to 18.9% in 2010 (Alvaredo et al., 2013).

Economic growth has always been the primary objective of economists over time; recently this has been also the case for the income distribution issue. The invigoration of the interest in income distribution is confirmed by the several works that have been devoted to this topic in last decades (see e.g. Stiglitz, 2012; Piketty, 2014; Atkinson, 2015; Atkinson and Bourguignon, 2015) as well as by the inclusion of this issue in the political agenda of many OECD countries. On December 4, 2013 in a speech sponsored by the Center for American Progress, the 44° President of the United States Barack Obama affirmed that the increasing income inequality is *"the defining challenge of our time"*.

Particularly, two main points of concerns in the literature have been raised. The first point of concern regards the causality between income inequality and economic growth. Does economic development determine income distribution? Does the income distribution affect economic growth? Second, the literature debates about the sign of the effect (if there is any) of income inequality on economic growth. Is there a positive association? Is income inequality harmful for growth? In order to address these questions, a wide body of theoretical and empirical literature has been advanced.

Following Galor' (2011) classification, this extensive theoretical and empirical literature might be divided into three main branches: the "Classical", "Neoclassical", and "Modern" approaches. The first approach, initiated by the studies of Lewis (1954) and Kaldor (1956), aims at proving that income inequality is beneficial for growth. On this point, we concentrate on the theory of Kaldor (1956) who indicates the marginal propensity of saving as an explanation of this positive association between income distribution and economic growth.

The Neoclassical approach rejects the hypothesis that income distribution affects the economic development process. This body of studies can be traced to the seminal work of Kuznets (1955) who sustains that the level of economic development determines the grade of income inequality in the long-term. Particularly, according to Kuznets (1955), the income distribution first, in an early economic phase, increased, then reached a peak, and finally shrank. A comprehensive survey on this topic has been provided by Fields (2001) who concludes that the degree of income distribution is not determined by the *stage* of economic growth.

The Modern approach has been developed in response to the hypothesis advanced by Kuznets, and more generally to the neoclassical viewpoint when suggests an irrelevance of the income distribution on the determination of economic growth. Galor and Zeira' (1993) and Alesina and Perotti' (1993) studies might be considered the pioneers of this literature. This approach aims at investigating how income distribution affects economic development. The dominant view is that the income distribution affects negatively growth. Particularly, two main lines of research in the literature have been proposed to explain this negative correlation: the non-political and the political economy. The non-political line of research we focus on is based on the "imperfect capital market" transmission channel. The idea is that income distribution affects negatively economic development via inefficient investment level in human capital due to the presence of imperfect capital market, which causes limitations on the option of borrowing for acquiring education, especially for disadvantaged individuals (Galor and Zeira, 1993; Galor and Moav, 2004).

At the risk of some oversimplification, the political economy line can be divided into two main channels: the "fiscal policy" and the "socio-political instability". In the fiscal policy channel, taxpayer preferences on taxation and redistributive expenditure levels are inversely related to his income level. Namely, the lower the taxpayer income is, the higher the preferred taxation and government expenditure levels are. In other words, considering a progressive taxation system, and given the median voter theorem, in equilibrium the lower the median income is, the higher the distortionary tax rate and expenditure levels are. In turn, higher taxation and government expenditure levels translate into lower economic prosperity. Thus, in the fiscal policy channel an increase in income inequality generates an increase in

distortionary taxation and redistributive government expenditure, and these latter redistributive fiscal policies cause a decrease in economic growth. In summary, in the fiscal policy channel an increase in income inequality affects negatively growth via its effects on redistributive public expenditure and distortionary taxation (Bertola, 1993; Perotti, 1993; Alesina and Rodrik, 1994; Persson and Tabellini, 1994).

In the socio-political instability channel the negative impact of an unequal income distribution on growth is driven from social unrest. In an economy with sustained unequal distribution of resources the arisen of socio-political instability - the establishment of interest groups who are able to extract part of someone else's wealth either by political manipulation or by rent-seeking behaviors - is more likely. In turn, the arisen of socio-political instability discourages investment, thus slow down economic growth. Hence, socio-political instability increases as inequality increases, and economic growth decreases as socio-political instability increases. Several works have formalized this transmission channel (e.g. Benhabib and Rustichini, 1996; Alesina and Perotti, 1996; Perotti, 1996). For the sake of brevity, in this study we focus on the fiscal policy channel.

Our survey shows in general mixed evidence, which differs across countries, time-span investigated, methodology employed, and especially database used for measuring income inequality.

The remainder of this chapter is organized as follows. Section 2 presents the Classical approach by illustrating the theory of Kaldor. Section 3 focuses on the Kuznets hypothesis, and the theoretical applications that have been suggested to explain how the level of economic development might determine the income inequality. Section 4 discusses the fiscal policy channel. Section 5 reports the main empirical findings. Finally, section 6 concludes this chapter.

2. Classical approach: the theory of distribution of Kaldor

The Classical approach suggests that income inequality is beneficial for economic growth. Many theoretical studies have predicted this positive association (see. e.g. Lewis, 1954; Kaldor, 1956; Pasinetti, 1962; Li and Zou, 1998). For instance, Nicholas Kaldor (1956) made a contribution to this theory of income distribution with his work "Alternative Theories of Distribution". In this theoretical study, he indicates the

marginal propensity of saving as an explanation of the possible positive association between income distribution and economic growth.

Kaldor's model assumes that there are only two employment classes: capitalist and workers. Then, the total net income *(Y)* is divided into two categories: wages *(W)* and profits *(P)*; and also the total net saving can be expressed into two broad forms: workers' savings and capitalists' savings. Moreover, he supposes that each class has a specific propensity to save: workers' propensity to save *(S_w)* is lower than capitalist' propensity to save *(S_c)*. In mathematically terms, it can be expressed as follows

$$Y = W + P \tag{1}$$

$$S = SwW \ + \ ScP \tag{2}$$

$$Sw \ < \ Sc \tag{3}$$

In this model, the amount of physical capital necessary in order to satisfy full employment condition is externally given, and is independent on propensities to save. In Kaldor's model, under full employment condition, equalizing saving and investment - condition of dynamic equilibrium - yields the only possible distribution of total income between wages and profits. From this investment-savings equilibrium condition follows that

$$I = S \tag{4}$$

$$I = SwW \ + \ ScP = \ SwY \ + (Sc - Sw)\ P \tag{5}$$

Then, the profit share in total income at which the equilibrium condition (4) remains satisfied over time can be expressed as follows

$$\frac{P}{Y} = \ \frac{1}{Sc-Sw}\ \frac{I}{Y} \ - \frac{Sw}{Sc-Sw} \tag{6}$$

The equation (6) shows the only distribution of income and share of profit which will keep the system in equilibrium over time.

In particular, the coefficient $\frac{1}{Sc-Sw}$ measures the changes in the distribution of income due to a unitary change in the investment rate. Assuming that Sw is very small, nearly zero, the equation (6) simply reduces to

$$P = \frac{1}{Sc} I \quad (7)$$

Kaldor states that the income distribution and the rate of profit above defined will not only exist but also will be the ones that the system tends to produce.

In Kaldor's model, thus, there is a specific distribution between wages and profits that makes savings equal to investment, and the share of profit is determined by the rate of investment. In particular, an increase in investment forces an increase in savings to keep the system in equilibrium. Since capitalists are assumed to have higher propensity to save, the system remains in equilibrium by increasing profits. This suggests a positive relationship between economic growth and inequality in the distribution of income. In the empirical evidence reported by Li and Zou (1998) and Forbes (2000), income distribution affects positively economic growth. This sign seems to be consistent with Kaldor's theory.

3. Neoclassical approach: the Kuznets hypothesis

The Neoclassical approach posits that income distribution does not affect the economic growth process. This literature can be traced down to the hypothesis postulated by Kuznets (1955), who states that the causal direction of this relationship goes from economic development to income distribution.

In his seminal paper about the relationship between the economic development and the income inequality, Kuznets argues that there was *"[ ] a long swing in the inequality characterizing the secular income structure: widening in the early phases of economic growth when the transition from pre-industrial to the industrial civilization was most rapid; becoming stabilized for a while; and then narrowing in the later phases"* (Kuznets, 1955, p.18). Therefore, according to Kuznets' hypothesis, the income distribution first, increased, then reached a peak, and finally shrank. This non-linear relationship, which is named either U-inverted or Kuznets curve, is illustrated in Figure 1.

[FIGURE 1 ABOUT HERE]

Kuznets ascribes the rising in income inequality to two main forces, which are taken to be synonymous of development. The first factor identified as a cause of this long-term increasing is the concentration of savings into the upper-income earners. On this point, it is suggested that only the top decile of income distribution is predisposed to save, whereas below the top decile the propensity of saving is almost equal to zero.

Moreover, the results of the process of industrialization and urbanization, associated with moving away from agriculture sector, are recognized by Kuznets as the second source of the increase in income inequality. On one hand, the movement from agriculture to urban sector, accompanied with shifting from rural to urban population, determines the increase in urban habitant share, which is considered as the more unequal part. On the other hand, the average per capita income associated with rural population is indicated to be lower than that of urban population. If this is so, the movement from agriculture to urban area is a source of increasing in income inequality.

However, according to Kuznets, the income distribution in developed countries moves towards equality in the first half of the twentieth century. Considering the evidence of the distribution of income in the United States, Kuznets reports that the income share of the two lowest quintiles increased, moving from 13.5% in 1929 to 18% in the end of 1950's; while the top quintiles income share decreased from 55% to 44%. A similar pace is registered in the United Kingdom and in Germany. At the same time, this shrinking of inequality in developed countries goes together with a rise in real income per capita – except in world wars periods. Given these considerations, Kuznets postulated two questions: *"First, why does the share of the top-income groups show no rise over time if the concentration of savings has a cumulative effect? Second, why does income inequality decline and particularly why does the share of the lower-income groups rise if both the weight of the more unequal urban income distribution and the relative difference between per capita urban and per capita rural incomes increase?"*

He indicates fourth main factors which offset the increase in upper income share. First, the tendencies of rising in income inequality in the pre-industrial economies were reversed by government interventions through redistributive policies. Second,

a contrary impact is played by demographic factors concerning the differential in natural growth population rates between the upper and the lower income earners, which differences were in force in the developed countries in the beginning of the 90s. Third, the technological progress and the consequent obsolescence process triggered an intergenerational change, where rapid growth for younger industries were fostered, promoting the advance of the middle-class; while rarely entrepreneurs' sons were still successfully in charge. Finally, the interindustry shifts of workers from low-paid to high-paid occupations played in favour of lower income earners, which had greater possibilities to increase their earnings than higher income earners, which were already in high-paid occupations. According to Kuznets, all these factors played against the increase in top income distribution, and especially all they were elements of economic growth process. Therefore, the economic development of societies is the fostering as well as the offsetting source of the increase in income inequality, as Kuznets argued.

3.1 Theoretical formalizations of the Kuznets curve

At the core of the Kuznets' story is the "swing" in the long-term income inequality evolution. Supporters of the Kuznets hypothesis have tried to provide theoretical models which may generate and justify the U-inverted relationship between economic development and income inequality. Banerjee and Newman (1993), Galor and Tsiddon (1996), and Aghion and Bolton (1997) have formalized economic mechanisms that are consistent with the evolution of income distribution predicted by Kuznets.

3.1.1 Banerjee-Newman model

Banerjee and Newman (1993) (BN) suggest that the explanation of the Kuznets curve may rely on the interaction between the occupational structure and the development process. BN assume a labor market with four occupational options: subsistence, working, self-employment, and entrepreneurship. Because of capital market imperfections, people's capacity to borrowing is limited; thus, high levels of investments, which are required for starting up both self-employment and

entrepreneurship occupations, are out of reach for poor people (*w** indicates the minimum wealth value necessary to being eligible for a loan large enough to self-employment, while *w*** is the minimum wealth level needed to being eligible for a loan large enough to entrepreneurship). Consequently, the initial income distribution determines the pattern of occupational choice: poor people are more likely to work for other wealthier employers instead of becoming either self-employed or entrepreneur in large-scale productions; while, rich people are more likely to choose wealthier occupations. In turn, the choice of occupations affects the people's propensity to save and to risk, and then determines the new distribution of income.

BN describe the three wealth intervals (*[0, w*), [w*, w**), [w**,$\overline{w}$]*) with three wealth classes: lower (L), middle (M), and upper (U); and identify the wealth distribution, that is the ratio of population of each class, by probability vectors ***p*** = (p_L, p_M, p_U). Depending on the ratio of the number of people in lower class and the number of people in the upper class the equilibrium wage and the occupation structure are defined. They stress that what matter is the initial income distribution in the sense that the occupational choices made by individuals, and then the new income distribution, rely only on the initial distribution of wealth and especially on the belonging wealth class, and are not affected both by individual abilities and by individual preferences. According to BN's model, thus, the long-term dynamic equilibrium of the economy depends on the initial income distribution.

To examine the equilibrium dynamics by using a phase diagram, BN derive the following system of linear differential equations:

$$\dot{p_L} = \begin{cases} 1 - q - (1-q)p_L + (q - q')p_U, & p_L > \mu p_U \\ 1 - q - \left(2 - q + \frac{q'}{\mu} - \frac{q}{\mu}\right)p_L, & p_L < \mu p_U \end{cases}^{1} \quad (1)$$

and

$$\dot{p_U} = \begin{cases} q - qp_L + (q' - q - 1)p_U, & p_L > \mu p_U \\ q - \left(q + \frac{q'}{\mu} - \frac{q}{\mu}\right)p_L - p_U, & p_L < \mu p_U \end{cases} \quad (2)$$

[1] $\mu > 1$ is a parameter indicating the maximum number of projects/workers that an entrepreneur can perfectly monitor. If the project succeeds, entrepreneur earns a random return *rl,* where *r* is r_0 or r_1 with probabilities *1-q* and *q*. Workers who undertakes projects generate random return *r'l'*, where *r'* is r'_0 or r'_1 with probabilities *1-q'* and *q'*.

In Figure 2, BN show that the economy might follow a development path as predicted by Kuznets. If we start at the point Y, the economy achieves a stable point of prosperity equilibrium (P) after an initial inequality increase and a subsequent decrease. BN state that *"Thus, as Kuznets suggested, while mean wealth rises along the entire development path, inequality first increases and then decreases"* (Banerjee and Newman, 1993 p. 295).

[FIGURE 2 ABOUT HERE

3.1.2 Aghion-Bolton theory

Aghion and Bolton (1997) propose the difference in investment behaviour among the poor and the rich due to capital market imperfection as an explanation of the Kuznets curve. The AB's idea is that the richer people have higher capacity to accumulate capital than poorer people. This difference in investment behaviour, which derives from imperfections of capital market, causes the increase in income inequality in the early stage of development; in turn, the accumulation of capital triggers a wealth trickle-down in favour of poorer people such that the income inequality narrows in the latter stage of economic development, as predicted by Kuznets.

In order to prove this Kuznets effect, Aghion and Bolton (1997) (AB) develop a closed-economy model where the agents do not differ for both preferences and abilities, but they diverge only for initial wealth distribution. In each period *t* an agent may choose either to take up a worker activity which does not involve capital investment or to undertake an entrepreneurial activity which requires a fixed capital investment. The investment return is given by

$$F(k,l) = \begin{cases} r \text{ with probability } p \\ 0 \text{ with probability } 1-p \end{cases} \quad \text{if } k \geq \hat{k} = 1 \text{ and } l \geq 1 \ ; \tag{1}$$

$$F(k,l) = 0 \text{ otherwise} \ ; \tag{2}$$

and the effort cost of getting the investment return *r* with probability *p* is denoted by the following quadratic cost function

$$C(p) = \frac{rp^2}{2a} \quad , \qquad \text{where } a \in (0,1] \tag{3}$$

Each agent either can use own initial endowment of capital to undertake an entrepreneurial activity or he can invest it in a mutual fund with gross return A_t, which is assumed to be certain. Chronology of agents' decision evolves in three steps: first, an agent decides how allocates her initial wealth and her labour unit; second, she obtains the investment return, if any. Then, at the end of lifetime, she decides how allocates the net wealth between consumption and bequests.

AB assume three wealth classes according to the initial wealth level, and prove that the agent's decision either of lending or borrowing depends on their initial wealth level. The agents with initial wealth $w > 1$, named the very wealthy lenders, can afford to invest both in their own entrepreneurial activity and in activity undertaken by other entrepreneurs via investment in capital market; the middle-class borrowers who have an initial wealth $w \in (\widehat{w}, 1)$ choose to start up their own entrepreneurial activity, though they need to obtain a loan as their initial wealth is not sufficient to cover the fixed capital cost of the investment; finally, the agents with initial wealth $w < \widehat{w}$, named the poor lenders, cannot afford to invest in their own activity.

They also assume that the probability of success *p* is not observable, and that the borrowers cannot repay her lenders more than her end of period wealth. Given these assumptions, AB suggest that an optimal lending contract have a repayment schedule *R(w)* such that

$$R(w) = \begin{cases} (1-w)\rho(w) & \text{if the project succeeds ;} \\ 0 & \text{if the project fails ;} \end{cases} \tag{4}$$

where $(1-w)$ is the amount of capital borrowed and $\rho(w)$ is the rate of repayment.

Borrower maximization problem is given by

$$\max_p \{pr - p(1-w)\rho(w) - C(p)\} \tag{5}$$

and the solution is given by

$$p(w) = a\left(1 - (1-w)\frac{\rho(w)}{r}\right) \tag{6}$$

From Equation (6) AB affirm that the bigger the amount of capital borrowed to invest, the lower the supply effort because a larger fraction of marginal effort returns are earned by the lenders; whereas the very wealthy, who can undertake an entrepreneurial activity without borrowing, supply the first-best level of effort because they gain all returns from such effort.

In equilibrium, the mutual fund return A_t must be such that

$$p(w)\rho(w) = A_t \tag{7}$$

Given Equation (6), it implies that

$$ap(w)\left[1 - (1-w)\frac{\rho(w)}{r}\right] = A_t \tag{8}$$

Form Equations (6), (7) and (8), AB derive that "*in equilibrium effort supply is increasing in w:* $p'(w) > 0$.

Since the effort supply is decreasing with the increasing in capital borrowed, the unit repayment $\rho(w)$ is increasing with the increasing in capital borrowed. As a consequence, all agents with wealth $w \in [0, w(A_t)]$ with $w(A_t) \equiv 1 - \frac{ar}{4A_t}$ cannot borrow though they want to do. This condition is named *credit rationing*.

However, in equilibrium when the capital cost A_t is high it may happen that no-credit rationing occurs because all agents with wealth $w \in [0, w(A_t)]$, the very poor, may prefer to lend. AB affirm that *there is credit rationing in equilibrium whenever* $\frac{ar}{4} < A_t < \frac{3}{8}ar - n - 1$.

In words, the poor lenders prefer to lend when the market return is high; while, when the capital cost is low, and then favorable to borrowers, they prefer to borrow but the credit access may be denied if their initial wealth level is insufficient to guarantee the investment return, as defined by the credit rationing rule.

Importantly, AB's proposition that credit rationing occurs when capital cost is low differs from the credit rationing predictions existing at that time. This divergent argument depends on the endogenously (instead exogenously) agents' decision of been either borrowers or lenders. Thus, in contrast to the referring literature, AB endogeneize the capital cost schedule which they sustain "*is determined*

endogenously by the interplay between the supply and demand for investment funds" (Aghion and Bolton, 1997 p. 152).

Endogeneizing the capital cost schedule implies that the evolution of the income distribution over time is influenced by the equilibrium cost of capital, and particularly entails an economic mechanism explaining the Kuznets curve.

Considering an economy in rapid capital accumulation, AB assume that the investment returns (*r)* and the saving propensity rates *(1-*δ) are both enough high such that

$$\frac{3}{8}ar(1-\delta) > 1+n. \tag{9}$$

They argue it implies that *"…whenever the equilibrium cost of capital* A_t *is strictly greater than 1,* A_{t+1}*will necessarily be strictly lower than* A_t *and more risky projects will be financed in period* $t+1$*. [ ]. Once all investment opportunities will end up being exploited growth tapers off and the cost of capital stays at the lower bound* $A=1$*"* (Aghion and Bolton, 1997 p. 161).

Therefore, the evolution of equilibrium cost of capital entails a trade-off between growth and inequality, à la Kuznets. Indeed, when in the early stages of economic development the cost of capital is high, and then favorable to lenders, the wealthy lenders ($w>1$) accumulate capital relatively faster. Consequently, the early stages of economic development are associated with a widening inequality, and a growth going up quickly. In the later stages of development the accumulation of capital determines that the cost of capital decreases becoming more favorable to borrowers. This mechanism triggers a trickle-down in favour to poorer people: the middle-class wealth increases and approaches to the wealth level of upper-wealth class, and the poor lenders' capacity to borrow increases such that they can invest in their own entrepreneurial activity. Consequently, the later stages of economic development are associated with a narrowing inequality and a growth tapering off. As AB argued, *"initial phases of growth tend to increase inequalities while later stages tend to reduce them"* (Aghion and Bolton, 1997 p. S162).

AB point out that their goal is not just to develop a model explaining the Kuznets curve, but it is also to present a model explaining the productivity efficiency of redistribution policies. They sustain that the trickle-down mechanism lowers the

moral hazard problem which causes an underinvestment level by poorer people, who, they state, *"need to borrow funds to invest"*, with respect to the first-best economy. They argue that *"There is a natural policy response available to correct this productive inefficiency: redistribute wealth permanently from the rich lenders to the poor and the middle-class who need to borrow funds to invest. The positive effect of such redistribution is to equalize opportunities by letting all agents have access to profitable activities on similar terms* (Aghion and Bolton, 1997 p. S166). Therefore, according to AB' model, redistribution is a desirable policy not just because of its relevance in terms of Pareto-improvement, but because of its positive effect in terms of lower productive inefficiency.

3.1.3 Galor-Tsiddon model: wage differentials theory

Galor and Tsiddon (1996) (GT) advance a model where the wage differentials between skilled and unskilled workers lead to a Kuznets relationship between economic development and income distribution.

GT develop an overlapping-generations model, where individual preferences and human capital technology are identical across generations, and the only difference comes from the parents' level of investment in human capital, and thus in the efficiency units of labour employed in the production process.

GT suppose that individuals live for three periods. In each period a generation with N individuals is born. In the first period individuals invest in human capital. In the absence of income, they borrow capital at the market interest rate $\bar{r}$. In the second period individuals offers their efficiency units of labour, and in the third period they retire using their savings for consumption.

An individual i of generation t who parents' level of human capital is h_t^i and who invest x^i units of capital and one unit of labour in the acquisition of human capital skills, obtains h_{t+1}^i efficiency units of labour that are offered in the labour market in the second stage of life

$$h_{t+1}^i = \mu + g\left(h_t^i\right)\phi\left(x_t^i, 1\right) \qquad \mu > 0 \tag{1}$$

where

$$\phi\left(x_t^i, 1\right) = \left(x_t^i\right)^{\alpha} 1^{1-\alpha} \qquad \alpha \in (0,1) \tag{2}$$

and

$$g\left(h_t^i\right) = \begin{cases} (h_t^i)^{\beta} & \forall h_t^i \leq \tilde{h} \\ \tilde{h}^{\beta} & \forall h_t^i \geq \tilde{h} \end{cases} \qquad \beta \in (0,1) \tag{3}$$

Thus, even if investments in the acquisition of human capital are not performed, individual i of generation t is endowed with $\mu > 0$ units of efficiency labour at time $t+1$. Therefore, the efficiency of the labour units that each individual offers in the labour market might increase up to an upper bound $\tilde{h}$, and depends on two factors. First, it depends on the amount of physical capital invested in human capital formation by the considered individual *i*; second, it depends on the level of human capital of individual *i*' parents. GT assume that parents' human capital level has a direct effect on offspring education formation, due to higher returns to schooling friendly environment, and an indirect effect on society as a whole because the larger the parents' human capital level, the bigger the average level of human capital in society. Importantly, an increase in current human capital level causes an increase in the labour-technological progress in the next generation, and fosters the investment in human capital by next dynasties.

In each period a single homogeneous good is produced using capital and efficiency labour in the production process. The production function is given by

$$Y_t = F(K_t, \lambda_t H_t) \equiv \lambda_t H_t f(k_t) \qquad k_t = \frac{K_t}{\lambda_t H_t} \tag{4}$$

where K_t and H_t indicate the amount of capital and efficiency labour employed in the production process at time t, respectively; while λ_t is the coefficient of technological change at time t.

Assuming a stationary rental rate at level $\bar{r}$, the ratio of capital to efficiency labour units is stationary at a level $\bar{k}$ and the wage rate per efficiency labour is $w_t = \lambda_t w\left(\bar{k}\right) = \lambda_t \bar{w}$.

An increase in human capital in current generation, h_t, causes an increase in the labour-technological progress in the next generation, λ_{t+1}. In particular, it is assumed that the coefficient of technological progress λ_{t+1} is stationary at level λ^1 till the level of human capital is below a certain threshold $\hat{h}$; once this threshold is

reached, the level of technological change jumps to a higher stationary level λ^2, and will remain at this level. Therefore, the level of production technology at time $t+1$ can be expressed as

$$\lambda_{t+1} = \lambda(h_t) = \begin{cases} \lambda^1 & if\ h_t < \hat{h} \\ \lambda^2 & if h_t \geq \hat{h} \end{cases} \tag{5}$$

Individuals' preferences are given by

$$u\left((c_{t+2}^t)^i\right) \qquad (c_{t+2}^t)^i = \bar{r}\left(\lambda_{t+1}\bar{w}h_{t+1}^i - \bar{r}x_t^i\right) \tag{6}$$

where x_t^i is the level of human capital of individual i at time t. Given $\bar{r}$, $\bar{w}$, λ_{t+1} and h_t^i and considering the equation (1), each individual i of generation t chooses the level of human capital, x_t^i, maximizing the utility function:

$$x_t^i = \arg\max u\ \left\{\bar{r}\left\{\lambda_{t+1}\bar{w}\left[\mu + g\left(h_t^i\right)(x_t^i)^\alpha\right] - \bar{r}x_t^i\right\}\right\} \tag{7}$$

The solution of the previous maximization problem is given by the necessary and sufficient condition

$$x_t^i = \left(\frac{\alpha\lambda^j\bar{w}g(h_t^i)}{\bar{r}}\right)^{1/(1-\alpha)} \tag{8}$$

where

$$\lambda^j = \begin{cases} \lambda^1 & if\ h_t < \hat{h} \\ \lambda^2 & if h_t \geq \hat{h} \end{cases} \tag{9}$$

Considering equations (1) and (8), GT conclude that the dynamical evolution of the economic system is determined by the following nonlinear difference equation

$$h_{t+1}^i = \begin{cases} \mu + \left(\frac{\lambda^j\alpha\bar{w}}{\bar{r}}\right)^{\alpha/(1-\alpha)}(h_t^i)^{\beta/(1-\alpha)} & \text{if } h_t^i < \tilde{h} \\ \mu + \left(\frac{\lambda^j\alpha\bar{w}}{\bar{r}}\right)^{\alpha/(1-\alpha)}(\tilde{h})^{\beta/(1-\alpha)} \equiv G\left(\lambda^j\right) & \text{if } h_t^i \geq \tilde{h} \end{cases} \tag{10}$$

Thus, given λ^1, as illustrated in Figure 3, they suggest that there is a level of μ such that the system may generate multiple steady-state equilibria: $h^a(\lambda^1)$ and $h^c(\lambda^1)$ which are locally stable equilibria, and $h^b(\lambda^1)$ that is unstable. Given λ^2 the economic system reaches a single steady-state equilibrium $h^c(\lambda^2)$.

[FIGURE 3 ABOUT HERE]

GT affirm that each dynasty influences the average level of human capital, and in turn affects the human capital of next dynasties through its effects on the technological progress. They sustain that the interplay between internal-dynasty externalities and the aggregate externalities causes the dynamics of the economic system evolving in the Kuznets curve.

For simplicity, GT assume two types of dynasties: dynasty H and dynasty L, where the former has a higher level of human capital than the latter. For a given level of technological change, each dynasty behaves as the previous ancestors. Dynasty H considers optimal to increase the investment level in human capital, while, the dynasty L finds optimal to invest the same level of human capital as previous ancestors. Since the wage level of dynasty H follows a high-investment equilibrium path, whereas the wage level of dynasty L remains stable, the income inequality between these dynasties becomes greater. Moreover, the increase in the average level of human capital brings about the rise in output, as depicted in the initial path of Kuznets curve.

The increase in labour-technological progress from λ^1 to λ^2 , due to the investment in human capital in dynasty H, translates into being beneficial for the society as a whole by inducing a rise in the optimal investment level in human capital by less educated members of dynasty L. The dynamical evolution of the economic system converges to the single steady-state equilibrium $h^c(\lambda^2)$ where the aggregate output is higher while the income inequality is narrower as both dynasties experience the same wage income. Thus, in a second moment the dynamics follow the final path predicted by Kuznets.

Galor and Tsiddon (1996) emphasize that this economic mechanism, associated with a widening and narrowing of the evolution of the wage differential between skilled and unskilled workers, is conformed with the Kuznets hypothesis, and stress that *"Thus, in accordance with the Kuznets hypothesis, during early stages of development output growth is associated with a widening income inequality, whereas in the latter stages output growth is accompanied by a more equal distribution of human capital and income"* (Galor and Tsiddon, 1996 p. S105).

4. The modern approach: non-political and political economy channels

The advanced Kuznets curve and the consequent implications on the causal relationship between economic development and income distribution have triggered a proliferation of studies interested in testing the validity of this U-inverted hypothesis. The Modern literature grows from this theoretical and empirical analyses explosion. The Modern approach, which might be traced back to Galor and Zeira' (1993) and Alesina and Perotti' (1993) contributions, sustains that the adverse effect of income inequality on growth is a key determinant of economic development, in contrary to what the Classical and Neoclassical approaches claim.

This Modern literature may be divided into two broad branches: the non-political and the political economy. The non-political line of research we focus on is based on the "imperfect capital market" transmission channel. The idea is that income distribution affects negatively economic development via inefficient investment level in human capital due to the presence of imperfect capital market, which causes limitations on the option of borrowing for acquiring education, especially for disadvantaged individuals. As representative theories, in what follows, we report the theories advanced by Galor and Zeira (1993) and Galor and Moav (2004).

In the political economy line we concentrate on the fiscal policy channel, and we illustrate the theory of Alesina and Rodrik (1994), according to which, the higher the income inequality the more the adoption of redistributive policies by the government, implying a reduction of *growth-producing assets*.

However, other important lines of research have been advanced. For instance, Murphy, Shleifer, Vishny (1989) suggests the size of domestic demand as a channel through which income inequality lowers economic growth. De la Croix and Doepke (2003) have proposed a theoretical model which attributes the adverse effect of inequality on growth to differences in fertility rates between the rich and the poor. Their idea is that poor parents tend to have a higher number of children and less education for them, which implies that fertility differential affects human capital accumulation because increases the average weight of low-skilled individuals in the

future. Consequently, an increase in income inequality lowers average level of education and, therefore, growth rates in the economy.

4.1 The non-political channel

The work of Galor and Zeira (1993) (GZ) is one of the seminal contributions to the non-political line of research. In particular, GZ sustain that due to capital market imperfections the income distribution affects negatively the economic development via the inefficient investment level in human capital.

GZ develop a two-period model with overlapping generations in an open economy. In the first period, the agents may either choose to invest in human capital or to work. In the second period, the agents who choose to invest in human capital undertake a skilled work; while, the agents who choose to work continue to work as unskilled. It is also assumed that the agents do not differ in preferences and abilities, but they differ only in initial wealth. Due to capital market imperfections with the interest rate in favour to lenders, the initial wealth distribution has an influence in human capital investment level and economic development. The idea is that the presence of imperfect capital market rules out the opportunity to invest in human capital by poor wealth class; then the poor wealth class chooses to work as unskilled, while the rich wealth class prefers to invest in human capital, and to work as skilled. Importantly, assuming the indivisibilities in investment in human capital, this wealth-classes' behaviour persists over time dynasty after dynasty. GZ' model predicts that in the long-run the population converges to two working groups: first, the skilled workers group, which consists of rich people that prefers to invest in human capital dynasty after dynasty; second, the unskilled workers group, which consists of poor people that cannot afford to invest in human capital, and end up working as unskilled generation after generation.

Galor and Zeira (1993) have argued that since the poor are likely to be most affected by credit market imperfections, which limit individuals to borrow against future income, to the extent that income inequality coexist with credit market imperfections poorer groups cannot afford to invest in their physical and human capital with adverse consequences for long run economic performance.

GZ conclude that *"an economy which is initially poor, ends up poor in the long run as well. An economy which is initially rich and its wealth is distributed among many, ends up rich. But an economy with a large initial amount of wealth, which is held by few, ends up poor in the long run"* (Galor and Zeira, 1993 p. 42). Therefore, the initial distribution of wealth has long-term repercussions on investment level in human capital with detrimental consequences for long term growth.

In the same vein, Galor and Moav (2004) (GM) develop a theoretical model exploring the effect of income inequality on long term economic growth in presence of capital market imperfections. According to GM' model, in the early stages of economic development the increase in income inequality is beneficial for growth because it promotes a faster physical capital accumulation; whereas, in the later stages of economic development the decrease in income inequality is beneficial for growth because it enhances the human capital accumulation. GM state that this theoretical model unifies two advanced approaches: the Classical approach, which suggests a positive correlation between inequality and economic growth, and the Credit Market Imperfection approach, which suggests a negative correlation between inequality and economic development.

GM develop an overlapping-generations model with a single good produced using physical and human capital. They sustain that the accumulation of these production factors promotes the economic development. It is assumed that the individuals do not differ in preferences and abilities, but only in their family wealth, which defines the investment level in human capital due to credit constraints. GM also suppose that the individuals live for two periods: during the first period the acquisition of human capital, which increases with increasing in investment in human capital, is obtained; and in the second period the individuals offer their labour unit in the labour market, and decide how allocates the gained income between consumption and bequest. Importantly, the population is classified in two groups: the poor who do not own any capital; and the rich who own initial capital.

In the early stages of economic development, due to low level of physical capital, the rate of return to human capital is lower than the rate of return to physical capital. Therefore, since the propensity to save of the rich people is higher than that of the poor people, the increase in inequality promotes a faster accumulation of physical

capital, and then fosters the economic growth. In this development phase, defined Regime I by GM, the economic development is only driven from the accumulation of physical capital. This accumulation by the wealthier people gradually determines the increase in wage rates.

This slow but steady process of wage rates increase determines that the rate of return to human capital reaches a sufficient level to induce investment in human capital not only by the rich but also the poor. According to GM, the economy enters in a development phase where the economic development is driven from the accumulation of physical capital as well as the accumulation of human capital. GM define this second development phase as Regime II, which is divided in three different stages.

In the first stage of Regime II, due to large credit bindings, the poor cannot afford to devote income to investment, though the increase in rate of return to human capital; thus, they still consume the entire income wage, and thus their offspring does. In contrast, the rich may devote part of their wealth to investment both in human and physical capital. Consequently, the wage rates experience further increases such that the rate of return to human capital becomes sufficiently high to induce the investment in both human and physical capital by the poor.

This change in investment behaviour by the poor marks the beginning of the second stage of Regime II. However, the investment level by the poor remains sub-optimal due to credit constraints. This source of inefficiency might be reduced through the rise in equality, which broadens the investment opportunity in the society. Therefore, in the Regime II, the increase in equality promotes an efficient allocation of investment between physical and human capital, and then fosters the economic growth. Due to this progressive participation of the poor in the aggregate investment process, and the consequent further increase in wage rates of the offspring, GM suggest that the economy enters in a third stage of Regime II where the credit constraint is not limited, and the inequality level does not affect the economic growth.

There are two GM' statements that well summarize these results. First, GM state that *"Inequality enhances the process development in Regime I since a higher concentration of wealth among members of group P (the Poor), would increase*

aggregate consumption, decrease aggregate intergenerational transfers, and thus would slow capital accumulation and the process of development" (Galor and Moav, 2004 p. 17). Also, they write that *"Inequality negatively affects the process development in Stage II of Regime II. A lower concentration of wealth among members of group R and a higher concentration of wealth among member of group P would not affect aggregate consumption, and aggregate intergenerational transfers, but due to liquidity constraints of members of group P would allow for a more efficient allocation of aggregate investment between physical and human capital"* (Galor and Moav, 2004 p. 21). Therefore, according to GM' model, the effect of inequality on economic development differs in regarding on which growth engine dominates: negative effect whether the physical capital accumulation is the dominant factor to enhance economic growth; and positive effect whether the human capital accumulation is the dominant element in this process.

4.2 The political economy channel

In the political economy line of research we focus on the fiscal policy channel, which suggests that an unequal distribution of resources is linked to redistributive pressures harmful to growth (e.g. higher taxation). An important contribution to this line of research is the theory of Alesina and Rodrik (1994) (AR).

These scholars develop a political economy model explaining how the income distribution influences the fiscal policies, and then the economic growth. In few words they expressed their consideration about the link between income distribution and economic growth writing that: *"[ ] distributive struggles harmful to growth are more likely to take place when resources are distributed unevenly"* (Alesina and Rodrik, 1994 p. 467).

AR assume an economy with a single good, which is produced by the adoption of two production factors: the aggregate stocks of capital and the labor factor. The capital factor is considered the only accumulated factor which enhances the economic growth. Moreover, it is assumed that this aggregate production needs the provision of "productive" public services by the government to be operated. Imaging that the government expenditure is financed by a tax on capital income, an important element of this model is, as AR argued, the interplay of redistributive

policies and growth-enhancing policies. In this setting a "low" tax on capital income is considered no-harmful for growth because of the offsetting positive effect of the productive public expenditure. However, a higher level of capital taxation may discourage individual investment causing lower economic growth rates.

Assuming that individuals differ in their capital endowment, the taxpayers whose income consists totally of capital prefer tax rates maximizing the economic growth rates; whereas, any other taxpayer prefers higher tax rates. Due to these differences in taxpayers' preferences on capital tax rates, the more unequal the income distribution in the economy, the greater the redistributive pressures, and then the higher the capital tax rate chosen by the government. This consideration comes from the idea that the government's choice of tax rates depends on the preferences of the majority of voters. AR incorporate this idea in the model by using the median voter theorem.

Given this assumptions, AR suggest that in a more equitable economy, in which the median voter owns a balanced share of capital, the government chooses the capital tax rate maximizing the economic growth. In contrast, in an economy with substantial income inequality, associated with a median voter who does not have access to the productive resource of the economy, namely the capital factor, the government prefers a capital tax rate higher than the economic growth maximizing rate: the less equal the income distribution, the higher the levied capital taxation. Hence, due to the higher redistributive pressures, the income inequality is harmful for economic growth via the distortion effect of taxation on investment decisions. AR' study provides empirical support to this theoretical result on the adverse correlation between income inequality and economic growth. We will come back to this evidence in the next section.

5. Empirical background

Empirical literature has widely concerned with the relationship between income distribution and economic development. In particular, several works have been devoted to addressing the validity of the Kuznets hypothesis, and then to investigating the causality between these two variables. Does economic development define the income inequality level, as Kuznets suggested? Or is income

inequality a key determinant of economic growth? If any, how income inequality affects economic growth, positively or negatively?

In this section we try to review the literature that has focused on these questions empirically. First, the empirical evidence on the validity of the Kuznets hypothesis is reported; and then a second section concerns the empirical literature which has been advanced to investigate the effect of income inequality on economic growth.

5.1 Has the U-inverted curve occurred?

At the core of the Kuznets' story is the "swing" in the long-term income inequality evolution. However, Kuznets himself recognizes that *"No adequate empirical evidence is available for checking this conjecture of a long secular swing in income inequality; nor can the phases be dated precisely"* (Kuznets, 1955, p.19). To cover this lack of *adequate* evidence, supporters of the Kuznets hypothesis have tried to provide empirical evidence of the Kuznets U-inverted relationship between economic development and income inequality.

Historical investigations into the distribution of income in some European and non-European countries tend to support this Kuznets' inequality-swing. First, Kuznets (1955) himself reports historical evidence of the U-inverted shaped pattern of inequality from the United States, England, and Germany. For example, in England he reports that the income inequality widened in the industrial revolution period 1780 to 1850, whereas a narrowing income inequality pattern started with the First World War. Also Williamson (1986) and Lindert (1986) provide historical evidence which supports the Kuznets Hypothesis. Lindert (1986) investigates data evidence from England and Wales between 1670 and approximately 1970 using different data sources. In line with the U-inverted curve, he shows that wealth and income inequality rose between 1750 and 1850, namely across Industrial Revolution era, experienced a phase of no strong variations around 1870 and 1913, and drifted down from 1913 to around 1970.

Ahluwalia (1975) uses econometrics technique to explore the relationship between the income inequality and the economic development. He does so by employing cross-section data, provided by Jain (1975), for 60 countries. The time-series data, though preferred, were not available at the work time, as Ahluwalia suggested. He

investigates the existence of a U-shaped curve by estimating an empirical equation in which the dependent variable, expressed as the income share of five alternative percentile groups, is regressed with respect to the per capita GNP, as measure of the level of development, and the quadratic form of the per capita GNP, included to estimate the turning point predicted by Kuznets. These cross-section regressions provide substantial support for the U-inverted relationship between income distribution and economic development: a worsening income inequality during the early stage of development, and then declining income inequality in the later stages.

Panel data analyses employed by Thornton (2001), for 96 countries, and Iradian (2005), for 82 countries over the period 1965 to 2003, both find an inverted-U relationship between income inequality and economic development, as hypothesized by Kuznets. Bahmani-Oskooee and Gelan (2008) and Shahbaz (2010) come to the same conclusion, while using long time-series data and error correction model (ECM) technique. Bahmani-Oskooee and Gelan (2008) estimate the relationship between income inequality and development in the case of the United States from 1957 to 2002, and Shahbaz (2010) in the case of Pakistan from 1971 to 2002.

Therefore, many studies that attempt to test the validity of the Kuznets' hypothesis end up with proving empirically what the Kuznets' law suggests (see also e.g. Papanek and Kyn, 1986; Jha, 1996; Lin, Huang, and Weng (2006); although the existence of the U-inverted curve ends up with been rejected in just as many. In particular, it appeared that the Kuznets bell-shaped curve tends to disappear when the availability of better data which allow controlling for country heterogeneity and the inclusion of other explanatory variables such as education occurred (Bourguignon and Morrison, 1990; Bourguignon, 2004).

In this respect, Deininger and Squire' (1996) study providing the compilation of a new data set will enable the research to make a qualitative leap. They developed an unbalanced panel with approximately 10-year observations for each country. As they argued, this work produced a "*better and more comprehensive data"*. Especially, they suggest that this data set allowed them to shed more convincing and robust light on the relationship between income distribution and economic development. Deininger and Squire' (1996) results are clearly remarked in a graphical

representation developed by Bourguignon (2004). This illustration is reported in Figure 4 (Bourguignon, 2004). As this Figure immediately illustrates, they find that the U-inverted curve is verified when pure cross-section data are used, the U-bell shape becomes less pronounced when time changes are estimated, and finally the Kuznets curve tends to disappear when country fixed effects are included.

[FIGURE 4 ABOUT HERE]

While studying the stylized facts of poverty, inequality, and growth for developing countries, the validity of the Kuznets hypothesis have been rejected by Field (1989) who uses micro survey time series data. He investigates the effect of development process on inequality by comparing this relationship both in low-income and in high-income countries. The conclusion he reaches is that *"In the data considered here, inequality increased with growth as frequently in low-income countries as in high-income countries. There appears to be no tendency for inequality to increase more in the early stages of economic development than in the later stages"* (Fields, 1989, p. 11). Thus, Fields' work provides no-evidence relative to Kuznets hypothesis.

Anand and Kanbur (1993) assess the Kuznets law by testing the validity of a crucial property associated with the *inequality-development process* predicted by Kuznets. In particular, they examine the Kuznets' statement of a shifting from rural sectors, characterized by low mean income and low income inequality, to urban sectors, characterized by high mean income and high income inequality - Kuznets himself considered this shifting in his numerical example. In the spirit of this property, Anand and Kanbur (1993) propose a formalization of the inequality-development process, and define the functional form and the condition for the turning point for six different inequality measures. For instance, they use the Theil's entropy index developed by Theil (1967), the Atkinson (1970) inequality index, and the Gini index. They estimate each functional form and condition for turning point for 60 countries by using data provided by Jain (1975). Importantly, out of the six inequality indices, the condition for the turning point is rejected for five indices. Moreover, it is worth noting that, according to the inequality index used as the dependent variable, different shapes and turning points are illustrated. Therefore, it seems that the

definition of the shape and the turning point of the inequality-growth relationship are susceptible to the choice of the index. Besides, Ravallion (1995) re-uses the Anand and Kanbur' (1993) "formalization" of the inequality-development process, and finds no significant evidence of the U-inverted curve. He underlines that this result displaying that growth does not influence inequality, neither negatively nor positively, is in line with Fields' (1989) empirical conclusions.

Exploiting panel data dimension Matyas, Konya, and Macquarie (1998) apply fixed and random effects technique for 47 and 62 countries over the period 1970 to 1993, and find no evidence to support the U-inverted curve postulated by Kuznets. Similarly, Angels (2010) estimates fixed effects regressions for panel of 226 countries over the period 1960 – 2005, and observes that his empirical results are not conformed to the Kuznets curve. Yet, while using panel data set of 29 regions of China for the period 1987 – 2001, Wan, Lu, and Chen (2006) reject the Kuznets hypothesis.

Garth Frazer (2006) assesses the relationship between income distribution and economic development applying both cross-country and within-country nonparametric econometric techniques. As in previous studies, he finds empirical evidence of the Kuznets curve when cross-country nonparametric method is used, but this evidence is not confirmed when within-country nonparametric technique is applied.

According to Bourguignon (2004), the empirical evidence which tends to not support the Kuznets hypothesis does not mean that economic growth does not affect income distribution. This indicates that the evidence of a "universal" cross-country Kuznets curve may probably fail because of the relevance of country specificity in the inequality-growth pattern; therefore, the Kuznets hypothesis might be verified for singular countries. Bourguignon (2004) reports the Brazilian case as example. Brazil shows an unchanged income inequality in front of slow economic growth from 1976 to 1996. It would appear that growth did not affect inequality. However, socio-demographic factors which imply lower inequality occurred. In particular, reduction in fertility rate and improvement in education achievement happened. Therefore, it could mean that economic growth has caused an increase in income inequality, which has offset the socio-demographic counteracting forces (Bourguignon, 2004).

Bourguignon (2004) emphasizes that the Brazilian case would mean that there exists room for redistributive policies to deal with the adverse inequality effect of growth.

5.2 Does income inequality affect long term economic growth?

The investigation into the effect of economic development on income distribution is just one side of the story. The reverse causality, namely the effect of income inequality on long term economic growth, is also extensively examined.

From an empirical standpoint, the dominant viewpoint is that the impact of income inequality on growth is negative due to both economic and political channels. A primary study by Alesina and Perotti (1993) on a cross-section of 70 countries for the period 1960 – 1985 finds that income inequality increases political instability and, in turn, political instability reduces investment. In particular, in a two-equation model which would address endogeneity issues the investment variable is regressed with respect to socio-political instability (SPI), as measured by a composite index assembled by Alesina and Perotti (1993), and then the SPI index is regressed with respect to investment and income inequality, as measured by the income share of the middle class. The results demonstrate that political instability reduces investment, and importantly that a *wealthy* middle class enhances investment, and consequently growth, due to the guarantee of political stability.

In the same vein, Perotti (1996) explores the correlation between income inequality and economic growth but especially investigates the transmission channel in action. For this reason, he first estimates the reduced form of the model, and then examines the validity of each single channel he considers using cross-section data of 67 countries for the period 1960 – 1985. Specifically, he takes into account two political channels (the fiscal policy and the socio-political instability channels) and two economic channels (the imperfect capital markets and the endogenous fertility channels). The results are supportive of both the socio-political instability channel and the endogenous fertility channel: the former, as already explained, suggests that inequality is conductive to social unrest that, in turn, reduces investment, and then economic growth; whereas, the latter predicts that less unequal societies have lower fertility rate and higher investment rate in human capital, and consequently higher economic growth rate.

Some other studies only focus on examining the reduced form of the model. For instance, Alesina and Rodrik (1994) assess whether income inequality affects negatively growth due to the adoption of redistributive policies. Specifically, they assume that more unequal countries are subject to higher redistributive policies, which are defined by Alesina and Rodrik (1994) as *"growth-retarding policies"* because of efficiency distortions. They develop a theoretical model that shapes this hypothesis, and then they test its validity on 54 countries for the period 1960 – 1985 by using a two-stage least square (TSLS) regression method. The results, which show a negative correlation between income inequality and growth, are supportive of their hypothesis.

Similarly, Persson and Tabellini (1994) develop a theory establishing that income inequality has a negative effect on growth due to redistributive pressures that act as reducing-investment policies. In order to test this hypothesis they use two data sets. First, they exploit historical panel data on 9 developed countries for the period 1830 – 1985, and apply both ordinary least square (OLS) and TSLS regression estimations. Second, they assess this hypothesis on cross-section of 56 developed and developing countries for the period 1960 – 1985 by using only OLS estimator. The empirical findings suggest that the income inequality is harmful for growth as predicted by the theory. However, they recognize that the empirical analysis only examines the reduced form but does not give clarifications about the specific channel of transmission.

From an historical data analysis, Berg and Ostry (2011) find a negative correlation between length of growth spell and income distribution. They underline that countries with more equal income distribution, like emerging Asian countries, have growth spells longer than countries with more unequal income distribution, like Latin American countries. This stylized fact, they argue, might explain the Asian *miracle*. However, they recognizes that it cannot be deducted any causal interpretation from their analysis, but it must be seen as a tentative to deduct some stylized facts from the data.

One of the main criticisms of the inequality-growth significant negative coefficient regards the application of cross-country estimators, which might be biased due to omitted variables. In particular, these regressions fail to control for preferences,

technological progress, institutions, and any other unobservable variable causing growth. When unobservable are constant over time, the use of panel dimension allowed controlling for omitted variable biases by including a country-specific effect. The use of the time dimension started to be possible thank to the availability of a new high-quality and expanded database compiled by Deininger and Squire (1996). Investigating the only reduced form of the model some scholars find that that inequality effect varies according to the country income level. For instance, Deininger and Squire (1998) find that lower land inequality is beneficial for growth in poor countries but not in rich countries. They argue that this evidence seems to be consistent with imperfect capital market theory. Moreover, Barro (2000) examines the causal relationship between income inequality and economic growth using a panel data set of 100 countries over the period 1960 – 1995, and finds that inequality affects negatively growth in poor countries but positively in rich countries. Similarly to Deininger and Squire (1998), he interprets this result as supporting evidence of the imperfect capital market theory. According to Barro (2000), since credit-market constraints are more stringent in low-income countries than in rich-income countries, the negative effect of inequality on growth prevails in the former but not in the latter, where the inequality *growth-promoting* effect dominates.

Knowles (2005) criticizes previous studies exploring inequality-growth relationship because of the use of no-comparable data. He sustains these analyses explore inequality-growth relationship by employing inequality measures based on gross income data for some countries and disposable income for other ones, and also using household and individual level data contemporaneously. While testing the validity of the negative coefficient claimed by most of the existing empirical literature, he estimates cross-country regressions, due to insufficient comparable data to conduct panel data analysis, for 30 countries during the period 1960 to 1990, and finds that the coefficient is not significant when only gross individual income distribution data are used, whereas a significant negative correlation exists when expenditure data are used as measure for income inequality.

While employing a household survey including 100 rural villages from China for the period 1986 to 1999, Benjamin, Brandt, and Giles (2006) find a negative correlation between initial inequality and subsequent household income growth. However, once

controlling for fixed-village-effects this significant relationship disappears. They suggest that the inclusion of village fixed-effects might capture only short-run aspects, and conclude that *"it is possible that the short-run links between inequality and growth are weak, and that whatever process drives growth, operates at longer run frequencies in the data"* (Benjamin et al., 2006, p. 28).

Other scholars note that the negative evidence is often accused of being unbelievable because results might be biased due to endogeneity. While facing with this issue and trying shedding some new light on this framework, the relationship between income distribution and economic growth has been investigated by using system generalized method of moments (sys-GMM) estimation technique. For instance, Castellò-Climent (2004) who uses a new inequality indicator that measures human capital inequality finds that inequality is harmful for growth by estimating sys-GMM regressions for 55 countries during the period 1965 to 2000. While using sys-GMM estimator for data sample presented in Solt (2009), Ostry, Berg, and Tsangarides (2014) discover that a reduction of disposable income inequality causes faster and more durable growth. Moreover, Cingano (2014) exploits time dimension, and estimates the relationship between income distribution and economic growth using a sys-GMM method of a panel data set covering 31 OECD countries over the period 1970 – 2010. The use of only OECD countries, he states, would avoid that the findings are biased by the differences in the level of economic development. Importantly, Cingano shows that the negative association between income distribution and economic growth still appears after controlling for these issues. In his study, performing both a macro and a micro level analysis, he finds strong support for the capital market imperfection theory proposed by Galor and Zeira (1993), according to which borrowing constraints limit the acquirement of education especially by *disadvantaged* individuals. Thus, he suggests that more income inequality is likely to lower growth due to underinvestment in human capital by "*low socio-economic groups"*.

Nevertheless, empirical evidence in defense for a positive correlation has also been provided. Forbes (2000) sustains that the empirical works establishing a negative correlation between inequality and growth are not robust. In particular, he states that these empirical studies suffer from measurement errors in inequality and

omitted variable biases. Especially, he stresses that these cross-country analysis show that country with lower inequality level growth faster but they do not show how the change in inequality within a country is associated with the internal-country growth rate. While addressing these issues through the use of panel data estimation and a more coherent data for income inequality measure, precisely the Deininger and Squire (1996) database, Forbes (2000) finds that an increase in income inequality is beneficial for subsequent economic growth. These results have been found by using a generalized method of moments (GMM) approach, developed by Arellano and Bond (1991), for 45 countries over the period 1966 – 1995. Similarly, while exploiting the Deininger and Squire (1996) database, Li and Zou (1998) examine the relationship between income distribution and economic growth using a panel data set of 46 developed and developing countries, and find a positive correlation between these two variables.

Same evidence is provided by Partridge' (2005) analysis which is performed by using state level data from the United States for the period 1960 to 2000. He shows a significant positive coefficient across alternative estimators (OLS, random effects, and between effects); however, this result is not confirmed by fixed effects regressions. Partridge (2005) suggests that fixed effects estimations capture only short-run effects, but inequality-growth relationship operates in the long-run, therefore, estimators such as both OLS and random effects are more consistent because they retain the cross-sectional variation. Moreover, Frank (2009) still employs annual data from 48 U.S. states for the period 1945 to 2004 when estimates three alternative dynamic panel error correction models which allow separating short-run and long-run effects: the fixed effects estimator, the mean group estimator of Pesaran and Smith (1995), and the pooled mean group estimator of Pesaran, Shin, and Smith (1999). From this analysis he concludes that the long-run relationship between inequality and economic growth is positive.

Voichovsky (2005) investigates the correlation between inequality and growth differently from presented empirical literature. She explores whether this relationship varies according to inequality in different parts of the income distribution. Namely, she examines whether the effect of inequality on growth has the same sign and magnitude regardless whether the inequality is at the top end of

the distribution or at the lower bound of the distribution. When applying a sys-GMM estimator to an unbalanced panel of 25 countries for the period 1975 to 2000, Voichovsky (2005) finds that inequality at the top end of the distribution causes more growth, whereas inequality at the lower bound of the income distribution is negatively associated with economic growth.

6. Conclusions

In this chapter, we have provided an overview of the literature regarding the relationship between income distribution and economic development. On one hand, we have seen the central work of Kuznets with his U-inverted hypothesis. On the other, much empirical work has been done to examine the impact of income inequality on long run economic growth. Galor and Zeira' (1993) and Alesina and Perotti' (1993) studies seemed to have shed some light on this topic, emphasizing the existence of a negative effect of income inequality on economic growth.

In his seminal work, Bénabou (1996) developed a wide body of theoretical models establishing a negative association between income distribution and growth, and summarized the main findings of 23 studies that deal with this relationship. The author concluded that these studies find a negative correlation between income inequality and economic growth, independently on data sets, time-periods, or income distribution measures used.

However, from an empirical point of view, our survey has reported mixed results. First, existing evidence reports controversial results about the existence of the Kuznets curve. On one hand, scholars provide evidence in favour of the Kuznets hypothesis. Conversely, other scholars assert that the U-inverted curve hypothesized by Kuznets tends to disappear when longitudinal data rather than cross-section data are used and other explanatory variables (such as education) are taken into consideration in the analysis (see e.g. Bourguignon and Morrison, 1990; Deininger and Squire, 1998; Bourguignon, 2004).

Second, empirical studies differ in finding positive, negative or not significant effects of income inequality on economic growth. It seems that studies' conclusions depend on the econometric technique performed, and the sample used. For instance, they vary due to sample coverage: a negative effect is more likely to emerge when using a

sample which includes developing countries rather than developed countries. Some studies have shown that the link between income inequality and economic growth is negative and significant in poor countries, but either insignificant (Deininger and Squire, 1998) or even positive (Barro, 2000) in rich countries. Moreover, they vary due to their estimation strategies: a negative effect usually emerges when referring to reduced form estimates performed using cross-sectional data, but a positive link is often found when within-country variation is investigated (e.g. Forbes, 2000; Li and Zou, 1998).

In particular, the literature (see e.g. Atkinson and Bourguignon, 2015) sustains that the lack of unanimous consensus is caused by the data quality, which is the main source of biased estimations in this field of studies. On this point, the Deininger and Squire' (1996) work demonstrated to be a turning point because it provided a *better and more comprehensive* database of income inequality measure. However, this database still maintained a crucial drawback concerning the *"differences in the method of measurement"* (Barro, 2000 p.17). Differences in measure unity, using both household and individual data, and income definition, using both gross/net income and expenditure data, are still present in Deininger and Squire Database, and are source of measurement errors (Barro, 2000; Knowles, 2005).

The absence of a database which provides an accurate measure of income inequality is a crucial limitation in the studies discussed in this chapter, and is still a challenge for economists.

References

Aghion P., Bolton P. (1997) A Theory of Trickle-Down Growth and Development. *The Review of Economic Studies*, 64:151-172.

Ahluwalia M.S. (1976) Inequality, Poverty and Development. *Journal of Development Economics*, 6: 307-342.

Alesina A., Perotti R. (1993) The Political Economy of Growth: A Critical Survey of the Recent Literature. *The World Bank Economic Review*, 8: 351-371.

Alesina A., Perotti R. (1996) Income Distribution, Political Instability, and Investment. *European Economic Review*, 40: 1203-1228.

Alesina A., Rodrik D. (1994) Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109: 465-490.

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

Anand S., Kanbur S.M.R. (1993) The Kuznets process and the inequality-development relationship. *Journal of Development Economics*, 40: 25-52.

Angeles L. (2010) An alternative test of Kuznets' hypothesis. *The Journal of Economic Inequality*, 8: 463-473.

Arellano M., Bond S. (1991) Some tests of specification for panel data: monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58: 277-297.

Atkinson A.B. (1970) On the measurement of inequality. *Journal of Economic Theory*, 2: 244-263.

Atkinson A.B. (1997) Bringing income distribution in from the cold. *The Economic Journal*, 107: 297-321.

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

Atkinson A.B., Bourguignon F. (2015) *Handbook of Income Distribution*. North-Holland, 2A ed., Elsevier.

Bahmani-Oskooee M., Gelan A. (2008) Kuznets inverted-U hypothesis revisited: a time-series approach using US data. *Applied Economics Letters*, 15: 677-681.

Banerjee A.V., Newman A.F. (1993) Occupational Choice and the Process of Development. *Journal of Political Economy*, 101: 274-298.

Barro R. (2000) Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5: 5-32.

Bénabou R. (1996) Inequality and Growth. NBER Chapters, in: *Macroeconomics Annual*, 11: 11-74.

Bénabou R. (2000) Unequal Societies: Income Distribution and the Social Contract. *The American Economic Review*, 90: 96-129.

Benhabib J., Rustichini A. (1996) Social Conflict and Growth. *Journal of Economic Growth*, 1: 125-142.

Benjamin D., Brandt L., Giles J. (2006) Inequality and Growth in Rural China: Does Higher Inequality Impede Growth. IZA discussion paper 2344.

Berg A., Ostry J.D. (2011) Inequality and Unsustainable Growth: Two Sides of the Same Coin?. IMF Staff Discussion Note SDN/11/08.

Bertola G. (1993) Factor Shares and Savings in Endogenous Growth. *American Economic Review*, 83: 1184-98.

Bourguignon F. (1990) Income distribution, development and foreign trade: a cross-sectional analysis. *European Economic Review*, 34: 1113-1132.

Bourguignon F. (2004) *The Poverty-Growth-Inequality Triangle*. Washington D.C.: The World Bank.

Castellò-Climent A. (2004) A Reassessment Of The Relationship Between Inequality And Growth: What Human Capital Inequality Data Say?. IVIE Working Paper no. 15.

Cingano F. (2014) Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment and Migration Working Papers No 163, OECD Publishing, Organization for Economic Co-operation and Development, Paris.

De la Croix D., Doepke M. (2003) Inequality and Growth: Why Differential Fertility Matters. *American Economic Review*, 93: 1091-1113.

Deininger K., Squire L. (1996) A New Data Set Measuring Income Inequality. *The World Bank Economic Review*, 10: 565-591.

Deininger K., Squire L. (1998) New ways of looking at old issues: inequality and growth. *Journal of Development Economics*, 57: 259-287.

Fields G.S. (1989) Changes in Poverty and Inequality in developing Countries. *The World Bank Research Observer*, 4: 167-185.

Fields G.S. (2001) *Distribution and Development: a New Look at the Developing World*. MIT Press.

Forbes K.J. (2000) A reassessment of the relationship between inequality and growth. *American Economic Review*, 90: 869-887.

Frank M.W. (2009) Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures. *Economic Inquiry*, 47: 55-68.

Frazer G. (2006) Inequality and Development Across and Within Countries. *World Development*, 34: 1459-1481.

Galor O. (2011) Inequality, Human Capital Formation and the Process of Development. NBER Working Paper No 17058.

Galor O., Moav O. (2004) From Physical to Human Capital Accumulation: Inequality and the Process of Development. *Review of Economic Studies*, 71: 1001-1026.

Galor O. Tsiddon D. (1996) Income Distribution and Growth: The Kuznets Hypothesis Revisited. *Economica*, 63: S103-S117.

Galor O., Zeira J. (1993) Income Distribution and Macroeconomics. *Review of Economic Studies*, 60: 35-52.

Jain S. (1975) *Size Distribution of Income: A Compilation of Data*. Washington D.C.: World Bank.

Jha S.K. (1996) The Kuznets Curve: A Reassessment. *World Development*, 24: 773-779.

Kaldor N. (1956) Alternative Theories of Distribution. *The Review of Economic Studies*, 23: 83-100.

Knowles S. (2005) Inequality and economic growth: the empirical relationship reconsidered in the light of comparable data. *The Journal of Development Studies*, 41: 135-159.

Kuznets S. (1955) Economic Growth and Income Inequality. *American Economic Review*, 45: 1-28.

Lewis W.A. (1954) Economic Development with Unlimited Supplies of Labour. *The Manchester School of Economic and Social Studies*, 22: 139-191.

Li H., Zou H. (1998) Income Inequality is not Harmful for Growth: Theory and Evidence. *Review of Development Economics*, 2: 318-334.

Lin S.C., Huang H.C., Weng H.W. (2006) A semi-parametric partially linear investigation of the Kuznets' hypothesis. *Journal of Comparative Economics*, 34: 634-647.

Lindert P.H. (1986) Unequal English Wealth since 1670. *The Journal of Political Economy*, 94: 1127-1162.

Matyas L., Konya L., Macquarie L. (1998) The Kuznets U-curve hypothesis: some panel data evidence. *Applied Economics Letters*, 5: 693-647.

Murphy K.M., Shleifer A., Vishny R. (1989) Income Distribution, Market Size, and Industrialization. *Quarterly Journal of Economics*, 104: 537-564.

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

Ostry M.J.D., Berg M.A., Tsangarides M.C.G. (2014) Redistribution, inequality, and growth. IMF Staff Discussion Note SDN/14/02.

Papanek G., Kyn O. (1986) The Effect on Income Distribution of Development, the Growth Rate, and Economic Strategy. *Journal of Development Economics*, 23: 55-65.

Partridge M.D. (2005) Does income distribution affect US state economic growth?. *Journal of Regional Science*, 45: 363-394.

Pasinetti L.L. (1962) Rate of Profit and Income Distribution in relation to the Rate of Economic Growth. *The Review of Economic Studies*, 29: 267-279.

Perotti R. (1993) Political Equilibrium, Income Distribution at Different Levels of Development: A Survey of Evidence. *Review of Economic Studies*, 60: 755-776.

Persson T., Tabellini G. (1994) Is Inequality Harmful for Growth? *American Economic Association*, 84: 600-621.

Pesaran M.H., Smith R.P. (1995) Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics*, 68: 79-113.

Pesaran M.H., Shin Y., Smith R.P. (1999) Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association*, 94: 621-624.

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

Piketty T., Saez E. (2003) Income Inequality in the United States, 1913-1998. *Quarterly Journal of Economics*, 14: 221-250.

Ravallion M. (1995) Growth and poverty: Evidence for developing countries in the 1980s. *Economic Letters*, 48: 411-417.

Rousseau J.J. (1775) *The Social Contract, A Discourse on the Origin of Inequality, and A Discourse on Political Economy*. Cole G.D.H., J.M. Dent and Sons.

Saez E. (2015) Striking it richer: the evolution of top incomes in the United States. Working Paper UC Berkeley.

Shahbaz M. (2010) Income inequality-economic growth and non-linearity: a case of Pakistan. *International Journal of Social Economics*, 37: 613-636.

Solt F. (2009) Standardizing the World Income Inequality Database. *Social Science Quarterly*, 90: 231-242.

Stiglitz J. (2012) *The price of Inequality: How Today's Divided Society Endangers Our Future*. W.W. Norton & Company.

Theil H. (1967) *Economics and Information Theory*. North-Holland: Amsterdam.

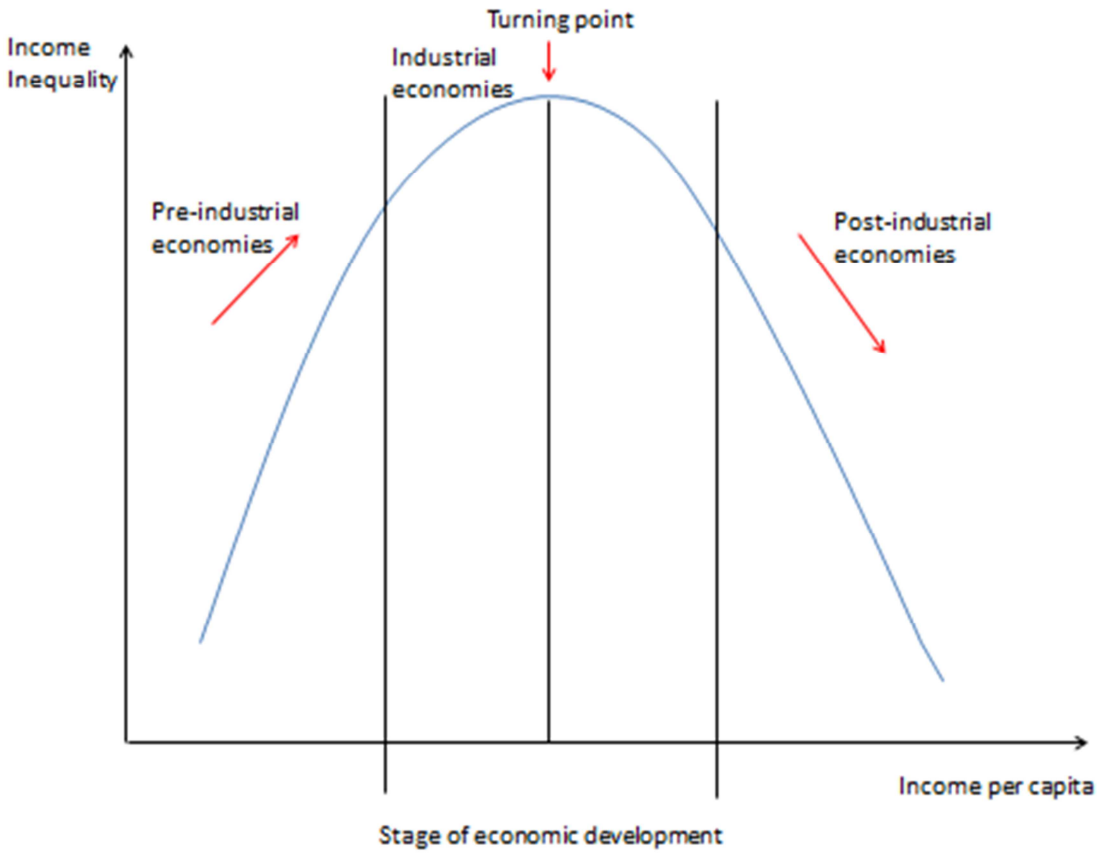
Thornton J. (2001) The Kuznets inverted-U hypothesis: panel data evidence from 96 countries. *Applied Economics Letters*, 8: 15-16.

Voitchovsky S. (2005) Does the profile of income inequality matter for economic growth?. *Journal if Economic Growth*, 10: 273-296.

Wan G., Lu M., Chen Z. (2006) The inequality-growth nexus in the short and long-run: Empirical evidence from China. *Journal of Comparative Economics*, 34: 654-667.

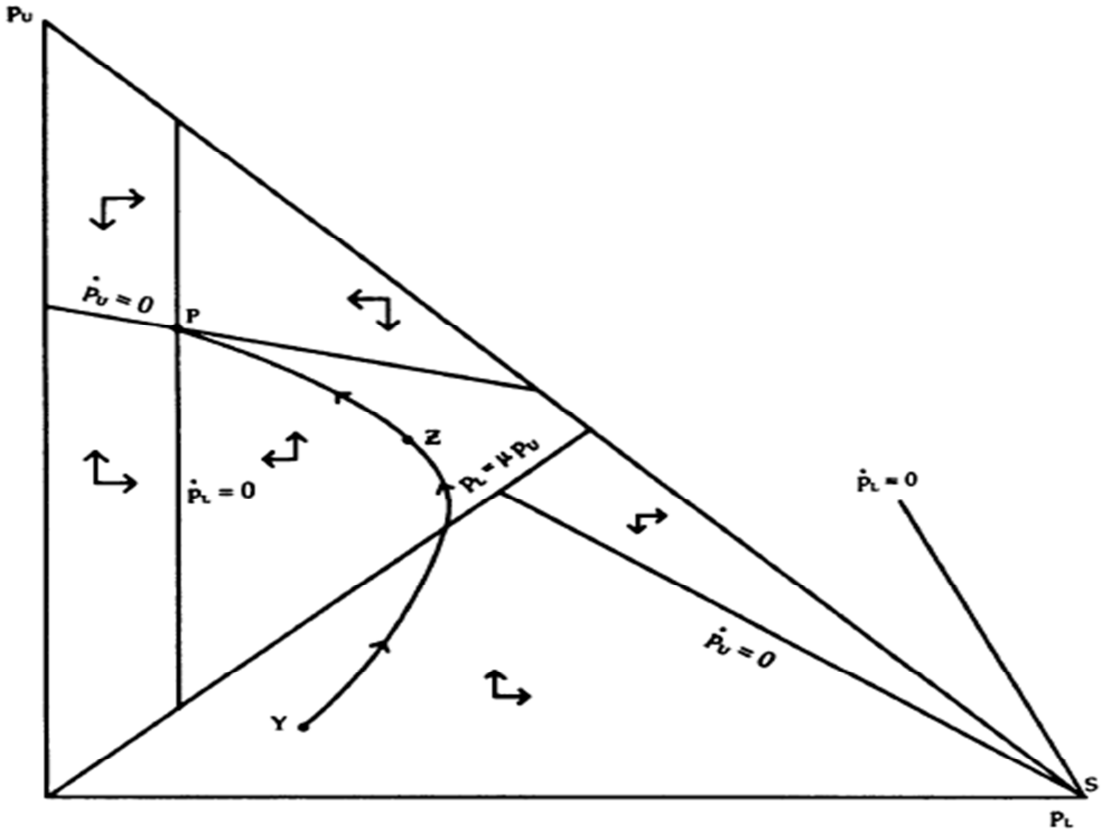
Williamson J.G. (1986) *Did British Capitalism Breed Inequality?*. Boston: Allen and Unwin.

Figure 1: Kuznets curve



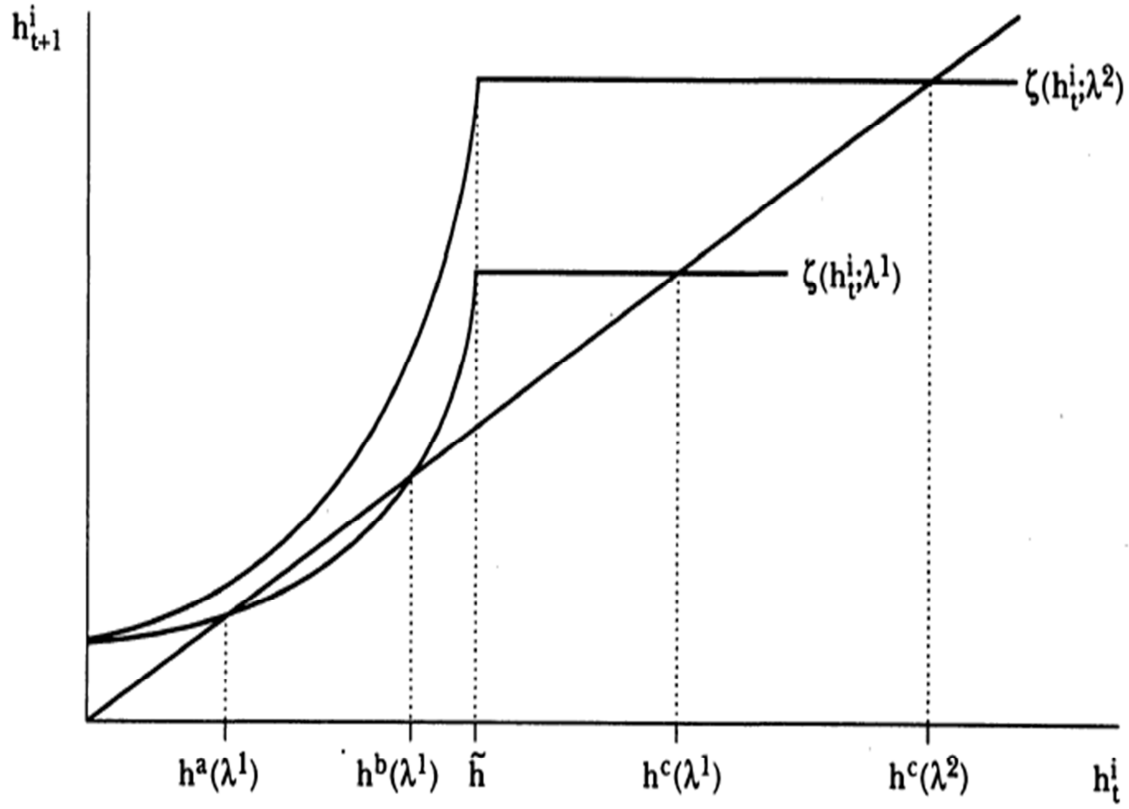
Source: our elaboration.

Figure 2: Kuznets curve development path



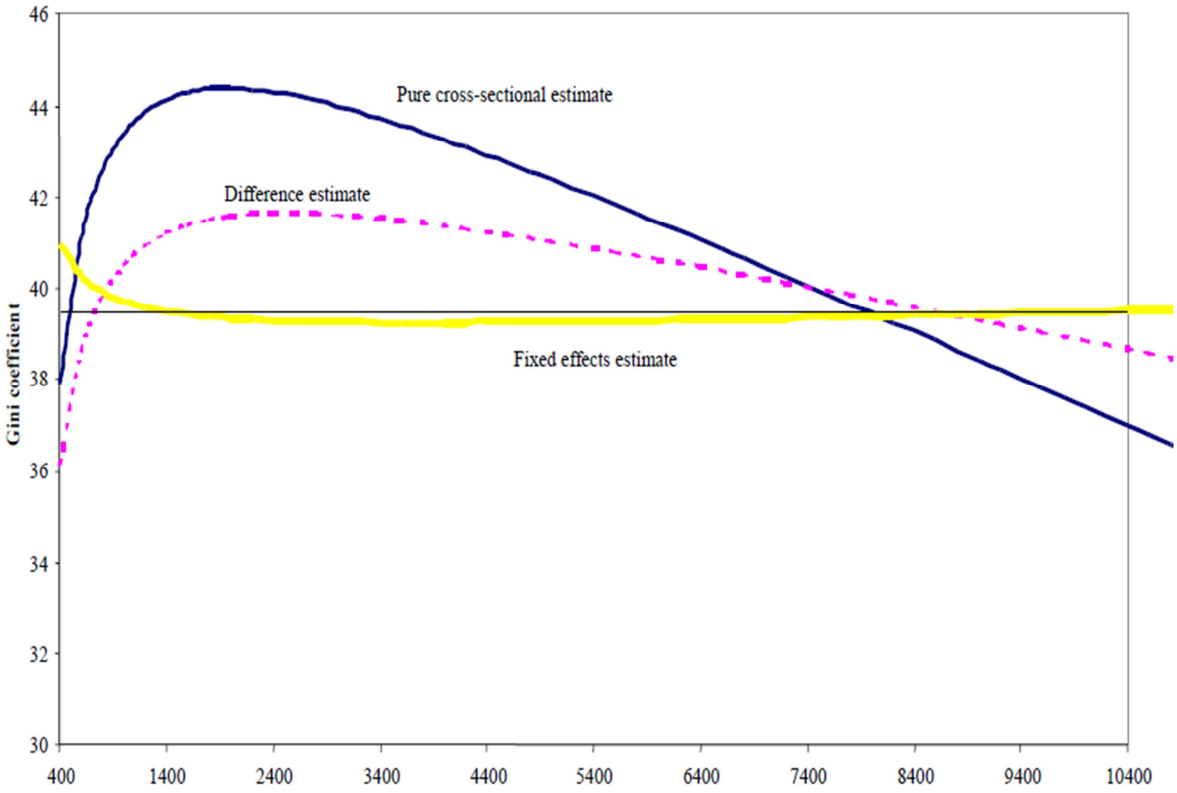
Source: Banerjee and Newman (1993, p.296).

Figure 3: System dynamical evolution



Source: Galor and Tsiddon (1996, p.S110).

Figure 4: Kuznets curve estimates



Source: Bourguignon (2004, p.15).

Chapter 2

Marginal tax rate and economic growth: accounting for the quality of institutions

1. Introduction

In the last decades, the impact of marginal tax rates on economic growth has received an enduring interest. On one hand, taxation plays a role in the recent European policy debate on the desirability of introducing growth-oriented tax reforms (Berg and Ostry, 2011; Cingano, 2014). The recent economic crisis has, undoubtedly, renewed interest in policies fostering macroeconomic performance and promoting a long-term growth and, for these purposes, in how tax system should be so that these objectives are best pursued. On the other hand, the economic literature reveals an interest in top marginal taxation because of the increase, in many OECD countries, of income inequality, of which tax regimes could be one of the causes (Alvaredo, Atkinson, Piketty, and Saez, 2013; Piketty, 2012; Piketty, Saez, and Stantcheva, 2014; Atkinson, 2015). A decrease in marginal income taxation is considered no-welfare improving because its impact on income inequality – cuts of top marginal tax rate are indicated as a driving force of the recent surge in upper incomes (Alvaredo et al., 2013; Piketty et al., 2014)[2] - largely outweighs increases in economic efficiency, whereas hardly marginal taxation rates can be considered relevant for economic growth.

[2] Recent upper income distribution has been explained in a number of ways. On one hand, economic factors have been indicated as driving forces of rising in inequality. This category would include explanations based on skill-biased technological change (SBTC) [Autor, Katz, and Krueger (1998), Autor and Katz (1999), Autor, Levy, and Murnane (2003), and Autor, Katz, and Krueger (2006)]. However, it is suggested that SBTC hypothesis might only narrowly explain income inequality distribution [Gordon and Becker (2005), and Piketty et al. (2014)]. International trade is also defined as an explanation of rising in income inequality due to labor demand shift favoring high-skilled workers [Borjas and Ramey (1995), and Feenstra and Hanson (1999)]. An alternative hypothesis of the increase in top percentile income share, related in some way to SBTC category, is based on the phenomenon of "superstar" proposed by Rosen (1981). Superstar is a term indicating person who has extraordinary ability in some field and is widely popular for it. Rosen has suggested that compensation for "superstars" grows more and more over time than compensation for others [Sherwin Rosen (1981), Gordon and Becker (2005)]. Besides, it is suggested that tax reforms involving an income shifting from corporate to personal tax base could explain the recent income distribution [Slemrod (1996) and Gordon and Slemrod (2000), Bakija (2013)]. Finally, it is suggested that upper incomes concentration has been driven by top earners' rent extraction behaviors favored by both social norms and institutional changes [Piketty and Saez (2007), Bakija, Cole, and Heim (2012), Bakija (2013), and Bell and Van Reenen (2014)] and drops in top marginal tax rates [Piketty et al. (2014)].

Traditionally, marginal tax rate cuts have been seen to be growth-oriented on the grounds that these tax reforms yield efficiency gains in terms of a higher labour supply by top earners. The first part of this argument, i.e. tax cuts have a positive effect on individual top earners labour supply, has been empirically validated by Lindsey (1987), Feenberg and Poterba (1993) and Feldstein (1995) who show that a cut of the marginal tax rate encourages top earners, who benefit of tax advantage, to make a greater productive effort.

However, the validity of the second part of the argument is challenged by the so-called Harberger's superneutrality conjecture – i.e. different tax structures impact on economic efficiency but not on macroeconomic long-run growth. In this case the evidence is mixed. The conjecture seems to be confirmed in a few cases. For instance Easterly and Rebelo (1993) who, once controlling for initial GDP per capita levels, do not find any significant effect related to income taxes; Mendoza, Milesi-Ferretti, and Asea (1997) who show that changes in so-called MRT tax rates, i.e. a measure of effective tax rates constructed by Mendoza, Razin, and Tesar (1994), do not affect long-term economic growth; Agell, Ohlsson, and Thoursie (2006) and Lee and Gordon (2005) who find that only corporate taxes seem to affect economic growth, but not personal income taxes. On the same vein, recently, no long-run effects have been found by Piketty et al. (2014), as far as top statutory income rates are concerned.

By contrast, negative effects from higher tax rates have been found by Kneller, Bleaney, and Gemmell (1999) and Bleaney, Gemmell, and Kneller (2001) for a set comprising taxation on income and profit, taxation on property, and no effects for taxation on domestic goods and services. Mertens and Ravn (2013), who estimate a structural vector autoregressions (SVAR) by using data on tax-policy changes classified exogenous by Romer and Romer (2010), show that an increase in average personal income tax rates affect negatively real Gdp per capita. While using the Romer and Romer (2010) tax-policy measure as an instrument, Barro and Redlick (2011) still find an inverse correlation between average marginal income tax rates and Gdp per capita.

Recent refinements of the empirical techniques adopted have not solved the puzzle. On one hand, Arnold (2008) shows that marginal tax rates are harmful for growth,

once applying pooled mean group (PMG) methodology proposed by Pesaran, Shin, and Smith (1999); Arachi, Bucci, and Casarico (2015) who, instead, apply the common correlated effects (CCE) estimator proposed by Pesaran (2006), do not find any significant effect when using implicit tax rates on labour.

Given that most studies make use of aggregate measures of the average tax burden (e.g. the tax revenue-GDP ratio), the previous studies may miss the focus on the size of the distortionary effect, i.e. the degree of diversion from optimal economic decisions, which is likely to depend on marginal tax rates. Following recent studies (e.g. Lee and Gordon, 2005; Piketty et al., 2014; Gemmell, Kneller, and Sanz, 2014), we use the top statutory personal income tax rates. We agree that the focus on top statutory rates is important when treating tax systems as progressive, and direct incentive effects and labour supply decisions are concerned. In examining the impact of taxation on Gdp growth, it is recognized that the choice of tax rates depends on the behavioural decision investigated: top personal statutory rates are likely to reflect the impact of increased progressivity, and capture the marginal rates which are more relevant for labour supply decisions of higher income earners; while effective and average tax rates, which are usually based on tax revenue data, are more likely to reflect investment decisions as capturing the level effects (Myles, 2007; Devereux, Lockwood, and Redoano, 2008; Gemmell et al., 2014; Kneller and Misch, 2017). To this end, Mertens (2015) finds no evidence of a direct effect of average tax rates on icome response, once controlling for marginal rates; also, Barro and Redlick (2011), who use top statutory personal tax rates, affirm that the effect on Gdp per capita mainly derives from marginal rather than average tax rates.

In this field of study, one aspect that has not been deeply investigated is whether the quality of institutions influences in some way the impact of marginal taxation on long-term growth. Douglass North (1991), in primis, has emphasized that institutions are primary engine of long-term economic growth. He asserts that

"Institutions are the humanly devised constraints that structure political, economic and social interaction. They consist of both informal constraints (sanctions, taboos, customs, traditions, and codes of conduct), and formal rules (constitutions, laws, property rights). Throughout history, institutions have been devised by human beings to create order and reduce uncertainty in exchange. Together with the standard

constraints of economics they define the choice set and therefore determine transaction and production costs and hence the profitability and feasibility of engaging in economic activity" (North, 1991, p.3).

The idea is that good economic institutions represent an inter-related *cluster of informal and formal rules* enhancing the prosperity of countries due to its effect on the setting of transaction and production costs. For example, since the protection of property rights ensures individuals who are generally risk adverse, their enforcement creates an incentive to invest, innovate, and participate to economic activity; hence, the security of property rights affects positively growth (Acemoglu, Johnson, and Robinson, 2001).[3]

Considering the bulk of empirical evidence proving the importance of institutions for economic (Knack and Keefer, 1995; Barro, 1991, 1997, 2013; Hall and Jones, 1999; Mahoney, 2001; Dollar and Kraay, 2003; Lee and Kim, 2009; Aisen and Veiga, 2013) and following the conventional wisdom valued that "overall quality" is a positive function of the existence of institutions such as "democracy", "enforcement of property rights", as well as "the protection from risk of government expropriation", the role of institutional quality is taken into particular consideration while trying to assess whether some new light can be shed on the poorly-established relationship between economic growth and marginal taxation. In this respect, our approach resembles that of Azman-Saini, Baharumshah, and Law (2010), in their anaysis on how the impact of foreign direct investments (FDI hereafter) on growth is mediated by economic freedom, and the approach adopted in a few studies dealing with the

[3] There are two main hypotheses explaining the differences in property rights across countries. A part of literature attributes these differences to differences in natural resources endowments (Acemoglu et al., 2001, 2002, and 2005). It is argued that European colonialism led to the establishment of extractive institutions, which were characterized by the concentration of power in the hands of a few in order to maximize the opportunity of expropriation in terms of natural resources and taxes, in places relatively prosperous before the European arrival; whereas, colonization introduced led social organizations aimed at the development of institutions ensuring the security of property rights, which are essential for successful economic performance, in relatively poor areas where the European were induced to settle. Another part of literature has proposed a second hypothesis which is based on the legal origin of countries (La Porta et al., 1998, 1999, 2000; Glaeser and Shleifer, 2002). This literature argues that the historical evolution of the legal system in France and England explain cross-country differences in protection of property rights, and then current institutions. Particularly, it is sustained that the differences in historic context and the division of power between the king and the nobles in these two countries brought about the adoption of legal systems - civil law in France and common law in England – which diverge in law enforcement and legal judgement. According to Glaeser and Shleifer (2002), France chose a law system which involved a rigid legal system controlled by the crown and less security of private property because of the risk of coercion and corruption of local justice by local feudal lords; in contrast, England chose a legal system which entailed the independence of judges from crown and moved towards the protection of private property. Levine (2005) provides an overview of the former *endowment* view and latter *law view*, and affirms that these two hypotheses should be considered are complementary rather than alternative.

well-known debates on natural resource and finance curse, where a key role by the institutional setting has been identified (see Mehlum, Moene, and Ragnar, 2006; Boschini, Pettersson, and Roine, 2007; 2013, for the former case; Law, Azman-Saini, and Ibrahim, 2013; Law, Kutan, and Naseem, 2017, for the finance case). We think that institutional quality can play a crucial role on the connection between economic outcome and marginal taxation whether because good institutions foster the growth of countries, or because the resilience of an economic system to the distortionary effects from taxation is likely to depend on the availability of the right institutions.

To this aim, in this analysis we try to answer the following questions. How top marginal taxation affects economic growth? Is the relationship between marginal tax rates and growth different in countries with *good institutions* than in countries with *bad institutions*? To address the aforementioned questions we explore the impact of top statutory tax rates on real Gdp per capita, as previously done by Piketty et al. (PSS hereafter, 2014), but we deal with this issue once controlling for a well-established set of growth determining economic factors. Namely, to make our study consistent with the bulk of the empirical growth literature (e.g. Mankiw, Romer, and Weil, 1992; Islam, 1995; Barro, 1991, 2000) our analysis accounts for the effects of gross investments share in Gdp, population growth, and human capital indicators.

Methodologically, three important problems may arise in growth context. A first issue concerns with isolating long-term from short-term effects. Second, unobserved heterogeneity and omitted variable through country-specific effects might be key sources of potential biases. Fixed effects estimators are usually used to overcome these kinds of technical problems (e.g. Islam, 1995). Another issue that may affect empirical studies of growth is endogeneity. Such endogenous correlation would arise in presence of the lagged dependent variable but not only. Endogeneity problems may also concerns fiscal variables. The causal relationship is that taxes affect growth, or more growth induces more taxes, or both? Recently, instrumental variable methods (Anderson and Hsiao, 1981; Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998) are addressed in order to account for the endogeneity problems (e.g. Bleaney et al., 2001; Azman-Saini et al., 2010; Boschini et al., 2013; Gemmell, Kneller, and Sanz, 2011; Law et al., 2017).

In order to face up to these shortcomings this work uses a dynamic fixed effects estimation strategy of the type formalized by Islam (1995) and Bond, Hoeffler, and Temple (2001), and adopted in a large number of studies on income growth (e.g. Bleaney et al., 2001; Lee and Kim, 2009; Azman-Saini et al., 2010; Aisen and Veiga, 2013; Law et al., 2017) - carried out using five-year averages panel data to extract long-term information – and deal with endogenenity and simultaneity bias by using system generalized method of moments (sys-GMM) techniques as developed by Arellano and Bover (1995) and Blundell and Bond (1998). We use a five-year averages panels in order to isolate long-term effects vis-à-vis short-term and rule out cyclical fluctuations (e.g. Bleaney et al., 2001; Durlauf, Johnson, and Temple, 2005; Esfahani and Ramirez, 2003; Acemoglu, Johnson, Robinson, and Yared, 2008; Temple 1999).

Far from being detrimental to growth, at first look, our regression analysis would suggest that marginal taxation has insignificant impacts on growth. Interestingly, once taking into account the quality of institutions, the result of Fixed-Effects estimations is that quality of institutions plays a clear role on this relationship. This analysis shows that higher marginal tax rates are harmful for growth in *bad institutions*, but this negative association disappears when *good institutions* are considered. However, this evidence is not robust to the methodological change – these results are not confirmed when sys-GMM estimator is performed instead of FE.

The remainder of the work is organized as follows. Section 2 illustrates the theoretical perspective concerned the definition of optimal income taxation schedule. In section 3 we summarize the predictions and results of empirical literature on the relationship between economic growth and marginal taxation, and institutions. In section 4 we illustrate our estimation strategy. The fifth section introduces the data, and discusses some descriptive evidence. The results are reported in section 6. Finally, in section 7 the concluding remarks are presented.

2. Optimal taxation and top marginal tax rate: theoretical perspective

A wide strand of literature concerns the definition of an optimal income taxation model. This animated participation derives from the trade-off between efficiency and equity arising in the definition of this model. On one hand, income taxation is a channel for wealth redistribution, consistent with principle of equity. With this in mind, a part of the literature argues that individuals with high incomes should be burdened proportionately more than those with lower incomes, such that the taxation system fulfills task of income redistribution, and thus yields an equality benefit. On the other hand, income taxation generates a strong distorting effect on labor supply, causing a disincentive in productivity effort and individual initiative. Given this loss in efficiency, another strand of economic literature suggests that top marginal tax rate should be as low as possible in order to limit the distorting effect on agent's labour decision.

The modern models of optimal income taxation are based on the theoretic framework developed by Mirrlees (1971), who was the first to formulate theoretically the distorting effect of income taxation on individual labor supply and to build a model of optimal income taxation expressing the key issue of equity-efficiency trade-off which government deals with. In a nutshell, Mirrlees (1971) illustrates that the level of optimal marginal tax rate must not be negative and greater than 100%. Considering the Mirrlees' (1971) intuitions, Seade (1977) indicates that optimal marginal tax rate should be zero at the income level of the top earners when condition of bounded skill distribution is assumed. However, Tuomala (1990) empirically shows that this result is very local, in the sense that it arises only at the single person which fills the position at the top of income distribution. Moreover, Saez (2001) writes that "*unbounded distributions are of much more interest than bounded distributions to address the high income optimal tax rate problem*", and suggests, as Diamond (1998) has done previously, that the Pareto distribution represents a good approximation of the shape of skill distribution.[4]

[4] See Saez 2001 for more details.

In Figure 1, the bold line illustrates the Pareto parameter estimated using US adjusted gross income data. This graph clearly shows that the upper tail of income distribution could be approximated by the Pareto parameter.

[FIGURE 1 ABOUT HERE]

Taking the Pareto distribution as a good approximation of skill distribution, Saez (2001) builds a theoretical model of optimal income taxation, and derives a formula for the optimal top marginal tax rate in terms of top marginal tax rate elasticity and Pareto parameter. Following, we briefly illustrate the model of optimal income taxation developed by Saez (2001).

First, Saez assumes that each taxpayer maximizes their utility function $u = u(c, z)$ which depends positively on consumption c and negatively on earnings z expressing the cost of producing income z. Yet, each taxpayer faces a linear budget constraint $c = z(1 - \tau) + R$, where the marginal tax rate is τ and R is the non-labour income.

Solving the first-order condition of the individual maximization problem, $(1 - \tau)u_c + u_z = 0$, Saez defines the uncompensated elasticity such that

$$\zeta^u = \frac{1-\tau}{z}\frac{\partial z}{\partial(1-\tau)}. \quad (1)$$

The income effect is

$$\eta = (1 - \tau)\frac{\partial z}{\partial R}. \quad (2)$$

The compensated elasticity that indicates the variation of the agent income keeping the same indifference curve is given by

$$\zeta^c = \left[\frac{1-\tau}{z}\frac{\partial z}{\partial(1-\tau)}\right]|_u. \quad (3)$$

These three concepts are linked by the Slutsky decomposition, which is written as

$$\zeta^c = \zeta^u - \eta. \quad (4)$$

Furthermore, Saez assumes a constant linear tax rate τ for incomes above a specified threshold $\bar{z}$, and a population with income above $\bar{z}$ equal to 1. In order to derive the optimal marginal tax rate, he imagines a small positive variation $d\tau$ in the marginal tax rate τ for incomes above $\bar{z}$. This implies two effect, named by Saez, *mechanical effect* and *behavioral effect*. The mechanical effect reports the positive variation in tax revenue generating when no behavioral response is allowed. The behavioral effect represents the change in tax revenue deriving from the decrease in earnings dz as a consequence of behavioral response. Using equations (1) and (2), Saez computes the earnings variation dz and the total tax revenue change B due to behavioral effect[5] such that

$$dz = -(\zeta^u - \eta\bar{z})\frac{d\tau}{1-\tau} \tag{5}$$

$$B = -(\overline{\zeta^u} z_m - \bar{\eta}\bar{z})\frac{\tau d\tau}{(1-\tau)} \tag{6}$$

where $\overline{\zeta^u}$ and $\bar{\eta}$ are the weighted average uncompensated elasticity and the average income effect, respectively. Instead z_m indicates the mean of incomes above $\bar{z}$.

By computing the welfare effect and by definition of the average marginal social welfare weight $\bar{g}$,[6] Saez suggests that a change of tax revenue by one more dollar decreases on average top earners social welfare by $\bar{g}$. Thus, he states that total welfare loss equals to $\bar{g}M$, and government chooses marginal tax rate solving the equation $(1-\bar{g})M + B = 0$, in other words, government chooses optimal marginal tax rate such that the welfare effect equals to the total revenue effect. Moreover, imaging that the skill distribution extends to infinity, Saez assumes that elasticity $\overline{\zeta^u}$ and income effect $\bar{\eta}$ converge when $\bar{z}$ rises, and the ratio $\frac{z_m}{\bar{z}}$ converges to $\frac{a}{(a-1)}$ where a is the Pareto parameter.[7]

Given those assumptions, Saez defines the optimal marginal tax rate as follows

[5] The total tax revenue change is given by summing the terms $\tau d\tau$ over all taxpayers with income above $\bar{z}$.

[6] The average marginal social welfare weight $\bar{g}$ expresses the government aims for redistribution, that is, it indicates the social value of increasing the utility of taxpayers with incomes above $\bar{z}$, measured in units of tax revenue.

[7] The Pareto parameter has the key property that the ratio $\frac{z_m}{\bar{z}}$ takes the same value, which equals to $\frac{a}{(a-1)}$ with *a>1*, for all incomes level above $\bar{z}$.

$$\bar{\tau} = \frac{1-\bar{g}}{1-\bar{g}+\overline{\zeta^u}+\overline{\zeta^c}(a-1)}\,. \qquad (7)$$

Therefore, the optimal marginal tax rate depends negatively on the average uncompensated elasticity $\overline{\zeta^u}$. Moreover, for a given compensated elasticity *($\overline{\zeta^c}$)*, the optimal marginal tax rate gets higher, the larger is the absolute income effect *($-\bar{\eta}$)* compared to the uncompensated elasticity *($\overline{\zeta^u}$)*, which means that optimal marginal tax rate could increase when income effect dominates substitution effect because taxpayers continue to work although marginal tax rate rises. Finally, the optimal marginal tax rate is a decreasing function of the Pareto coefficient, which expresses the thinness of the top distribution. Indeed, if the top distribution is thick, a large number of persons would be affected by the distorting effect deriving from a higher marginal tax rate, implying a remarkable loss of efficiency.

PSS (2014) present a theoretical model of optimal income taxation which has the advantage to formulate an optimal top marginal tax rate capturing the three channels of response through which top marginal tax rate influences top earners labor supply: the standard supply-side response, the tax-avoidance response, and the bargaining-compensation response.[8] In fact, economic literature indicates three main channels of response. The first channel, named supply-side elasticity, was proposed by Lindsey (1987) and Feldstein (1995), who suggested the fewer the top marginal tax rate is, the grater the economic activity of top taxpayers is, in other words, a cut of top marginal tax rate encourages top earners, which benefit of tax advantage, to make a greater productive effort, and therefore it may be a conductor of higher income growth. The second channel has been noted by Slemrod (1996), and involves a tax avoidance elasticity for which a fewer top marginal tax rate entails a less incentive of taxpayers to shift part of their taxable income in another form or in another time period that is subject to more favorable taxation. The third channel, highlighted by PSS, consists of a bargaining compensation elasticity for which the fewer the top marginal tax rate is, the greater the incentive of top taxpayers to act a more aggressive bargaining is. According to PSS, if top marginal tax rates are very

[8] PSS consider tax-avoidance response, and derive a formula of optimal marginal tax rate as a function of supply-side and tax-avoidance response. They indicate that the optimal tax system should minimize tax-avoidance elasticity, and suggest that, considering just capital gains as a tax-avoidance channel, tax-avoidance elasticity appears to be small in the long run. Given those considerations, for conciseness, we omit the theoretical model of tax-avoidance response. See Piketty et al. (2014) for more details.

high, the net reward for more aggressive bargaining will be modest, while if top marginal tax rates are very low, it will be considerable, and then top taxpayers could be incentivized to bargain aggressively for raising their compensations. Therefore, according to bargaining compensation elasticity, a positive top income variation depends on a more aggressive bargaining effort rather a greater productive effort, linked to a bigger economic growth.

Following Saez' (2001) model, PSS illustrates a standard model reporting only the *supply-side elasticity*. They assume that each taxpayer maximizes their utility function

$$u = u(c, z) = c - h_i(z). \tag{8}$$

Clearly, the agent's utility depends positively on consumption $c = z - T(z)$, that is, the disposable income, and negatively on the cost of realizing z earnings, $h_i(z)$. Solving the first-order condition of the individual maximization problem, $h_i'(z) = 1 - \tau$, they define *supply-side elasticity* e_1 as

$$e_1 = \frac{1-\tau}{y} \frac{dy}{d(1-\tau)}. \tag{9}$$

Considering only the supply-side elasticity, the optimal marginal tax rate is defined as follows

$$\widetilde{\tau} = \frac{1}{1+ ae_1} \tag{10}$$

where a still represents the Pareto parameter.[9]

Furthermore, PSS capture bargaining effect. The idea that bargaining effort could play a significant role in determining top earners compensation derives from PSS' consideration on divergence between pay and marginal economic product for top income earners. According to job-matching models, there is a Nash bargaining between employee and employer over the division of the positive surplus created by the match. Alvaredo et al. (2013) state that in presence of high top marginal tax

[9] Differently from previous equation (7), built by Saez (2001), PSS also assumes that the average social marginal welfare weight $\bar{g}$ equals to zero. The smaller the average marginal welfare weight is, the more the government cares about redistribution. They indicate that this assumption allows them to compute the high income tax rate maximizing tax revenue. Moreover, they consider a model without income effects implying $\overline{\zeta^u} = \overline{\zeta^c}$ in Saez' (2001) framework.

rates the net beneficial deriving from more bargaining effort for more compensation was modest; however, after tax cuts reforms top earners have been incentivized to bargain more aggressively in order to increase their individual earnings. Divergence between payment and marginal product can also occur in presence of imperfect information because it arises issues in estimating individual marginal product, and it occurs especially for top income earners, such as executives and management positions, because they are occupations with more or less unique tasks, in which it is difficult to get full information about their individual marginal product.[10]

In the bargaining model, the taxpayer utility function is

$$u = u(c,\ \eta,\ y) = c - h_i(y) - k_i(\eta). \tag{11}$$

Thus, the agent's utility depends positively on consumption, and negatively on the cost of realizing real y product, $h_i(y)$, and the cost of receiving a fraction η of his real product through bargaining effort, $k_i(\eta)$.

PSS express $E(b)$ as the average bargained earnings in the economy, and assume that the government rules out every gain in $E(b)$.[11]

Denoting $b = (\eta - 1)y$ as the fraction of earnings obtained through bargaining rather than productive effort, PSS express the earnings given by $z = \eta y = y + b$, namely the received earnings, and suppose that the government burdened those earnings with nonlinear taxation.

Given those assumptions, PSS derive an expression for the total compensation elasticity e as

$$e = \frac{1-\tau}{z}\frac{dz}{d(1-\tau)}. \tag{12}$$

Moreover, they define fraction of behavioral response deriving from bargaining effort s and *bargaining elasticity* e_3 as

$$s = \frac{db/d(1-\tau)}{dz/d(1-\tau)} + \frac{dy}{d(1-\tau)}. \tag{13}$$

[10] For more details see Alvaredo et al. (2013), Piketty and Saez (2013), and Piketty et al. (2014).

[11] Since in the aggregate the total product should equal to the total compensation, PSS sustain that the overpayment $E(b)$ comes at the expense of others if $E(b) > 0$, and the underpayment is gained by somebody if $E(b) < 0$. For simplicity, they assume that any gain deriving from bargaining effort comes homogeneously at the expense of everyone else in the economy.

$$e_3 = s * e = \frac{1-\tau}{z} \frac{db}{d(1-\tau)}. \quad (14)$$

There are two channels through which marginal tax rate variations may affect government' budget. First, it may derive from change of the average bargained earnings in the economy $E(b)$; second, it arises from variation of tax collection from top earners bracket $(z - \bar{z})$.

The government's problem is to choose τ to maximize the $T = \tau [(z(1-\tau) - \bar{z}] - N * E(b)$, defining total population as N. Solving the first order condition with respect to τ, PSS obtain the optimal marginal tax rate maximizing tax revenue as a function of *supply-side elasticity* and *bargaining elasticity*

$$\tilde{\tau} = \frac{1+ae_3}{1+ae} = 1 - \frac{a(\frac{y}{z})e_1}{1+ae}.^{12} \quad (15)$$

Therefore, PSS suggest that the optimal marginal tax rate depends negatively on the total compensation elasticity e, keeping e_3 constant, and on the supply-side elasticity e_1, keeping e and $\frac{y}{z}$ constant. On the other hand, optimal marginal taxation depends positively on bargaining elasticity e_3, keeping e constant, and on the level of overpayment $\eta = \frac{y}{z}$, keeping e and e_1 constant. At an extreme point, PSS state that $\tilde{\tau} = 1$ if $e_1 = 0$.

Examining the relationship between the earnings z and the productivity y of top earners, two alternative distribution statuses are drawn by PSS:

1. *trickle-up* situation in which $z >$ y - overpayment in favour of top earners:

- $s > 0$;
- $e_3 > 0$;
- The optimal marginal tax rate is higher than the standard supply-side elasticity;
- Marginal taxation reduction triggers an income shifting from the bottom toward the upper income share.

2. *trickle-down* situation in which $z <$ y - underpayment at expense of top earners:

- $s < 0$;

[12] Proof: $[z - \bar{z}] = \frac{(\tau-s)}{dz/d(1-\tau)}$; e$\left[\frac{(\tau-s)}{(1-\tau)}\right]$ = $\frac{(z-\bar{z})}{z}$ = $\frac{1}{a}$; since $e = \frac{e_3}{s}$, then e_3 = s$\left[\frac{(1-\tau)}{(\tau-s)}\right]\left[\frac{1}{a}\right]$; $a\big((\tau - s)e_3\big) = s(1-\tau)$; $\tau a e_3 + s\tau = s + s a e_3$; since $e_3 = se$ and $\tau(ase + s) = s(1 + ae_3)$, hence $\tilde{\tau} = \frac{1+ae_3}{1+ae}$.

- $e_3 < 0$;
- The optimal marginal tax rate is lower than the standard supply-side elasticity;
- Top marginal tax rate reduction triggers an income shifting towards the bottom of the income distribution.

3. Empirical literature review

In this section, we first present the literature examining the relationship between marginal taxation and economic growth. In order to proceed with the red line of this work, i.e. assessing whether accounting for institutional quality may contribute to fill the gap about current knowledge of the effects of marginal taxation on growth, we need to focus on the literature on institutions and growth. To this aim, the literature exploring the direct and indirect effect of institutions on growth is illustrated in the second part of this section.

3.1 Marginal taxation and growth

A number of empirical studies has examined the impact of taxation on growth. Most have focused on average tax rates (see e.g. Barro, 1990; Angelopoulos, Economides, and Kammas, 2007; Bania, Gray, and Stone, 2007; Reed, 2008; Bergh and Karlsson, 2010; Gemmell et al., 2011) while fewer researches have used the marginal tax rates (see e.g. Lee and Gordon, 2005; Arnold, 2008; Barro and Redlick, 2011; Piketty et al., 2014; Gemmell et al., 2014; Mertens, 2015).

When focusing only on the literature that has studied the effects of marginal taxation on long-run growth, no clear-cut evidence can be established. Recent contributions of the empirical literature where marginal taxation has been found to impact negatively the economic growth include Gemmell et al. (2014) and Mertens (2015). The former who use Pooled Mean Group (PMG) estimator proposed by Pesaran, Shin, and Smith (1999) find negative evidence when using top statutory marginal tax rates. Their results imply that a 10 percentage point cut in the marginal rates increases the annual Gdp growth by 0.6 percentage points. The latter use a structural vector autoregressive (SVAR) model, and finds that a cut in the marginal rates, while determining as expected an increase in average top 1% income share, also leads to a rise in real Gdp per capita of up to 0.44 percent. This positive effect

on real Gdp, together with a positive spillover on incomes outside of the top 1%, is seen as evidence that the increase in upper incomes derives from a larger productive effort rather than a shifting from corporate to personal income, namely tax avoidance. In this analyses, a standard supply-side elasticity mechanism is playing a major role – that is, the lower the top marginal tax rate, the grater the economic activity by top taxpayers (who make a greater productive effort), and the higher overall Gdp growth.

These results are consistent with the findings of Lindsey (1987), Feenberg and Poterba (1993) and Feldstein (1995, 2006) who show that tax cuts have a positive impact on top earners productivity. For instance, Lindsey (1987) analyzed behavioral response of American top earners after a fiscal policy change (ERTA) acted in 1981, finding a rise of their taxable income share relative to the one before the cut of marginal tax rate. Feenberg and Poterba (1993), using tax return data from 1951 to 1990, examined the surge in top income share in the 1980s highlighting that while it increased gradually in the early 1980s, it rose considerably in the end of 1980s. This would suggest that a fraction of the increase in the upper incomes share derived from the reduction of marginal taxation acted in the 1986 Tax Reform Act. Also, Feldstein (1995) has compared individual tax returns before and after the 1986 tax reform, which reduced tax rates significantly for upper earners, finding a substantially higher taxable income elasticity of top taxpayers, indicating that high marginal tax rates determine relevant deadweight losses.

In another strand of the literature, however, the researchers have found an insignificant effect of marginal taxation on growth, in line with the studies by Easterly and Rebelo (2003) and Mendoza et al. (1997). Piketty et al. (2014) do not find any statistically significant correlation between marginal rates and growth in real Gdp per capita when using yearly data on 18 OECD countries from 1960 to 2010. On the same vein, Lee and Gordon (2005) while using cross-country data during 1970-1997. Dahlby and Ferede (2012), exploiting data from 10 Canadian provinces over the period 1977-2006, and Gale, Krupkin, and Rueben (2015) do not find statistically significant effect of top personal income tax rates on economic growth.[13]

On the whole, at least for marginal income tax rate, it seems still valid the statement

[13] At an extreme point, Milasi (2013) finds that an increase in marginal tax rate is positively correlated to an increase in economic growth.

by Easterly and Rebelo (1993) that ‘‘the evidence that tax rates matter for economic growth is disturbingly fragile.”

This no-significant evidence is consistent with the PSS’ bargaining-compensation hypothesis suggesting that upper incomes concentration has been mostly driven by top earners’ rent extraction behaviors favored by drops in marginal taxation rather than greater productivity effort.[14] Campbell and Lusher (2015) who use a difference-in-difference approach examine how much international trade and skill-biased technological change (SBTC hereafter) have contributed to the increase in income inequality, considering the manufacturing sectors that are more exposed to international trade. They find that both international trade and SBTC do not explain to a greater extent the rising in inequality. Hence, they use an international dataset of 18 countries to test alternative causes. Their results show that the level of top 1% income share depends on lagged values of marginal tax rate, and indicate these as supporting evidence for bargaining compensation hypothesis.[15]

It is also probably important focusing on the different nature of some determinants of taxable income elasticity, additional or even substitutive of labor supply reductions. Taking into account the different nature of some “sheltering behaviors”, which do not reduce labor supply, such as tax avoidance and tax evasion, has relevant consequence on the use of the taxable income elasticity in the computation of welfare losses (Slemrod, 1996). For example, in response to a tax increase some

[14] Another strand of literature suggest that rent extraction behaviors of top earners are favored by social norms and institutions changes (Piketty and Saez, 2007; Bakija et al., 2012; Bakija, 2013; and Bell and Van Reenen 2014). Social norms and institutions changes, such as labor and financial deregulations, would have increased rent-seeking behavior because it would have reduced preexisting limit in paying high compensations at “managers” (Piketty and Saez, 2007; and Bakija, 2013). Bell and Van Reenen (2014) indicate that rents-extraction may derive from both lack of product market competition within the sector and implicit or explicit guarantees of bailout by governments, which might encourage bankers to take risky asset as shareholders could benefit from huge positive gains, and at the same time shareholders could suffer small losses thanks to subsidies and garantees received in case of negative trend.

[15] According to Kaplan (2012), who compared compensation growth of CEO and other highly paid workers, CEO’ pay “has remained relatively constant or declined”. He indicates these as evidence that CEO’ rising pay was a result of increasing in demand for CEO’ skills and other highly paid occupations rather rent-seeking behavior. Yet, Bakija et al. (2012) and Bakija (2013) analyse the percentage change of pre-tax income going to top 1% by occupation, and show that income growth rates across occupations in the top 1% present different patterns. Especially, they stress that 70% of the pre-tax income growth going to top 1% concerns executives, managers, supervisors, entrepreneurs and financial professionals. Given those considerations, Bakija (2013) suggests that the heterogeneity in pre-tax income variation within top 1% cannot be explained by tax reasons, while, it would partially depend on institutional factors, such as corporate governance issues, which might have encouraged behavior aimed at seeking rent.

individuals may avoid taxation through charitable contributions (Chetty, 2009).[16] For the top-earner this can certainly be a sub-optimal allocation but, looking at overall output level, the main point is that many sheltering behaviors are not reducing Gdp.[17]

On the whole, accounting for that part of taxable income elasticity not related to "real variations" in labour supply, together with the idea that the lower marginal taxation causes the surge of upper incomes, at least in part, in form of rent-seeking at the expense of lower incomes in contrast to additional productive effort, might lead to wrong evaluation of efficiency-equity costs of marginal taxation variations, and have relevant consequence on the sound of some policy recommendations.

3.2 Institutions and growth

Over the last decades research on comparative development has emphasized the role of institutions. Since the term *institutions* includes a wide range of institutional aspects, the proxy for institutions refers diverse indicators, which include measures that describe the characteristics of institutions, such as measure of political regime (democracy, dictatorship) and political instability (revolutions, coups), and those that evaluate the performance of institutions, such as the measure of quality of institutions (the enforcement of property rights).

Several studies examining economic growth determinants have investigated the influence of political regime characteristics, such as the presence or absence of democracy, and have found mixed evidence. Gerring, Bond, Barndt, and Moreno (2005, p.323) review the empirical literature until the mid-2000s, and state that *"the predominant view is that democracy has either a negative effect on GDP growth or no overall effect"*. In this line of research are Barro (1999), Tavares and Wacziarg (2001), and Mulligan, Gil and Sala-i-Martin (2004). In contrast, others scholars have suggested that democracy plays a significant and positive role (Persson and Tabellini, 2006; Papaioannou and Siourounis, 2008). For instance, Papaioannou and Siourounis

[16] The relevance of tax avoidance response among high-income earners, and its implications on policy recommendations have been stressed by severe researchers (see, e.g., Gordon and Slemrod, 2000; Saez and Slemrod, 2012; Slemrod and Gillitzer, 2014).

[17] By using a Finnish Tax Administration data set, Harju and Matikka (2016) split overall taxable income elasticity (ETI) in "real" and income-shifting part. The latter is found to have a prominent role on determining the size of overall taxable income elasticity.

(2008) using a new dataset of political transitions in the 1960 – 2005 period (Papaioannou and Siourounis, 2007) find that democratization is associated with a higher level of growth over the long run. Lee and Kim (2009) still find a positive association. While exploiting a measure of political constraints of executives provided by Polity IV Project, they examine the determinants of long-run economic growth using fixed effects and sys-GMM panel estimations of 32 developed and developing countries in the period 1965 – 2002, and find a positive correlation between institutions and economic performance. Moreover, Persson and Tabellini (2006) who perform a difference-in-difference strategy to examine the effect of democracy on Gdp per capita for 150 countries from 1960 to 2000 provide evidence that democracy boosts Gdp growth by almost 0.75 percentage points.

Some other studies focus on political instability effect. It is shown that the political instability influences economic performance by reducing the efficiency of investment. For instance, while estimating cross-sectional regressions for 98 countries in the period 1960 – 1985, Barro (1991) who uses both the number of revolutions and coups and the number of political assassinations from Bank's (1979) data set as proxy for political instability shows that the presence of political violence affects economic growth negatively. When applying cross-sectional regressions, Mauro (1995), Alesina and Perotti (1996) and Perotti (1996) still find that political instability affects economic performance negatively. To overcome a potential reverse causality issue arising in the association between political instability and economic outcome, Abadie and Gardeazabal (2003) who model a quasi-experimental analysis comparing the Gdp per capita in the Basque Country with that of a synthetic control region without political terrorism show that political instability exercises adverse influence on economic prosperity. Furthermore, considering the count of cabinet change as indicator of political instability, Aisen and Veiga (2013) provide similar findings while using a sys-GMM strategy for a panel data of 169 countries in the 5-year periods from 1960 to 2004.

However, the adverse effects of political instability appear insignificant once controlling for the quality of institutions (Barro, 1997). Knack and Keefer (1995) suggest that measures of political instability, such as the count of revolutions and coups and the number of assassinations, might poorly capture the security of

property rights; whereas, performance measures, evaluating the quality of institutions in terms of efficiency and effectiveness in the enforcement of property rights, would fit better the economic institutions as defined by North (1991). Quoting Knack and Keefer (1995, p.210) *"Unlike the Gastil data, these two sources* [International Country Risk Guide (ICRG) and Business Environmental Risk Intelligence (BERI)] *provide detailed ratings for large samples on disaggregated dimensions of property rights that are closely related to those institutions emphasized by North (1990), Weingast (1993), Olson (1982) and others"*.

Several scholars have documented that the institutional quality, as the enforcement of property rights as well as the protection from risk of government expropriation, matter for economic development. Most of these studies examining whether institutional quality influences economic performance are cross-sectional regressions (see e.g. Knack and Keefer, 1995; Acemoglu, Johnson, and Robinson, 2001; Acemoglu, Johnson, Robinson, and Thaicharoen, 2003; Rodrik, Subramanian, and Trebbi, 2004; Mehlum et al., 2006; Brunnschweiler, 2007). Increasingly, however, growth analyses have been carried out using panel data (see e.g. Busse and Hefeker, 2007; Azman-Saini et al., 2010; Boschini et al., 2013). Some scholars have used ordinary least square (OLS) estimator (Knack and Keefer, 1995; Barro, 1996, 2013), while others have applied instrumental variable strategies (either two stage least square or generalized method of moments estimators) to deal with potential endogeneity (Hall and Jones, 1999; Acemoglu et al., 2001; Dollar and Kraay, 2003; Brunnschweiler, 2007; Acemoglu, Gallego, and Robinson, 2014). Indeed, the relationship between economic institutions and growth might suffer from reverse causality – i.e. institutions may cause economic growth, but also rich countries may be conducive to the growth of better institutions.

Under the OLS cross-sectional perspective, Knack and Keefer (1995) find a positive association when using the International Country Risk Guide (ICRG) index as proxy for institutional quality for 97 countries in the period 1974 – 1989. Also, Barro (2013) who applies OLS cross-sectional estimations in the period 1960 - 1995 to investigate determinants of economic growth point out that quality of institutions, measured with ICRG indicator, has a positive effect.

Studies using instrumental variables methods show similar findings. Acemoglu et al. (2001) explore the relationship between economic growth and institutions by applying an IV analysis which, they state, allows them to identify the causal effect of institutions on economic prosperity. They do so by using the indicator of institutional quality provided by Political Risk Services, namely the ICRG index, and by instrumenting the latter by settler mortality rates in the colonies more than 100 years ago.[18] While estimating a cross-sectional two-stage least square (TSLS) model for 64 countries in the 1995, the authors show that economic institutions cause economic prosperity, positively. Likewise, Acemoglu et al. (2014), answering to the Glaeser and Shleifer (2002) critique,[19] provide additional evidence that the quality of institutions, measured by the rule of law index provided by Worldwide Governance Indicators (Kaufmann et al., 2010), matter for economic development once controlling for human capital. In this analysis, they apply three estimation methods (OLS, TSLS, and semi-structural models) using both cross-country data of 62 colonies and cross-regional data of 670 sub-national colonial regions in the 2005. In the same vein, Rodrik et al. (2004) who compare the role of economic institutions, geography, and trade in determining economic development by using cross-sectional OLS and TSLS methods for 137 countries in the 1995 show that the institutional quality *"trumps"* both geography and trade factors.

Some other studies use cross-regional variation to explain cross-country Gdp differences. Banerjee and Iyer (2005) exploit cross-regional differences in the political environment in India, and find a positive correlation between security of property rights and economic development. During the British colonialism, some regions applied landlord-based systems, which attribute the liability to set and collect the land taxes at the landlord class, while in other regions, which have an individual cultivator-based system, this liability laid directly with the individual

[18] The AJR (2001) idea is that Europeans introduced better economic institutions at the extent they could benefit from that. This means that they created or maintained weak economic institutions in places with enormous opportunity of resources extraction, that is areas previously-rich in endowments and densely-populated; while, they shaped economic institutions aimed at enforcing the property rights in areas previously-poor in endowments and in indigenous population density, and especially optimal for Europeans settlement. Since the initial conditions may influence the Europeans decision on the establishment of settlement, AJR suggest the disease environment as an exogenous source of variation in current economic institutions, and then the settler mortality rates as a valid instrument. They defend the exogeneity of their instrument by arguing that the settler mortality rates affect current economic growth only via their effect on institutions development.

[19] Glaeser et al. (2004) criticize the exclusion of human capital in the analysis on institution and development of AJR (2001). They sustain that institutions does not seem to affect economic growth once including human capital in the AJR (2001) framework.

cultivator. According to Banerjee and Iyer, in landlord-based areas the peasant property rights were relatively insecure and cultivator's investments were discouraged from the risk of expropriations by landlord class. On the other hand, in non- landlord areas a written contract ensured the security of property rights of peasants. They suggest that the remarked political environment persist into the post-independence period, and find that the effects of this institutional framework explain the current differences in investments and economic development across Indian regions when estimating OLS and TSLS panel regressions in the period 1956 – 1985. Moreover, Di Liberto and Sideri (2015) examine the relationship between institutional quality and economic performance by exploiting Northern - Southern region heterogeneity in firm's productivity or per capita GDP in Italy. While using historical dominations which have ruled Italian regions in the past centuries as exogenous instrument for current institutional performance, they find evidence of the significant and positive role of the quality of institutions in explaining economic development (see also Michalopoulos and Papaioannou, 2013; Agostino, Nifo, Trivieri, and Vecchione 2016).

Other scholars investigate the connection between economic development and institutions by focusing on the legal origins of countries. The idea is that differences in legal traditions shaped in the past explain cross-country differences in protection of property rights, and then current institutions (La Porta, Lopez-de-Silanes, and Shleifer, 1998, 2008; La Porta, Lopez-de-Silanes, Shleifer, and Vishny 1999, 2000; Glaeser and Shleifer, 2002). This literature sustains that the differences in the historic context and in the division of power between the king and the nobles in France and England brought about the adoption of legal systems - civil law in France and common law in England – which diverge in private property enforcement and legal judgement. According to Glaeser and Shleifer (2002), France chose a law system which involved a rigid legal system controlled by the crown and less security of private property because of the risk of coercion and corruption of local justice by feudal lords; in contrast, England chose a legal system which entailed the independence of judges from crown and moved towards the protection of private property. While using cross-section data for 152 countries, La Porta, Lopez-de-Silanes, Shleifer and Vishny (1999) find that legal systems with more state

interventionism, such as socialist or French civil law, are associated with inferior economic outcomes compared to those in common law countries. Mahoney (2001) shows that common law countries, which exhibit greater protection of property rights, have faster economic growth than civil law countries when estimating generalized method of moments (GMM) regressions for 102 countries in the period 1960 – 1992. Furthermore, Agbor (2015) who investigates the importance of colonial legal origins on current economic growth by applying Hausman-Taylor estimation method for 36 Sub-Saharan African countries in the period 1960 – 2000 suggests that former British colonies exhibit marginally faster growth rates than former French ones.

3.2.1 Indirect institutions effects

In a few studies, the focus is on the indirect effect of institutions on growth, via their effects on other determinants. Azman-Saini et al. (2010) who uses a proxy of institutional quality provided by the Fraser Institute estimate sys-GMM regressions for 85 countries in the period 1976 – 2004 and find that FDI effects on growth are conditional on the quality of institutions in the host countries - a positive effect on growth is detected in countries with high protection of property rights, while in those with low property rights enforcement FDI does not exercise effects.

In the same vein, Boschini et al. (2013) take into account the role of institutions on the relationship between economic growth and natural resource abundance. They do so by estimating both cross-section and 5-year averages panel regressions for 75 countries during the period 1965-2005. The natural resource curse hypothesis seems to be confirmed in countries with low institutional quality but a reversal effect emerge when institutional quality is good enough.

Law et al. (2017) examine the indirect effect of institutions on the interplay between economic growth and finance development instead. A U-shape relationship supporting the finance curse view appears when analyzing the direct effect between these variables; however, above a certain minimum level of institutional quality, a positive connection between finance and growth is found.

In the rest of the paper, we explicitly reconnect our analysis to these just mentioned studies on the indirect role of institutions on natural resource exploitation, FDI and finance in a growth context.

4. Empirical strategy

While using a panel data set of 26 OECD countries over the period 1981-2015, we perform our analysis starting from a growth regression model specified as follows:

$$y_{it} = \alpha + \delta\, y_{it-1} + \beta mtr_{it} + \lambda inst_{it} + \gamma X_{it} + \eta_t + \mu_i + \varepsilon_{it}\,. \qquad (16)$$

To make our analysis consistent with main empirical growth approaches, all variables are expressed in logarithms and are five-year averages. In equation (16) the dependent variable (y_{it}) is the log of real Gdp per capita in country *i* and at time *t*. In the right hand-side we include the log of the lagged dependent variable (y_{it-1}) to capture the process of country convergence. The variable mtr_{it} denotes the log of marginal tax rates and $inst_{it}$ represents our measure of the institutional quality, our variables of interest. The vector of control variables, $X_{it,}$ includes physical capital investment, human capital, population growth, average tax rate, tax decentralization, and income inequality. The variable η_t is a set of time-specific effects, μ_i denotes a set of country-specific effects, and ε_{it} is the idiosyncratic error term.

Different methods are available for the estimation of an empirical growth model. The panel data approach can be one. Longitudinal data offer significant advantages because more information are available, hence, estimates are more efficient, but especially because panel data allow controlling for unobserved heterogeneity by means of fixed effects estimation methods. However, technical issues may arise, particularly, in the context of growth, as well highlighted, for example, by Caselli, Esquivel, and Lefort (1996) and Durlauf, Johnson, and Temple (2005). At this regards, starting from Islam (1995) the empirical literature suggests that growth models should include some dynamics in lagged output, to capture convergence processes. Consequently, the resulting structure of the error term in a fixed effects framework entails well-known endogeneity problems. An additional source of bias may arise

due to endogeneity of some explanatory variables (e.g. investments and fiscal variables: e.g. Bleaney et al., 2001; Gemmell et al., 2014).

Assume to estimate the following regression with panel data

$$y_{it} = \rho y_{it-1} + \beta x_{it} + u_{it} \quad (17)$$

$$i = 1, \dots, N \; ; t = 2, \dots, T$$

$$y_{it-1} = \rho y_{it-2} + \beta x_{it-1} + u_{it-1} \quad (18)$$

$$u_{it} = a_i + \varepsilon_{it} \; ; \; u_{it-1} = a_i + \varepsilon_{it-1} \, . \quad (19)$$

OLS estimates would be biased because of unobserved heterogeneity, that is, the standard OLS assumption of no-correlation between the error term and the explanatory variables would be violated, since a_i is part of the process that generates y_{it-1}.

Fixed effects (FE) estimators ensures that any unobservable time-invariant heterogeneity as well as omitted variables, which are constant over time, will not bias the estimations, although there is correlation between the omitted and the explanatory variables (Islam, 1995).

As indicated in equation (19), the error term is constituted of two components: an individual-specific error term a_i and an idiosyncratic error term ε_{it}, where the individual-specific error term does not change over time, namely, it is a time-constant factor, and represents the time-constant unobserved heterogeneity.

FE strategy is based on the within transformation, which consists in averaging the regression model (17) over time for each *i* such that

$$y_i = \rho {y_i}^{-1} + \beta x_i + a_i + \varepsilon_i \, . \quad (20)$$

Then, the average equation (20) is subtracted from the regression model (17) such that

$$y_{it} - y_i = \rho (y_{it-1} - {y_i}^{-1}) + \beta (x_{it} - x_i) + (\varepsilon_{it} - \varepsilon_i) \, . \quad (21)$$

This model can be estimated by FE estimators. In particular, as can be seen, the within estimator has the advantage to remove the time-constant error term from the transformed equation, hence, it allows for overcoming the time-constant

unobserved heterogeneity issue. Though, a drawback of implementing FE estimators in growth models is the loss of all between variation, as argued by Barro (1997) and Temple (1999).

The use of FE, however, does not rule out panel biases deriving from endogeneity and simultaneity. As argued by Nickell (1981), Caselli et al. (1996), and Durlauf et al. (2005), the FE estimator could still achieve biased estimations because of the non-negligible problems of endogeneity. Indeed, we have that

$$y_{it-1} - y_i{}^{-1} = y_{it-1} - \frac{y_{it} + \cdots + y_{it-1} + \cdots + y_{iT-1}}{\mathrm{T}-1} \tag{22}$$

$$\varepsilon_{it} - \varepsilon_i = \varepsilon_{it} - \frac{\varepsilon_{i2} + \cdots + \varepsilon_{it-1} + \cdots + \varepsilon_{iT}}{\mathrm{T}-1} \tag{23}$$

$$Cov\left(-\frac{y_{it-1}}{T-1}\,;\,-\frac{\varepsilon_{it-1}}{T-1}\right) \neq 0 \tag{24}$$

Since FE estimator has to satisfy the standard assumption of strict exogeneity implying no-correlation between independent variables and idiosyncratic error term, biased estimations are still possible as the correlation between transformed lagged dependent variable and transformed error term does not disappear when N will tend to infinity.[20]

Hence, we follow the indications of Bond, Hoeffler, and Temple (2001) and complement the *à la* Islam fixed effects approach by applying GMM as developed by Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998). There are several reasons because the Difference and sys-GMM estimators fit for growth context. In particular, GMM estimators allow for the presence of fixed individual effects, which are likely in panel set-up, and fit well with dynamic processes and endogenous regressors, which are the case for growth models.

The Difference GMM (diff-GMM) method applies the first-difference transformation which eliminates the fixed effects, as does also Within Group estimator, but further predetermined and endogenous transformed variables are instrumented with untransformed lags, which are orthogonal to the error term. However, diff-GMM, which was developed by Arellano – Bond (1991), built on Holtz-Eakin, Newey, and

[20] See Nickell (1981) for more details.

Rosen (1988) work, performs poorly in presence of persistent series because lagged variables in level are weak instruments for differenced-variables.[21]

System GMM estimator (sys-GMM), as proposed by Arellano and Bover (1995) and Blundell and Bond (1998), augments the Arellano – Bond (1991) method with original equations in levels, such that the standard set of equations in first-differences, instrumented with lagged levels, is combined with a set of equations in levels, instrumented with lagged first-differences.[22] In addition to the standard assumptions required by the Arellano-bond model

$$E[y_{it-s}\Delta\varepsilon_{it}] = 0,\ E[mrt_{it-s}\Delta\varepsilon_{it}] = 0,\ E[inst_{it-s}\Delta\varepsilon_{it}] = 0,\ E[X_{it-s}\Delta\varepsilon_{it}] = 0 \qquad (25)$$

$$\text{for } t = 3, \dots, T \text{ and } s \geq 2$$

Sys-GMM require the stationary assumption

$$E[\mu_i\Delta y_{it}] = 0,\ E[\mu_i\Delta mrt_{it}] = 0,\ E[\mu_i\Delta inst_{it}] = 0,\ E[\mu_i\Delta X_{it}] = 0 \qquad (26)$$

Namely, the covariance between the unobservable country-specific effects and the changes in the instrumenting variables is constant over time. Though this stationary means condition may not be reasonable for the per capita Gdp growth context, the problem can be overcome by means of the "*the inclusion of time dummies* [that transforms the variables into deviations from the means and] *allows for common long-run growth in per capita GDP, consistent with common technical progress, without violating the validity of the additional moment restrictions used by the system GMM estimator" (*Bond et al., 2001 p.16)*.* [23]

However, sys-GMM estimator performs poorly when cross-section dimension is small, as our panel is (Agell et al., 2006). Roodman (2009a p.128) affirms that *"if T is large, dynamic panel bias becomes insignificant, and a more straightforward fixed-effects estimator works"*. To this aim, we apply not only sys-GMM but also FE estimator.

In order to assess the interplay between institutions and marginal tax rates, we augment our baseline specification in Eq (16) with an interaction term given by the

[21] See Roodman (2009a, 2009b) and Bond et al. (2001).

[22] The additional information provided by the latter instruments can be important. As Roodman (2009a p.114) states "for random walk-like variables, past changes may indeed be more predictive of current levels than past levels are of current changes".

[23] See Bond et al. (2001), and Roodman (2009a) for more details.

product between *mtr* and a time-invariant dummy which takes value 1 when the country average value of institutional quality is larger than the sample average (zero otherwise) *Dinst*, while also including the two variables independently.[24]

$$y_{it} = \alpha + \delta\, y_{it-1} + \beta mtr_{it} + \lambda inst_i + \gamma X_{it} + \delta(mtr_{it} * Dinst_i) + \eta_t + \mu_i + \varepsilon_{it}\,. \qquad (27)$$

For reason of simplicity, countries with institutional quality higher than sample average are defined *good institutions*; while countries with quality of institutions lower than sample average are briefly indicated as *bad institutions*.

In order to isolate long-term from short-term information, and control for economic cycle fluctuations, FE and sys-GMM estimates of models (16) and (27) are carried out on a five-year averaged panel data set of 26 OECD countries over the period 1981 – 2015 (Mendoza et al., 1997; Lee and Kim, 2009; Azman-Saini et al., 2010; Boschini et al. 2013; Arachi et al., 2015).

5. Data and descriptive evidence

In this section, we first illustrate our data sources and report some basic summary statistics. In a second paragraph, it is presented some evidence both on the evolution of marginal taxation and real Gdp per capita over the long-run for selected OECD countries and on the cross-country relationship between marginal tax rate and economic growth and between institutional quality and economic growth.

5.1 Data

In our empirical analysis, we use different data sources to construct our five-year averaged panel data set of 26 OECD countries. As summarized in Table 1, the main source of our dataset is the OECD Statistics database. Our data on marginal taxation, real Gdp per capita, and average tax rates are obtained from this OECD Statistics database. Population growth rates information come from World Population Prospects 2017. Tax decentralization variable is the ratio between local tax revenue and total tax revenue, which computation is based on data from the OECD Statistics database as well. Data on investments come from World Development Indicators

[24] We are following here Azman-Saini et al (2010), who insert the product of FDI with an index of economic freedom in a standard growth regression.

Database (WDI), provided by World Bank. Data on years of schooling, as proxy for human capital, derive from Penn World Table (PWT). Data on inequality measure are obtained from the Standardized World Inequality Indicators Database (SWIID), constructed by Frederick Solt (Solt, 2009).The source of our proxy for institutional quality is the International Country Risk Guide (ICRG) index, which covers the period from 1984 to 2014, and is obtained from QOG Standard Database.

Our main variables of interest are marginal tax rate and institutional quality. In detail, we identify the marginal tax rate as the top statutory personal income tax rate, excluding both payroll and consumption taxes. This allows us to make our work consistent with Devereux et al. (2008), Myles (2007), Gemmell et al. (2014), and Kneller and Misch (2017) who suggest top statutory tax rates as capturing distortionary effects of increased progressivity, and recent studies that used analogous information (Lee and Gordon, 2005; Barro and Redlick, 2011; Gemmell et al., 2014; Piketty et al, 2014; Mertens, 2015). As measure of institutional quality we use the International Country Risk Guide (ICRG) index. These data are provided by the QOG Standard Database, and they relate to years ranging from 1984 to 2015. In detail, the ICRG index is composed by five governmental aspects: the quality of the bureaucracy, the political corruption, the likelihood of government repudiation of contracts, risk of government expropriation, and overall maintenance of the rule of law. [25]

We now turn to describe our remaining additional controls. Why do some countries grow much faster than others? Though literature has gained considerable knowledge about what explains growth, this remains yet a crucial question in the field of economic growth and development. Traditional neoclassical growth theories explain cross-country differences in income per capita as different paths in propensity to save, population growth rates, and total factor productivity (Solow, 1956; Cass, 1956; and Koopmans, 1965). Crucial prediction of neoclassical models, relying on the assumption of diminishing returns to capital, is that long-run growth depends only

[25] Glaeser et al. (2004) criticize the measurement of institutional quality from subjective survey indicators provided by International Country Risk Guide, and recently used in cross-country studies by AJR (2001) since"…these measures reflect what actually happened in a country rather than some permanent rules of the game" (Glaeser et al., 2004, p.276). They argue that these measures, named performance measures, could be influenced by income path and thus the analysis could suffer from reverse causality.

on exogenous technological change. However, long-run growth analyses provide evidence of persistent positive income growth rates.

A second wave of models, starting with Romer (1986, 1990) and Lucas (1988), provide an endogenous technical progress theory which ties the economic growth to the investment in human capital by using the Arrow' (1962) concept, described with the expression *learning by doing*. In these endogenous models, investment in human capital, innovation and know-how are significant contributors to the economic growth; particularly, as Romer (1990) argued, through investment in human capital workers become more productive because capture and generate more ideas due to technological spillovers, and hence the economy grows.

In accordance with these observations and with the empirical studies, which combines basic features of both neoclassical and endogenous growth models (Mankiw et al., 1992; Islam, 1995; Barro, 1991, 2000) we include a proxy for physical and human capital and population growth. Physical capital is measured by the gross capital formation as a percentage of Gdp; human capital is measured by years of schooling; and the population growth is the annual population growth rate.

Regarding other factors that shape a country's tax system, we include a proxy for average tax rate, measured as total tax revenue over Gdp, which indicates the share of a country's output that is collected by the government through taxes. The omission of the average tax rates could bias our analysis as the marginal taxation indicator could capture their effects on economic growth. In this regard, we agree with Gemmell et al. (2014) and Kneller and Misch (2017) who point out the different effects captured by average and marginal income taxes – the former the level effects instead the latter the structural effects – and the importance of including average tax rates in the regression when estimating the impact of increased progressivity through marginal tax rates.

For this variable, the survey by Bergh and Henrekson (2011) shows that a negative sign is generally expected in the case of developed countries, whereas positive effected is sometimes detected for the less developed countries, where public expenditure mostly finances basic infrastructure and education.

A theoretical explanation of this negative impact of government size on growth is provided by Olson (1982), who suggests that rich countries, with large public sectors,

are more likely fertile lands in which organized interested groups, which tend to shift public resources from providing goods of public interest to activities of no-public interest, could emerge.[26] Since our sample is composed of developed countries, it is expected a negative sign.

We also include a proxy for tax decentralization, which is computed as the ratio between local tax revenue and total tax revenue. Traditional theory sustains that subnational fiscal autonomy leads to efficient allocative outcomes, thus fiscal decentralization affects positively economic performances (Tiebout, 1956). Tiebout stated that mobile consumers "vote with their feet", in the sense that they choose to reside in the jurisdiction offering a fiscal system which suits better their personal preferences. In the same vein, Oates' Decentralization Theorem predicts that *"in the absence of cost-savings from the centralized provision of a "local public" good and of interjurisdictional externalities, the level of welfare will always be at least as high (and typically higher) if Pareto-efficient levels of consumption is maintained across all jurisdictions"* (Oates, 1972, p.54). The justification of this normative proposition is that the Pareto-efficient levels of a "local public" good likely diverge across jurisdictions because of differences in preferences and marginal costs; consequently, when the public good provision is not centralized, the overall social welfare is maximized (Oates, 1999).

While fiscal federalism theory sustains that subnational fiscal autonomy leads to efficient allocative outcomes, however, the empirical evidence on growth effects is ambiguous. A vast literature detects a positive link between fiscal decentralization and economic growth (Zhang and Zou, 1998; Lin and Liu, 2000; Thiessen, 2000; Akai and Sakata, 2001; Brueckner, 2006). Gemmell, Kneller, and Sanz (2013) analyse the effect on economic growth of both spending and revenue decentralization for a panel dataset of 23 OECD countries during 1972 – 2005, and find that spending decentralization causes lower growth while revenue decentralization is associated with higher growth. Other scholars find that fiscal decentralization negatively influences economic growth. Some references are Xie, Zou, and Davoodi (1999) for the United States, Davoodi and Zou (1998) for a sample including developed and developing countries, Zhang and Zou (1998) and Jin and Zou (2005) for the case of

[26] See Bergh and Henrekson (2011) for more details.

China. Further studies fail to find a significant relationship (Martinez-Vazquez and McNab, 2003; Thornton, 2007; Baskaran and Feld, 2013).

Possible explaination of a negative effect of tax decentralization on growth is the so-called flypaper effect. Assuming that central government gives local jurisdictions funds earmarked for education, but the jurisdictions use most of the grant for some personal interests, consequently the maximization of central government's objective has not been met. This practice was named "flypaper effect" by Arthur Okun who noted that *"money sticks where it hits"*. Numerous empirical studies have investigated the effect on local spending of various types of grants, and they have revealed that public spending is stimulated by grants-in-aid more than by voter income, contrary to theoretical expectations (Gramlich and Galper, 1973; Fisher, 1982; Hines and Thaler, 1995; Knight, 2002; Reinikka and Svensson, 2004; Inman, 2008; Fafchamps, McKenzie, and Quinn, 2014). Caselli, Esquivel, and Lefort (1996) measured flypaper effect by using federal grants data for 48 states of the United States over the period 1970 – 1985; their study finds that, ceteris paribus, an extra dollar of personal income increases government spending by 0.07 cents, instead an equivalent dollar obtained in grants increases states spending by roughly 65 cents. Turnbull (1992) suggests fiscal illusion as a cause of flypaper effect. His idea is that intergovernmental grants are less observable by voters, which are not able to link what is spent on what. Thus, local jurisdictions tend to spend most of grants to increase their popularity for being reelected rather than maximize voter's utility.

A last control considered in our analysis is income inequality, measured by the net Gini index on the grounds that an increasing attention has been recently devoted to the link between the income concentration and economic performance. As it is well known, also for the relationship between income inequality and economic growth the empirical evidence is mixed. On one hand, some studies find that an increase in income inequality is beneficial for subsequent economic growth (Li and Zhou, 1998; Forbes, 2000; Frank, 2009). This evidence is consistent with theoretical literature predicting that reducing income inequality could be harmful for economic growth. The idea is that saving propensity rates of the wealthy people are larger than those of the poor one, thus the income redistribution involves a decreasing in saving rates, and then in investments and economic growth (Kaldor, 1956).

However, the mainstream empirical evidence predicts a negative impact of income inequality on economic growth (Alesina and Perotti, 1993; Alesina and Rodrik, 1994; Deininger and Squire, 1998; Barro, 2000; Berg and Ostry, 2011; Cingano, 2014). These results support the theoretical predictions of a negative connection between income inequality and growth. It is pointed out that income inequality has a negative effect on growth because the increase in income inequality might bring about redistributive pressures or social unrest issues, which in turn negatively affect economic growth (Persson and Tabellini, 1994; Alesina and Perotti, 1993, 1996; Alesina and Rodrik, 1994; Bénabou, 1996, 2000). Moreover, due to the imperfection of capital markets, the income distribution affects negatively the economic development via inefficient investment level in human capital - the presence of imperfect capital market causes limitations on the option of borrowing for acquiring education, especially for disadvantaged individuals (Galor and Zeira, 1993; Galor and Moav, 2004; Knowles, 2005).

The main summary statistics of all these indicators are reported on annual and fifthly terms in Table 1.

[TABLE 1 ABOUT HERE]

5.2 Economic growth, marginal taxation and institutions

In table 2 we present economic growth rates in the period 1975-2015. Variations in marginal taxation and institutional quality are also illustrated.[27] In the elaboration of this table we exploit data for the period 1975-2015 in order to examine the evolution of real Gdp per capita and the variation of marginal taxation before and after the main cuts of marginal tax rate that were mostly implemented from 1979 onwards. On the same vein, also figures 2 and 3 are computed by considering the cut period 1975 – 2015.

From Table 2, it is noticeable the fact that countries which have implemented large cuts of top marginal tax rates are those countries experiencing growth rates both larger and smaller than the sample average. For instance, both Ireland and the

[27] Data limitation on ICRG index limits the descriptive analysis to time period 1984-2015.

United Kingdom, as well as the United States, acted a considerable cut of marginal taxation, nevertheless, Ireland lives a huge increase in Gdp level, while, a more modest increase in real Gdp is reported by both the United Kingdom and the United States. When considering the variation of quality of institutions, the differences in the pattern of institutional quality over time seems to be the cause of the differences in growth rates across countries. Quality of institutions in Ireland achieves a higher level, while the United Kingdom and the United States report a decrease in the level of quality of institutions over the period 1985 – 2015.

[TABLE 2 ABOUT HERE]

For the scope of this paper, it is worth noting that countries which acted a considerable reduction of top marginal taxation, such as New Zealand, the United Kingdom, and the United States, report an economic growth similar or even smaller than other countries, such as Belgium, the Netherlands, and Sweden, which have not experienced a sharp decrease in top marginal tax rates. Same pattern also emerges in Figures 2.a and 2.b, which present a graphical analysis of the relationship between economic growth and marginal tax rate by countries.[28]

[FIGURES 2.a TO 2.b ABOUT HERE]

Figure 3 reexamined the relationship between the economic growth and the variation of marginal tax rates by exploiting a cross-country perspective. Consistent with previous analytical and graphical analysis, this figure still displays that countries which report a huge decrease in marginal taxation (e.g. Australia, New Zealand, the United Kingdom, and the United States) have an economic growth as similar as other countries which have not experienced a sharp decrease in marginal tax rates (e.g. Belgium, Finland, and Germany). Importantly, we find that the correlation between growth and the change in top marginal tax rates (-0.36) shows the sign predicted by traditional literature, which indicates that larger economic growth is linked to lower

[28]Figure 2.a illustrates the relationship between real Gdp per capita and marginal tax rates for the nine countries which have experienced the largest top tax rate cut over the period 1975 – 2015; while Figure 2.b describes this correlation for the nine countries which have applied the smaller top tax rate cut.

top marginal tax rates as a consequence of increasing in productive effort by upper earners.

[FIGURE 3 ABOUT HERE]

Finally, in Figure 4 we identify a clear and high (0.54) positive correlation between Gdp per capita growth rates and our measure of institutional quality. This result supports the the idea that good economic institutions represent an inter-related *cluster of informal and formal rules* enhancing the prosperity of countries due to its effect on the setting of transaction and production costs (North, 1991; Acemoglu et al., 2001); namely, the better the institutional quality is, the more economically successful the societies are.

[FIGURE 4 ABOUT HERE]

6. Results

6.1 First approach: Fixed effects

In this section we describe results obtained by estimating Equations (16) and (27) using the fixed effects estimator which controls for unobserved time-invariant heterogeneity. Time effects are accounted for by including time dummy variables in each specification.

In Table 3 we start introducing the most parsimonious specification which estimates the real Gdp per capita with respect to lagged outcome variable, physical and human capital and to population growth rate. This construction mostly relies on empirical growth contributions of Islam (1995) and Mankiw et al. (1992). Each specification also includes the measure of marginal taxation, and in columns 2 to 3 the proxy for quality of institutions is introduced. Moreover, we enter our additional controls in Model 3: average tax rate, tax decentralization, and net inequality. In specifications 4 and 5 the role of institutional quality on the impact of marginal tax rate on economic performance is investigated. To test the existence of this effect, an interaction term

given by the product between *mtr* and our time-invariant institutions dummy *Dinst* is estimated, while also including the two variables independently.

In columns 1 through 3, focusing on our taxation variable, the estimated coefficients always display no-statistically significance. These results may induce interesting implications concerning on the relationship between marginal taxation and top earners labor offer. Importantly, the supply-side elasticity hypothesis is not supported by last evidence.

Instead, this no-significant evidence would be consistent with rent-extraction theory and PSS' statements suggesting that upper incomes concentration comes mostly from redistribution in favour of top income earners due to the bargaining elasticity rather than an actual increase in economic activity. As already mentioned, the idea is that low marginal rates ensure great gains from more aggressive bargaining effort, and then represent an incentive to practice rent-seeking by top earners. In contrast, relative higher marginal taxation would work as rent-extraction limitation due to small gains. No-significant correlation between marginal tax rates and economic growth is also in line with the arguments by Chetty (2009) who sustains that an important part of the elasticity of taxable income could be explained by sheltering behaviors and that omitting the equivalence between marginal social costs of sheltering and tax rate causes an overestimation of deadweight losses.

[TABLE 3 ABOUT HERE]

In Model 2 we consider the effect of institutional quality on our outcome measure. This specification reports that the estimated coefficient is positive and statistically significant. These results suggest that quality of institutions is a key source of economic growth, consistently with the existing literature. However, when all controls are included (Model 3) our institutional quality measure do not show a significant coefficient.

In columns 1 through 3, we observe that the coefficients related to the additional controls are never significant, with the exception of investments and tax decentralization. Coefficients of our investments indicator are always highly significant and positive. This result is consistent with empirical literature (Mankiw et

al., 1992; and Islam, 1995) suggesting that investments have a positive effect on long-run economic growth. Instead, tax decentralization shows a significant and negative coefficient. This sign supports the flypaper effect – literature (Oates, 1999) paraphrases this result saying *"money sticks where it lands"*.[29] Most importantly, in Models 2 to 3 it is worth noting that the inclusion of our controls leaves no significant role for our marginal taxation indicator.

As said above, in this study we focus on examining whether the marginal taxation effect on growth may differ according to institutional quality. Column 4 in Table 3 shows the results of this analysis when the real Gdp per capita is estimated with respect to the interaction terms and to our set of controls (lagged dependent variable, physical and human capital, population growth rate, average tax rate and tax decentralization, and then net inequality). We observe that a significant and negative relationship between marginal taxation and growth emerges in countries with lower than average institutional quality; while the estimated coefficient is significant and positive (even if almost quantitatively negligible) when *good institutions* are considered. Thus, our analysis offers a picture where the institutional quality influences the interplay between marginal tax rate and economic performance. In particular, it is evident that in *bad institutions* increasing effects on real Gdp per capita are of order of 1-2% after a 10% decrease in marginal tax rates. Hence, this result seems to suggest that in *bad institutions* a reduction of marginal taxation is growth-enhancing as predicted by standard supply-side theory. However, an opposite trend, with coefficient becoming negligible, appears in *good institutions*.

In column 5 we check whether the evidence of this opposite sign of the impact of marginal taxation on growth due to the quality of institutions holds when we multiply our institutions dummy (*Dinst*) with each continuous regressor, with the exception of the lagged dependent variable. As illustrated in specification 5, even when included jointly, the impact of marginal taxation on growth still changes according to the quality of institutions. Therefore, the result of a different effect of marginal taxation on growth according to institutional quality is confirmed.

[29] The concept of flypaper effect was first introduced by Arthur Okun to describe the phenomenon that government grants and individual income would raise local spending differently. According to theoretical expectations, receiving grants from federal government and an increase in local individual income of equal value have the same effect on local spending. However, this is not the case. Flypaper effect illustrates this anomaly. As said in section 6, Turnbull (1992) indicates fiscal illusion as an explanation of flypaper effect.

The evidence of this different effect might be explained by an efficiency problem. Assuming that countries with institutional quality level higher than sample average benefit from both the presence of efficient institutional structure and the guarantee of property rights, it is possible that the distortionary effect of taxation on top earners is offset by growth-enhancing effects of an efficient public spending. In contrast, countries with institutional quality level lower than sample average might be only exposed to efficiency costs of higher income progressivity, namely the distortionary effect of taxation, but not to the benefits.

Interestingly, Model 5 shows that the impact of tax decentralization also varies according to the quality of institutions: an increase in tax decentralization has a negative effect on growth in *bad institutions*, while this effect becomes positive in *good institutions*. This result seems to support the idea that the distortionary effect of marginal taxation might be offset by efficiency gains in public spending in *good institutions*. Namely, when the public good provision is not centralized, the overall social welfare is maximized, hence, an increase in local revenue, which might derive from higher marginal tax rates, could entail positive effects on economic performance, as argued by traditional literature (Oates, 1999).

6.2 Second approach: GMM system

Accounting for the potential endogeneity of our regressors, the fixed effects approach may be inconsistent in our framework. Bleaney et al. (2001) and Gemmell et al. (2011) show that this technical drawback can be overcome by using instrumental variable methods. In this spirit, we apply system GMM estimations. Results are summarized in Table 4.

The consistency of sys-GMM estimates relies on the presence of no-serial correlation in the idiosyncratic error terms and the validity of the instruments set. Regarding the first requirement, we report the Arellano-Bond (AR) tests of serial correlation which allows for assessing the respect of serial independence in the error terms. About the second condition, the Hansen test is performed, once controlling for the problem of instrument proliferation (collapsing the instrument matrix) which could cause unreliability of the test results (Roodman, 2009b).

For each specification, tests results are reported in Table 4. The Hansen test results do not reject the joint null hypothesis of overidentification restrictions. This implies that our instruments are valid. Moreover, in each regression the p-values of the AR tests always fail to reject the null hypothesis, suggesting that estimates are free from autocorrelation problems.

Table 4 replicates the previous Table 3 analysis changing the estimator applied – sys-GMM instead of FE. Considering our taxation variable, we always observe a coefficient no-statistically significant. Thus, these results seem to confirm previous evidence that assigns no-significant role of marginal taxation on growth, and thus seem to not support the supply-side elasticity hypothesis sustaining that marginal taxation increases are harmful for growth due to disincentive effects on top earners labour offer.

[TABLE 4 ABOUT HERE]

However, unlike our FE results that shows positive effect of quality of institutions on economic outcomes, sys-GMM analysis is not consistent with those results found in recent studies which suggest a positive role of institutional quality on growth (see for example Acemoglu et al., 2001; Rodrik et al., 2004; Mehlum et al., 2006; Brunnschweiler, 2007; Azman-Saini et al., 2010; Acemoglu et al., 2014). Moreover, conversely to FE estimations, the evidence of an indirect role of institutional quality on the relationship between marginal taxation and growth is not long-established in Table 4. Considering the other controls, we observe that as before our investments coefficients are positive and significant in all specifications. Results are also confirmed for our tax decentralization variable in Model 4 where coefficient is negative and significant.

Overall, the p-values of the coefficient on our main taxation indicator display no-statistically significance in both FE and sys-GMM analysis. These values would imply that marginal taxation does not contribute to explain the differences in the economic performance across our 26 OECD countries. Therefore, our analysis seem to be in line with the empirical literature sustaining that marginal taxation does not

affect economic growth, and that, given this evidence of no-significant relationship, marginal taxation cuts do not bring about enhancements in growth levels.

6.3 Robustness checks

In this section, we consider a series of robustness checks that address the validity of the reported results. To measure robustness, several checks are used: treating alternative indicators of institutional quality (Tables 5 to 6); reconstructing our institutions dummy by quartiles (Tables 7 to 8); reselecting different time period (Tables 9 to 10); and re-running FE and sys-GMM estimations using an alternative country sample, which consists of the 18 OECD countries selected in PSS (Tables 11 to 12).

[TABLES 5 TO 12 ABOUT HERE]

Firstly, an alternative proxy of institutional quality, based on data for information and accountability transparency by Williams Andrews (2014), has been exploited. Dataset covers all countries but diverges for time period: 1980-2010. These specification checks do not find evidence that the estimates are biased on FE analysis. Table 5 confirms the existence of a different effect of marginal taxation on growth when quality of institutions varies. This result, however, does not still find confirmation from sys-GMM regressions.

Yet, we test results by using an alternative time-varying institutions dummy. Quality of institutions measure, based on ICRG index, is divided in two quartiles, and then each quartile is interacted with marginal tax rate. Values of our proxy of institutional quality below 0.85 belong to first quartile; and remaining higher values belong to second quartile. Results reported in table 3 do not hold when time-variant construction of institutions dummy is applied. Furthermore, specifications in tables 4 and 5 are regressed using a different time period from 1975 to 2015. FE estimation results are confirmed.

Finally, we re-estimate our FE and sys-GMM models using an alternative country sample, which includes the same 18 OECD countries selected in Piketty et al. (2014).

In specifications 1 through 3 in Table 11, coefficients of marginal tax rate are still insignificant, as also PSS found in their more robust estimation. In Models 2 through 3, the coefficients of our measure of institutional quality are positive and highly significant. In column 5, the evidence of a different effect of marginal tax rates on growth according to institutional quality still holds, with a negative effect in *bad institutions*, and a positive but negligible one in *good institutions*. As before, sys-GMM estimations do not confirm that the effect of marginal taxation on growth changes when quality of institutions varies. Differently from previous evidence, Table 12 shows a positive and significant coefficient of our marginal taxation variable (Model 1); though, this result disappears in remaining specifications (Model 2 to 3).

7. Conclusions

The recent economic crisis has renewed interest in policies protecting macroeconomic performance and promoting a long-term growth. On the other, given the increase in inequality among many OECD countries, the spotlight is also on income inequality. Many recommended tax reforms are aimed at increasing the tax progressivity through higher marginal tax rates (PSS, 2014). However, empirical evidence is not clear-cut on the effect of marginal taxation on long-term economic growth.

This work tried to shed some light on the impact of marginal tax rate on economic growth by using a panel data set of 26 OECD countries over the period 1981 - 2015 once controlling for a well-established set of growth determinants. Far from being detrimental to growth, at first look, our analysis has suggested that marginal taxation has an insignificant effect on growth. This result has found confirmation when we have used different estimation strategy and specifications.

In particular, this no-significant evidence would be consistent with rent-extraction theory and PSS' statements suggesting that upper incomes concentration comes mostly from redistribution in favour of top income earners due to the bargaining elasticity rather than an actual increase in economic activity. The idea is that low marginal rates ensure great gains from more aggressive bargaining effort, and then represent an incentive to practice rent-seeking by top earners. In contrast, relative higher marginal taxation would work as rent-extraction limitation due to small gains.

No-significant correlation between marginal tax rates and economic growth is also in line with the arguments by Chetty (2009) who sustains that an important part of the elasticity of taxable income could be explained by sheltering behaviors and that omitting the equivalence between marginal social costs of sheltering and tax rate causes an overestimation of deadweight losses.

However, an enriched picture came from investigating whether the effect of marginal taxation on growth varies according to the quality of institutions. From FE analysis it is emerged that in countries with institutional quality lower than sample average a reduction of marginal tax rate is growth-enhancing, as predicted by standard supply-side theory, but this distortionary effect of marginal taxation becomes negligible, or even marginally positive, in *good institutions*. This evidence is not robust to methodological changes, though.

Following the recent literature, an open question concerns the impact of marginal tax rate on current upper income distribution. In particular, the evidence of no-significant correlation between marginal taxation and growth that we have found might suggest that a cut of marginal taxation does not determine a higher productive effort among top earners which would explain the surge in upper incomes. At the same time, our analysis do not imply that a cut of marginal rates has caused the huge increase in top 1% income share due to a more aggressive bargaining effort acted by top earners. Indeed, considering the different picture emerging from our investigation into the impact of marginal taxation on growth once taking into account the quality of institutions (even if not robust to methodological changes) the impact of marginal tax rates on economic performance does not appear very clear.

In conclusion, our analysis has not revealed clear evidence about the impact of marginal tax rate on economic growth. Nevertheless, our results certainly suggest that quality of institutions represents a key determinant in explaining economic growth, and that its interrelations with fiscal policies and on economic growth deserves serious consideration.

References

Abadie A., Gardeazabal J. (2003) The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*. 93: 113-132.

Acemoglu D., Gallego F.A., Robinson J.A. (2014) Institutions, Human Capital, and Development. *Annual Review of Economics*. 6: 875-912.

Acemoglu D., Johnson S., Robinson J.A. (2001) The colonial origins of comparative development: an empirical investigation. *American Economic Review*, American Economic Association, 91: 1369-1401.

Acemoglu D., Johnson S., Robinson J.A. (2002) Reversal of fortune: geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, 117: 1231-1294.

Acemoglu D., Johnson S., Robinson J., Thaicharoen Y. (2003) Institutional causes, macroeconomic symptoms: volatility, crises and growth. *Journal of Monetary Economics*. 50: 49-123.

Acemoglu D., Johnson S., Robinson J.A., Yared P. (2008) Income and democracy. *American Economic Review*. 98: 808–842.

Agbor J.A. (2015) How does colonial origin matter for economic performance in sub-Saharan Africa? UNU-WIDER Working Paper.

Agell J., Ohlsson H., and Thoursie P.S. (2006). Growth effects of government expenditure and taxation in rich countries: A comment. *European Economic Review* 50: 211-8.

Agostino M., Nifo A., Trivieri F., Vecchione G. (2016) Total factor productivity heterogeneity: channeling the impact of institutions. MPRA Working Paper.

Aisen A., Veiga F.J. (2013) How does political instability affect economic growth? *European Journal of Political Economy*. 29: 151-167.

Akai N., Sakata M. (2002) Fiscal decentralization contributes to economic growth: evidence from state-level cross-section data for the United States. *Journal of Urban Economics*. 52: 93-108.

Alesina A., Perotti R. (1993) The Political Economy of Growth: A Critical Survey of the Recent Literature. *The World Bank Economic Review*, 8: 351-371.

Alesina A., Perotti R. (1996) Income Distribution, Political Instability, and Investment. *European Economic Review*, 40: 1203-1228.

Alesina A., Rodrik D. (1994) Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109: 465-490.

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

Anderson T.W., Hsiao C. (1981) Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76: 589-606.

Angelopoulos K., Economides G., Kammas P. (2007) Tax-spending policies and economic growth: Theoretical predictions and evidence from the OECD. *European Journal of Political Economy*. 23: 885-902.

Arachi G., Bucci V., Casarico A. (2015) Tax structure and macroeconomic performance. *International Tax and Public Finance*, 22: 635-662.

Arnold J. (2008) Do Tax Structures Affect Aggregate Economic Growth?: Empirical Evidence from a Panel of OECD Countries. OECD Economics Department Working Papers.

Arellano M., Bond S. (1991) Some tests of specification for panel data: monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58: 277-297.

Arellano M., Bover O. (1995) Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68: 29-51.

Arrow K.J. (1962) The Economic Implications of Learning by Doing. *The Review of Economic Studies*. 29: 155-173.

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

Autor D.H., Katz L.F., Krueger A.B. (1998) Computing inequality: have computers changed the labor market?. *The Quarterly Journal of Economics*, Oxford University Press, 113: 1169-1213.

Autor D.H., Katz L.F. (1999) Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, 3: 1463-1555.

Autor D.H., Levy F., Murnane R.J. (2003) The skill content of recent technological change: an empirical exploration. *The Quarterly Journal of Economics*, Oxford University Press, 118: 1279-1333.

Autor D.H., Katz L., Kearney M. (2006) The polarization of the U.S. labor market. *American Economic Review Papers and Proceedings*, 96: 189-194.

Azman-Saini W.N.W., Baharumshah A.Z., Law. S.H. (2010) Foreign direct investment, economic freedom and economic growth: International evidence. *Economic Modelling*. 27: 1079-1089.

Bakija J., Cole A., Heim, B. (2012) Jobs and income growth of top earners and the causes of changing income inequality: evidence from U.S. tax return data. *Working Paper*, Williams College.

Bakija J. (2013) Tax policy and philanthropy: a primer on the evidence for the U.S. and its implications. *Working Paper*, Williams College.

Banerjee A., Iyer L. (2005) History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *The American Economic Review*. 95: 1190-1213.

Barro R.J. (1990) Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98: 103-117.

Barro R.J. (1991) Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106: 407-443.

Barro R.J. (1997) *The Determinants of Economic Growth: A Cross-Country Empirical Study*. MIT Press, Cambridge.

Barro R.J. (1999) Determinants of democracy. *Journal of Political Economy*. 107: S158-S183.

Barro R. (2000) Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5: 5-32.

Barro R. (2013) Education and Economic Growth. *Annals of Economics and Finance*. 14: 301-328.

Barro R.J., Redlick C.J. (2011) The Macroeconomic Effects of Government Purchases and Taxes. *Quarterly Journal of Economics*. 126: 51–102.

Baskaran T., Feld L.P. (2013) Fiscal decentralization and economic growth in OECD countries: is there a relationship? *Public Finance Review*. 41: 421-445.

Bell B., Van Reenen J. (2014) Bankers and their bonuses. *Economic Journal*, 124: F1-F21.

Bénabou R. (1996) Inequality and Growth. NBER Chapters, in: *Macroeconomics Annual*, 11: 11-74.

Bénabou R. (2000) Unequal Societies: Income Distribution and the Social Contract. *The American Economic Review*, 90: 96-129.

Berg A., Henrekson M. (2011) Government size and growth: a survey and interpretation of the evidence. *Journal of Economic Survey*. 25: 872-897.

Berg A., Ostry J.D. (2011) Inequality and Unsustainable Growth: Two Sides of the Same Coin?. IMF Staff Discussion Note SDN/11/08.

Bergh A., Karlsson M. (2010) Government size and growth: Accounting for economic freedom and globalization. *Public Choice*. 142: 195-213.

Bleaney M.F., Gemmell N., Kneller R. (2001) Testing the endogenous growth model: Public expenditure, taxation and growth over the long run. *Canadian Journal of Economics*, 34: 383-403.

Blundell R., Bond. S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87: 115-143.

Bond S.R., Hoeffler A., Temple J. (2001) GMM estimation of empirical growth models. CEPR Discussion Papers.

Borjas G.J., Ramey V.A. (1995) Foreign competition, market power, and wage inequality. *The Quarterly Journal of Economics*, 110: 1075-1110.

Boschini A., Pettersson J., Roine J. (2007) Resource curse or not: A question of appropriability. *Scandinavian Journal of Economics*. 109: 593–617.

Boschini A., Pettersson J., Roine J. (2013) The Resource Curse and its Potential Reversal. *World Development*. 43: 19-41.

Brueckner J.K. (2006) Fiscal federalism and economic growth. *Journal of Public Economics*. 90: 2107-2120.

Brunnschweiler C.N. (2007) Cursing the Blessings? Natural Resource Abundance, Institutions, and Economic Growth. *World Development*. 36: 399-419.

Busse M., Hefeker C. (2007) Political risk, institutions and foreign direct investment. *European Journal of Political Economy*. 23: 397-415.

Campbell D., Lusher L. (2015) *The rise in inequality: technology and trade, or history and taxes?. Working Paper*.

Case A.C., Harvey S. R. (1993) Budget spillovers and fiscal policy interdependence. *Journal of Public Economics*. 52: 285-307.

Caselli F., Esquivel G., Lefort F. (1996) Reopening the convergence debate: a new look at cross-country growth empirics. *Journal of Economic Growth*, 1: 363-89.

Cass D. (1956) Optimum growth in an aggregative model of capital accumulation. *The Review of Economic Studies*. 32: 233-240.

Chetty Raj (2009) Is the taxable income elasticity sufficient to calculate deadweight loss? the implications of evasion and avoidance. *American Economic Journal: Economic Policy*, 1: 31-52.

Cingano F. (2014) Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment and Migration Working Papers No 163, OECD Publishing, Organization for Economic Co-operation and Development, Paris.

Dahlby B., Ferede E. (2012) The effects of tax rate changes on tax bases and the marginal cost of public funds for Canadian provincial governments. *International Tax and Public Finance*, 19: 844-883.

Davoodi H., Zou H. (1998) Fiscal Decentralization and Economic Growth: A Cross-Country Study. *Journal of Urban Economics*. 43: 244-257.

Deininger K., Squire L. (1998) New ways of looking at old issues: inequality and growth. *Journal of Development Economics*, 57: 259-287.

Devereux M.P., Lockwood B., Redoano M. (2008) Do Countries Compete over Tax Rates? *Journal of Public Economics*, 92: 1210–35.

Di Liberto A., Sideri M. (2015) Past Dominations, Current Institutions and the Italian Regional Economic Performance. IZA Discussion Paper.

Diamond P.A. (1998) Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *The American Economic Review*. 88: 83-95.

Dollar D., Kraay A. (2003) Institutions, trade, and growth. *Journal of Monetary Economics*. 50: 133-162.

Durlauf S.N., Johnson P.A., Temple J.R.W. (2005) Growth econometrics. *Handbook of Economic Growth*, 1: 555-677.

Easterly W., Rebelo S. (1993) Marginal income tax rates and economic growth in developing countries. *European Economic Review*. 37: 409-417.

Esfahani H.S., Ramírez M.a.T. (2003) Institutions, infrastructure, and economic growth. *Journal of Development Economics*. 70: 443–477.

Fafchamps M., McKenzie D., Quinn S., Woodruff C. (2014) Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in Ghana. *Journal of Development Economics*. 196: 211-226.

Feenberg D.R., Poterba J.M. (1993) Income inequality and the incomes of very high-income taxpayers: evidence from tax returns. NBER Chapters, in: *Tax Policy and the Economy*, 7: 145-177.

Feenstra R., Hanson G. (1999) The impact of outsourcing and high technological capital on wages: estimates for the United States 1979-1990. *Quarterly Journal of Economics*, 114: 907-940.

Feenstra R.C., Inklaar R., Timmer M.P. (2015) The next generation of the Penn World Table. forthcoming in *American Economic Review*.

Feldstein M. (1995) The effect of marginal tax rates on taxable income: a panel study of the 1986 tax reform act. *Journal of Political Economy*, 103: 551-572.

Feldstein M. (2006) The effect of taxes on efficiency and growth. *NBER Working Paper*.

Fisher R. (1982) Income and Grant Effects on Local Expenditures: The Flypaper Effect and Other Difficulties. *Journal of Urban Economics*. 17: 51-70.

Forbes K.J. (2000) A reassessment of the relationship between inequality and growth. *American Economic Review*, 90: 869-887.

Frank M.W. (2009) Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures. *Economic Inquiry*, 47: 55-68.

Gale W.G., Krupkin A., Rueben K. (2015) The relationship between taxes and growth at the state level: new evidence. *National Tax Journal*, 68: 919-942.

Galor O., Moav O. (2004) From Physical to Human Capital Accumulation: Inequality and the Process of Development. *Review of Economic Studies*, 71: 1001-1026.

Galor O., Zeira J. (1993) Income Distribution and Macroeconomics. *Review of Economic Studies*, 60: 35-52.

Gemmell N., Kneller R., Sanz I. (2011) The timing and persistence of fiscal policy impacts on growth: Evidence from OECD countries. *Economic Journal*, 121: F33-F58.

Gemmell N., Kneller R., Sanz I. (2013) Fiscal decentralization and economic growth: spending versus revenue decentralization. *Economic Inquiry*. 51: 1915-1931.

Gemmell N., Kneller R., Sanz I. (2014) The growth effects of tax rates in the OECD. *Canadian Journal of Economics*, 47: 1217-1255.

Gerring J., Bond P., Barndt W.T., Moreno C. (2005) Democracy and Economic Growth: A Historical Perspective. *World Politics*. 57: 323-364.

Glaeser E.L., Shleifer A. (2002) Legal origins. *The Quarterly Journal of Economics*. 117: 1193-1229.

Gordon R.J., Dew-Becker I.L. (2005) Where did productivity growth go? inflation dynamics and the distribution of income. *NBER Working Paper*.

Gordon R.H., Slemrod J. (2000) Are "Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?. In Slemrod J. (2000) *Does atlas shrug:*

the economic consequences of taxing the rich, Russel Sage Foundation, Harvard University Press. Also available as NBER Working Paper.

Gramlich M., Galper H. (1973) State and local fiscal behavior and federal grant policy. *Brookings Papers Econ. Activity*. 1: 15-58.

Gwartney J., Lawson R., Hall J., (2015) Economic freedom dataset. *Economic Freedom of the World: 2015 Annual Report*.

Hall R.E., Jones C.I. (1999) Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics*. 114: 83-116.

Harju J., Matikka T. (2014) The elasticity of taxable income and income-shifting: what is "real" and what is not?. *CESifo Working Paper*.

Hines J., Thaler R. (1995) Anomalies: The Flypaper Effect. *Journal of Economic Perspectives*. 9: 217-226.

Holtz-Eakin D., Newey W., Rosen H.S. (1988) Estimating vector autoregressions with panel data. *Econometrica*, 56: 1371-1395.

Inman R.P. (2008) The Flypaper effect. NBER Working Paper.

Islam N. (1995) Growth empirics: a panel data approach. *The Quarterly of Economics*, 110: 1127-1170.

Jin J., Zou H. (2005) Fiscal decentralization, revenue and expenditure assignments, and growth in China. *Journal of Asian Economics*. 16: 1047-1064.

Kaldor N. (1956) Alternative Theories of Distribution. *The Review of Economic Studies*, 23: 83-100.

Kaplan S.N. (2012) Executive compensation and corporate governance in the U.S.: perceptions, facts and challenges. *NBER Working Paper*.

Kaufmann D., Kraay A., Mastruzzi M. (2010) The Worldwide Governance Indicators: Methodology and Analitical Issues. World Bank Policy Reseacrh Working Paper.

Knack S., Keefer P. (1995) Institutions and economic performance: cross-country tests using alternative institutional measures. *Economics and Politics*. 7: 207-227.

Kneller R., Bleaney M. F., Gemmell N. (1999) Fiscal policy and growth: evidence from OECD countries. *Journal of Public Economics*, 74: 171-190.

Knight B. (2002) Endogenous Federal Grants and Crowd-out of State Government Spending: Theory and Evidence from the Federal Highway Aid Program. *The American Economic Review*. 92: 71-92.

Knowles S. (2005) Inequality and economic growth: the empirical relationship reconsidered in the light of comparable data. *The Journal of Development Studies*, 41: 135-159.

Koopmans T.C. (1965) On the concept of optimal economic growth. In: *The Econometric Approach to Development Planning*. Amsterdam: North-Holland.

La Porta R., Lopez-de-Silanes F., Shleifer A. (1998) Law and Finance. *Journal of Political Economy*. 106: 1113-1155.

La Porta R., Lopez-de-Silanes F., Shleifer A., Vishny R. (1999) The Quality of Government. *The Journal of Law, Economics, & Organization*. 15: 222-279.

La Porta R., Lopez-de-Silanes F., Shleifer A., Vishny R. (2000) Investor protection and corporate governance. *Journal of Financial Economics*. 58: 3-27.

La Porta R., Lopez-de-Silanes F., Shleifer A. (2008) The Economic Consequences of Legal Origins. *Journal of Economic Literature*. 46: 285-332.

Lee Y., Gordon R.H. (2005) Tax structure and economic growth. *Journal of Public Economics*, 89: 1027-1043.

Lee K., Kim B.-Y. (2009) Both Institutions and Policies Matter but Differently for Different Income Groups of Countries: Determinants of Long-Run Economic Growth Revised. *World Development*. 37: 533-549.

Levine R. (2005) Law, Endowments and Property Rights. *Journal of Economic Perspectives*. 19: 81-88.

Lin J.Y., Liu Z. (2000) Fiscal decentralization and economic growth in China. *Economic Development and Cultural Change*. 49: 1-21.

Lindsey L.B. (1987) Individual taxpayer response to tax cuts: 1982-84: with implications for the revenue maximizing tax rate. *Journal of Public Economics*, 33: 173-206.

Lucas R.E. (1988) On the mechanics of economic development. *Journal of Monetary Economics*. 22: 3-42.

Mahoney P.G. (2001) The common law and economic growth: Hayek might be right. *The Journal of Legal Studies*. 30: 503-525.

Mankiw G.N., Romer D., Weil D.N. (1992) A contribution to empirics of economic growth. *The Quarterly Journal of Economics*, 107: 407-437.

Martinez-Vazquez J. (2003) Fiscal Decentralization and Economic Growth. *World Development*. 31: 1597-1616.

Mauro P. (1995) Corruption and growth. *The Quarterly Journal of* Economics. 110: 681-712.

Mehlum H., Moene K., Ragnar T. (2006) Institutions and the Resource Curse. *The Economic Journal*. 116: 1-20.

Mendoza E.G., Razin A., Tesar L.L. (1994) Effective tax rates in Macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. Journal of Monetary Economics, 34: 297-323.

Mendoza E., Milesi-Ferretti G., Asea P. (1997) On the ineffectiveness of tax policy in altering long-run growth: Harberger's superneutrality hypothesis. *Journal of Monetary Economics*, 34: 297-323.

Mertens K. (2015) Marginal tax rates and income: new time series evidence. *NBER Working Paper.*

Mertens K., Ravn M. (2013) The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. *American Economic Review*. 103: 1212–1247.

Michalopoulos S.. Papaioannou E. (2013) Pre-colonial Ethnic Institutions and Contemporary African Development. *Econometrica*. 81: 113-152.

Milasi S. (2013) Top marginal taxation and economic growth. *Centre for Economic and International Studies*, University of Tor Vergata, 11.

Mulligan C.B., Gil R., Sala-i-Martin X. (2004) Do Democracies Have Different Public Policies than Nondemocracies? *Journal of Economic Perspectives*. 18: 51-74.

Myles D. (2007) Economic Growth and the Role of Taxation. Prepared for the OECD. Unpublished Report, University of Exeter and IFS.

Nickell S.J. (1981) Biases in dynamic models with fixed effects. *Econometrica*, 49: 1417-26.

North D.C. (1991) Institutions. *The Journal of Economic Perspectives*. 5: 97-112.

Oates W.E. (1972) *Fiscal Federalism*. New York: Harcout Brace Janovich.

Oates W.E. (1999) An Essay on Fiscal Federalism. *Journal of Economic Literature*. 37: 1120-1149.

Papaioannou E., Siourounis G. (2008) Democratisation and Growth. *The Economic Journal*. 118: 1520-1551.

Perotti R. (1996) Growth, income distribution, and democracy: what data say. *Journal of Economic Growth*. 1: 149-187.

Persson T., Tabellini G. (1994) Is Inequality Harmful for Growth? *American Economic Association*, 84: 600-621.

Persson T., Tabellini G. (2006) Democracy and Development. *American Economic Association*. 96: 319-324.

Piketty T., Saez E., Stantcheva S. (2014) Optimal taxation of top labor incomes: a tale of three elasticities. *American Economic Journal: Economic Policy*, 6: 230-271.

Piketty T., Saez E. (2007) How progressive is the U.S. federal tax system? a historical and international perspective. *Journal of Economic Perspectives*, 21: 3-24.

Reed W.R. (2008) The robust relationship between taxes and U.S. state income growth. *National Tax Journal*, 61: 57-80.

Reinikka R., Svensson J. (2004) Local capture: evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics*. 119: 679-705.

Rodrik D., Subramanian A., Trebbi F. (2004) Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth*. 9: 131-165.

Romer P.M. (1986) Increasing Returns and Long-Run Growth. *Journal of Political Economy*. 94: 1002-1037.

Romer P.M. (1990) Endogenous Technological Change. *Journal of Political Economy*. 98: S71-S102.

Romer C.D., Romer D.H. (2010) The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. *American Economic Review*. 100: 763–801.

Roodman D. (2009a) How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9: 86-136.

Roodman D. (2009b) A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, University of Oxford, 71: 135-158.

Rosen S. (1981) The economics of superstars. *American Economic Review*, 71: 845-58.

Saez E. (2001) Using Elasticities to Derive Optimal Income Tax Rates. *The Review of Economic Studies*. 68: 205-229.

Saez E., Slemrod J., Giertz S.H. (2012) The elasticity of taxable income with respect to marginal tax rates: a critical review. *Journal of Economic Literature*, 50: 3-50.

Seade J.K. On the shape of optimal tax schedules. *Journal of Public Economics*. 7: 203-235.

Slemrod J. (1996) High income families and the tax changes of the 1980s: the anatomy of behavioral response. NBER Chapters in: *Empirical Foundations of Household Taxation*, 169-92.

Slemrod J., Gillitzer C. (2014) Insights from a tax-system perspective. *CESifo Economic Studies*, 60: 1-31.

Solow R.M. (1956) A contribution to the theory of economic growth. *The Quarterly Journal of Economics*. 70:65-94.

Solt F. (2009) Standardizing the world income inequality database. *Social Science Quarterly*, 90: 231-242.

Tavares J., Wacziarg R. (2001) How democracy affects growth. *European Economic Review*. 45: 1341-1378.

Temple J. (1999) The new growth evidence. *Journal of Economic Literature*, 37: 112-156.

Teorell J., Dahlberg S., Holmberg S., Rothstein B., Khomenko A., Svensson R. (2016) The quality of government standard dataset. University of Gothenburg: The Quality of Government Institute, version Jan16.

Thiessen U. (2000) Fiscal federalism in Western European and selected other countries: centralization or decentralization? What is better for economic growth? DIW Discussion Paper.

Thornton J. (2007) Fiscal decentralization and economic growth reconsidered. *Journal of Urban Economics*. 61: 64-70.

Tiebout C.M. (1956) A pure theory of local expenditures. *Journal of political economy*. 64: 416-424.

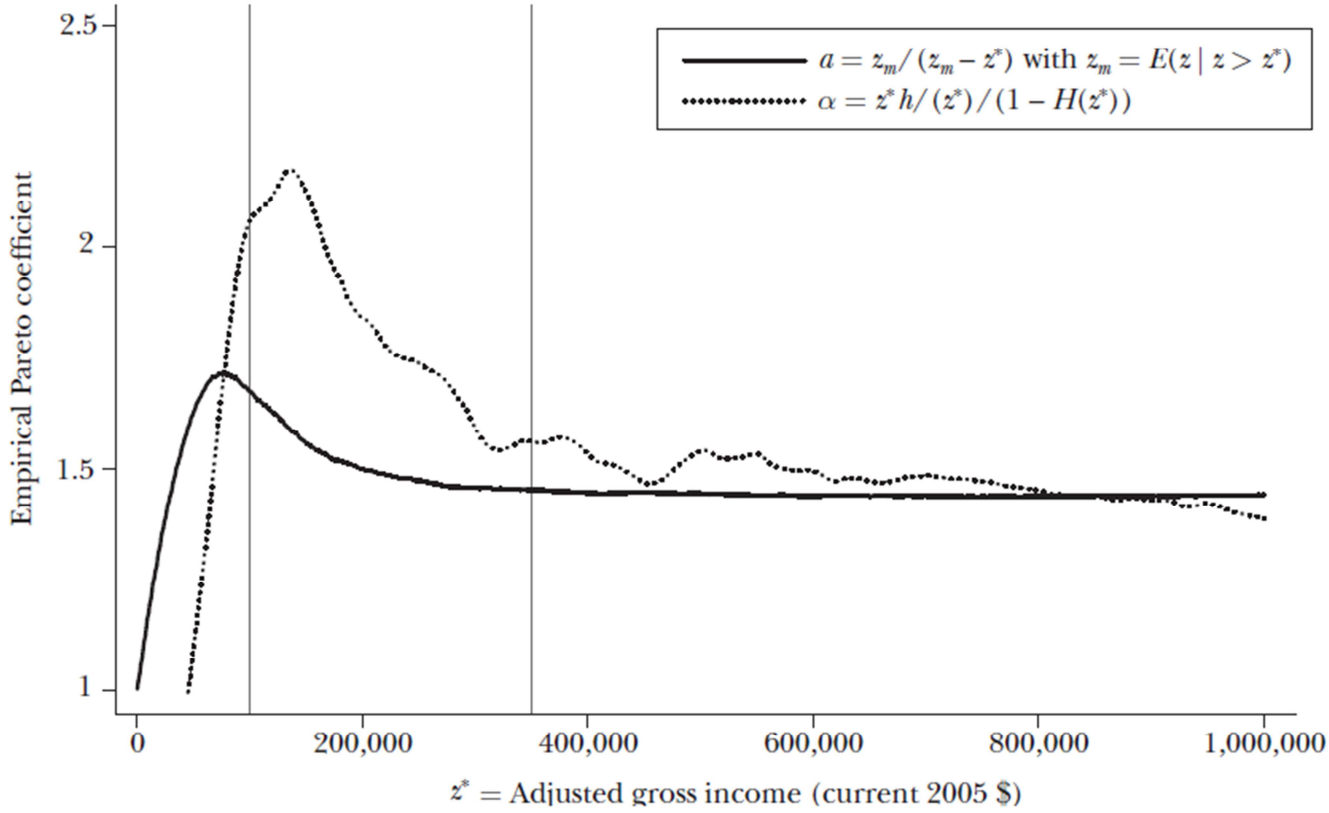
Tuomala M. (1990) *Optimal Income Tax and Redistribution.* London: Oxford University Press.

Turnbull G.K. (1992) Fiscal illusion, uncertainty, and the flypaper effect. *Journal of Public Economics*. 48: 207-223.

Xie D., Zou H., Davoodi H. (1999) Fiscal Decentralization and Economic Growth in the United States. *Journal of Urban Economics*. 45: 228-239.

Zhang T., Zou H. (1998) Fiscal decentralization, public spending, and economic growth in China. *Journal of Public Economics*. 67: 221-240.

Figure 1: US adjusted gross income distribution and Pareto parameter



Source: Diamond and Saez (2011).

Table 1: Statistics and source

Variables	Obs	Mean	Standard Deviation	Obs	Mean	Standard Deviation	Source
	Annual Data			*Fifthly Data*			
Real Gdp per capita	806	33771	12635	182	32533	12589	OECD Statistics
Marginal tax rate	783	0.46	0.11	181	0.48	0.12	OECD Statistics
Investments	754	22.99	3.79	182	23.01	3.48	World Bank (WDI)
Population growth rate	754	0.71	0.59	182	0.71	0.56	World Population Prospects 2017
Human capital	780	3.04	0.44	182	3.01	0.46	Penn World Table (PWT)
Institutional quality	806	0.85	0.15	182	0.86	0.15	QOG Database (ICRG)
Average tax rate	804	33.66	8.13	182	33.32	8.22	OECD Statistics
Tax decentralization	771	0.11	0.09	181	0.11	0.09	Elaborations on data from OECD statistics
Net inequality	646	29.91	17.59	177	29.98	5.90	Standardized World Inequality Indicators

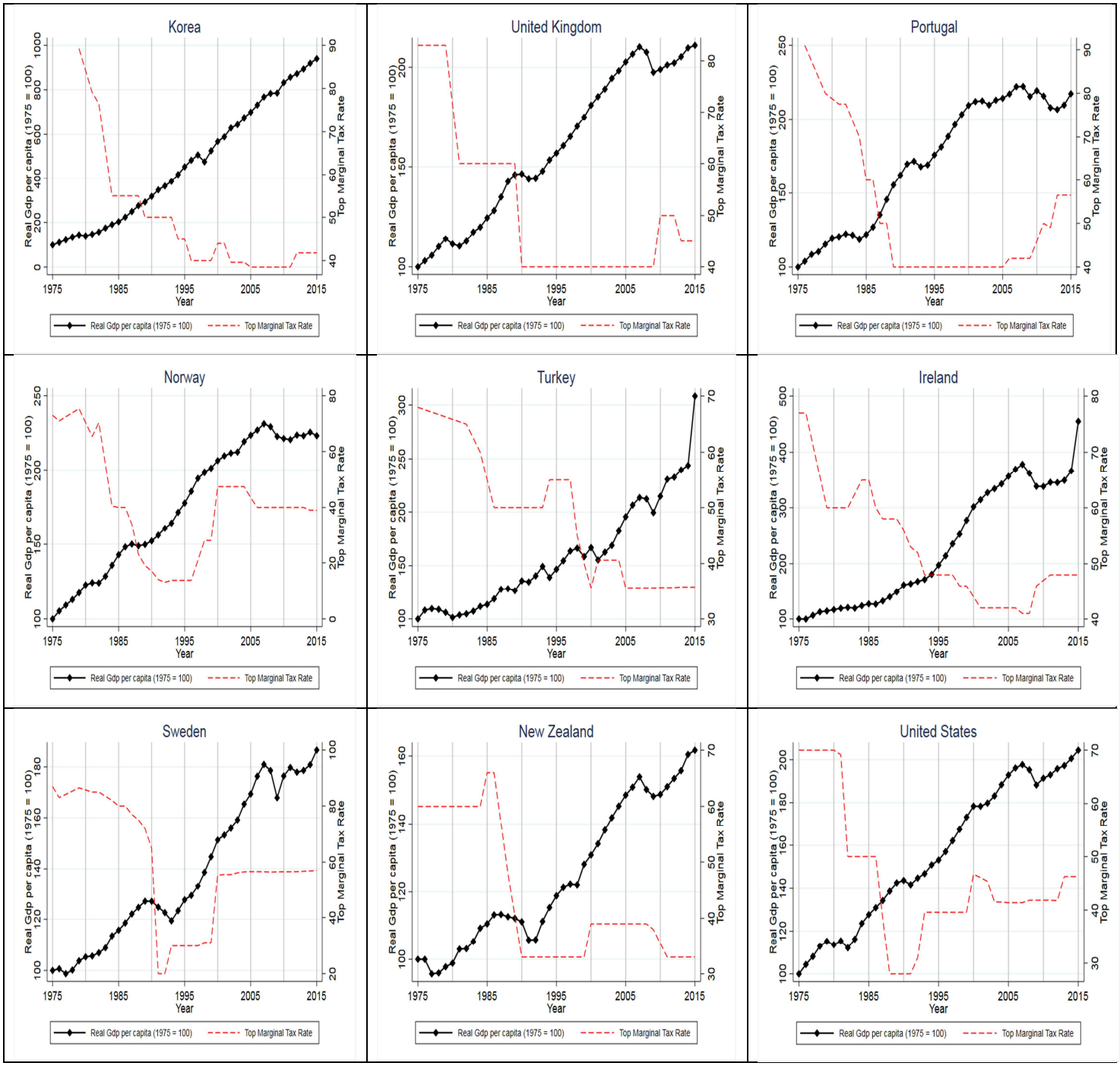
Sources: Our Elaborations.

Table 2: Variations in economic growth, marginal tax rates and quality of institutions

Country	Economic Growth Rate	Marginal Tax Rate	Marginal Tax Rate	Marginal Tax Rate	Quality of Institutions	Quality of Institutions
	1975-2015	*1975*	*1985*	*2015*	*1985*	*2015*
Australia	1.74	0.65	0.6	0.47	0.94	0.9
Austria	1.91	0.62	0.62	0.48	0.9	0.92
Belgium	1.71	0.6	0.74	0.54	1	0.88
Canada	1.53	0.47	0.45	0.49	1	0.92
Denmark	1.49	0.4	0.55	0.56	1	0.97
Finland	1.97	0.51	0.51	0.51	1	0.98
France	1.52	0.6	0.64	0.54	0.94	0.77
Germany	1.79	0.56	0.56	0.48	0.87	0.88
Greece	0.93	0.63	0.62	0.48	0.51	0.62
Iceland	2.05	0.38*	0.38	0.46	1	0.95
Ireland	3.45	0.77	0.63	0.48	0.79	0.89
Italy	1.44	0.72	0.69	0.49	0.75	0.57
Japan	1.97	0.75	0.83	0.52	0.89	0.86
Korea	5.84	0.89*	0.66	0.41	0.55	0.69
Luxembourg	2.51	0.57	0.58	0.43	1	0.94
Mexico	1.3	0.53*	0.55	0.32	0.51	0.45
Netherlands	1.65	0.71	0.72	0.52	1	0.94
New Zealand	1.17	0.6	0.61	0.33	1	0.95
Norway	2.2	0.73	0.54	0.4	0.98	0.97
Portugal	1.84	0.86*	0.71	0.54	0.65	0.74
Spain	1.63	0.62	0.67	0.5	0.71	0.74
Sweden	1.61	0.87	0.83	0.57	1	0.97
Switzerland	1.98	0.44	0.26	0.42	1	0.88
Turkey	2.46	0.68	0.6	0.36	0.5	0.49
United Kingdom	1.82	0.83	0.6	0.47	0.96	0.86
United States	1.73	0.7	0.54	0.45	0.96	0.83
Sample average	1.9073	0.63	0.6	0.47	0.87	0.83

Notes: The asterisks indicate that data are from either 1980 or 1985 instead of 1975 because of missing value. Table 2 is elaborated by using fifthly average observations.
Source: Our elaboration.

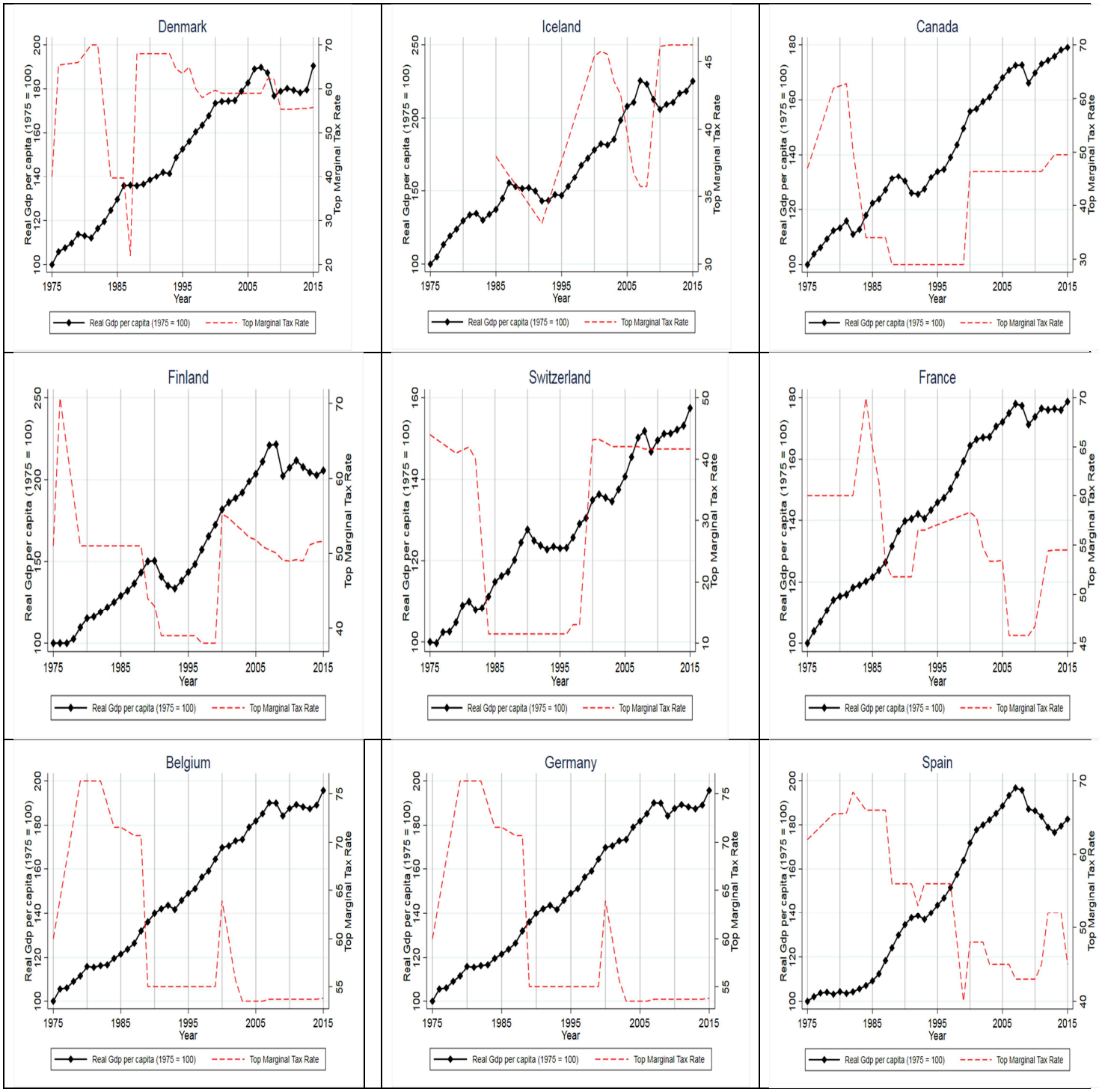
Figure 2.a: Real Gdp per capita and marginal tax rate by countries



Notes: Figure 2.a illustrates the relationship between real Gdp per capita and marginal tax rates for the nine countries which have experienced the largest top tax rate cut over the period 1975 – 2015. This figure is elaborated by using annual data.

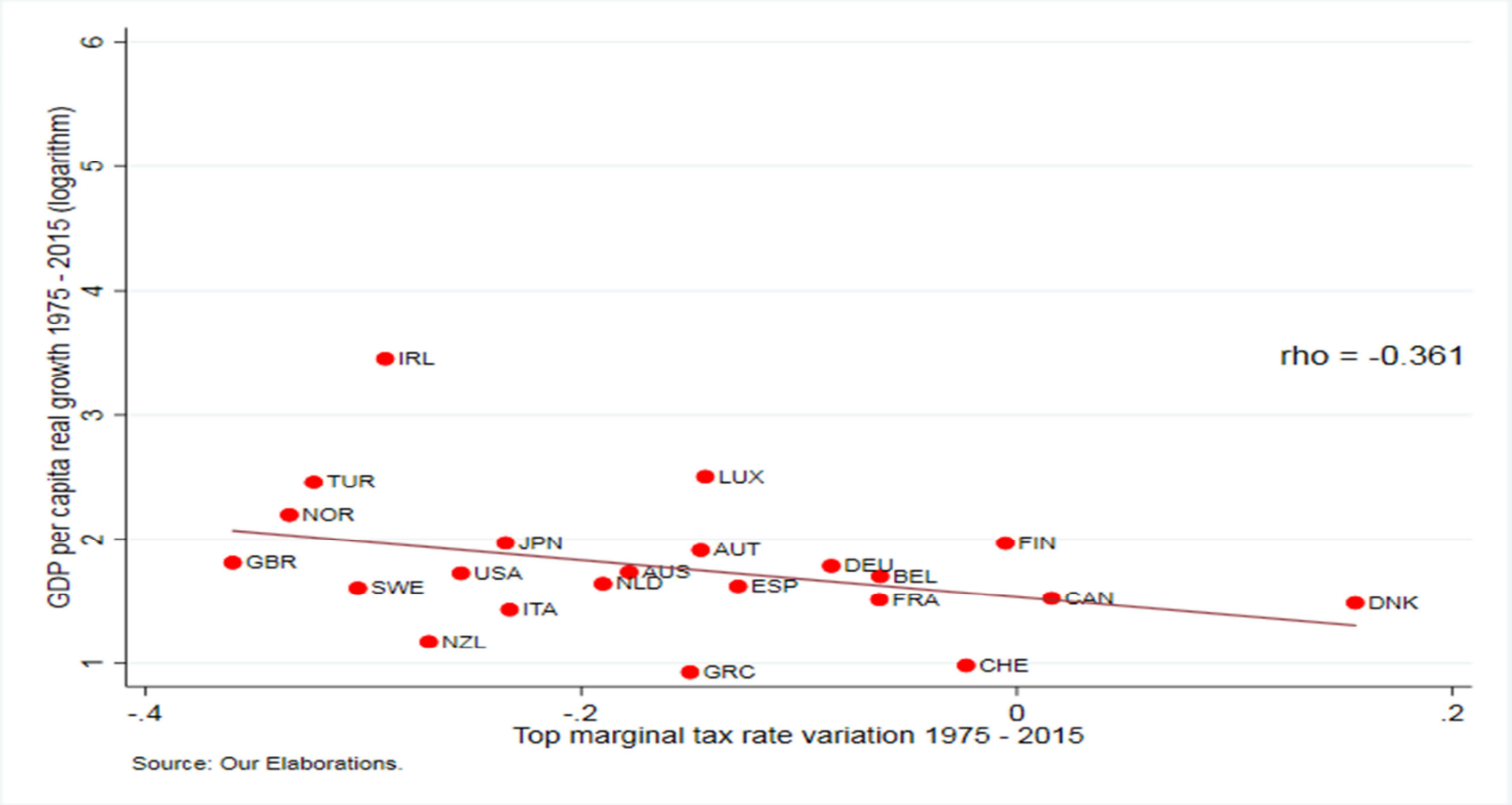
Source: Our elaboration.

Figure 2.a: Real Gdp per capita and marginal tax rate by countries



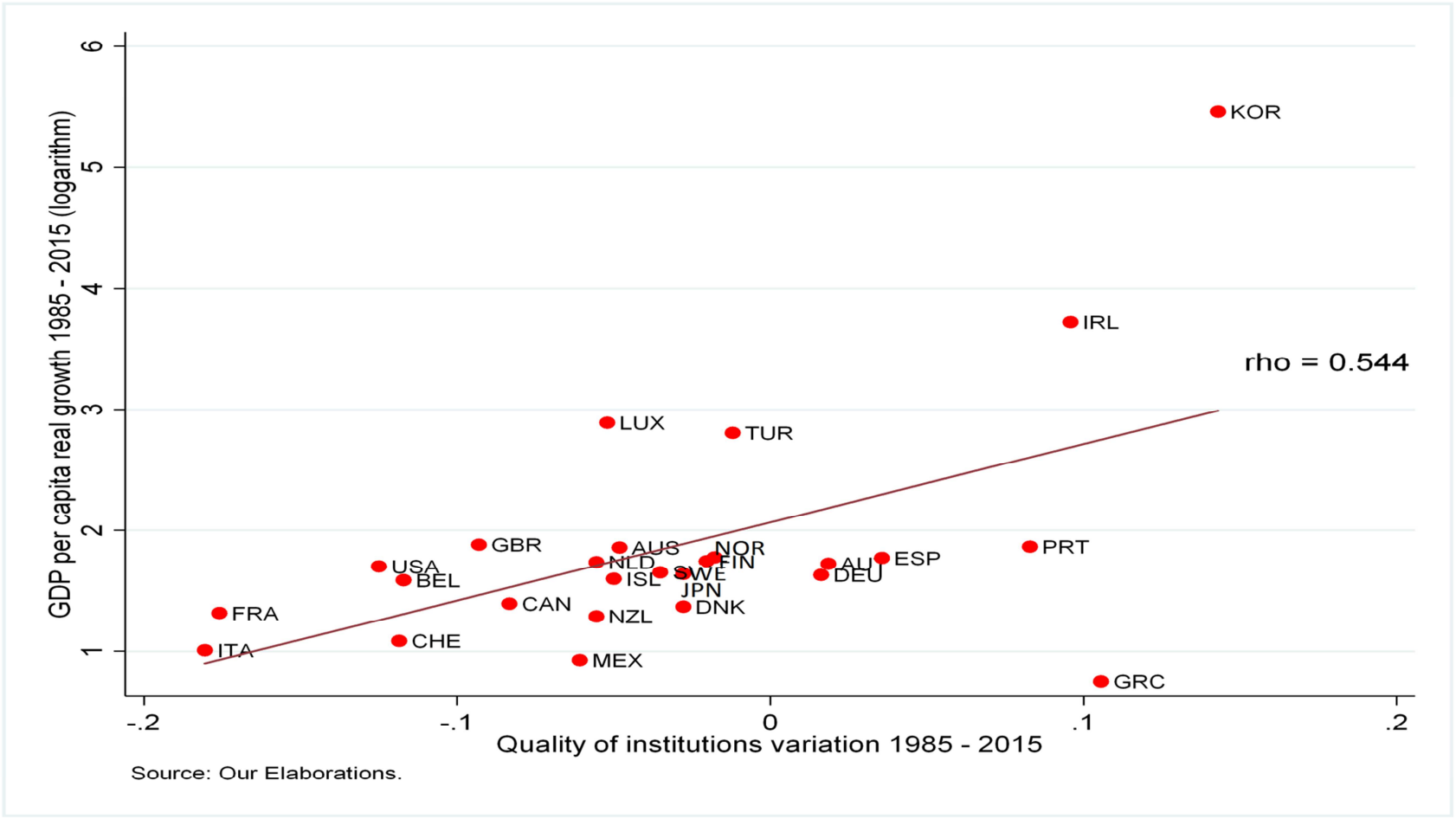
Notes: Figure 2.a illustrates the relationship between real Gdp per capita and marginal tax rates for the nine countries which have experienced the lowest top tax rate cut over the period 1975 – 2015. This figure is elaborated by using annual data.
Source: Our elaboration.

Figure 3: Gdp per capita growth and marginal tax rate



Notes: Figure 3 is elaborated by using five-year averaged data.
Source: Our elaboration.

Figure 4: Gdp per capita growth and quality of institutions



Notes: Figure 4 is elaborated by using five-year averaged data.
Source: Our elaboration.

Table 3: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015					
Variables	(Fe) (Model 1)	(Fe) (Model 2)	(Fe) (Model 3)	(Fe) (Model 4)	(Fe) (Model 5)
Marginal Taxation	-0.0021	-0.0035	-0.0101	-0.1434*	-0.1731**
	[0.0179]	[0.0184]	[0.0186]	[0.0783]	[0.0764]
Lagged Depndent Variable	0.7672***	0.7371***	0.7047***	0.8089***	0.7569***
	[0.0545]	[0.0445]	[0.0261]	[0.0473]	[0.0394]
Investments	0.2480***	0.2347***	0.2531***	0.1958***	0.1800***
	[0.0320]	[0.0362]	[0.0312]	[0.0333]	[0.0335]
Population Growth Rate	-0.005	0.0028	0.0044	-0.0005	0.002
	[0.0123]	[0.0121]	[0.0135]	[0.0176]	[0.0235]
Human Capital	-0.0052	-0.0374	-0.181	0.1361	0.2006
	[0.1753]	[0.1718]	[0.1191]	[0.1952]	[0.2314]
Istitutional Quality		0.1585**	0.0767		
		[0.0672]	[0.0629]		
Average Tax Rate			0.1051	0.0002	0.0116
			[0.0630]	[0.0710]	[0.0664]
Tax Decentralization			-0.0284*	-0.0392**	-0.0520***
			[0.0140]	[0.0142]	[0.0094]
Net Inequality			-0.0332	0.0945	-0.3462***
			[0.0790]	[0.0778]	[0.1095]
Interaction marginal taxation * Dinst				0.1565*	0.2038**
				[0.0782]	[0.0770]
Interaction investments * Dinst					-0.0326
					[0.0564]
Interaction population growth rate * Dinst					0.0234
					[0.0311]
Interaction human capital * Dinst					0.0569
					[0.1537]
Interaction average tax rate * Dinst					-0.1146
					[0.1108]
Interaction tax decentralization * Dinst					0.0709***
					[0.0225]
Interaction net inequality * Dinst					0.6210***
					[0.1261]
Constant	1.7187***	2.1412***	2.2464***	0.8102*	1.7781***
	[0.4417]	[0.4089]	[0.2752]	[0.4639]	[0.4560]
F statistic	489.54	473.7	1,742.34	389.27	281,972.69
Adjusted R-squared	0.96	0.97	0.97	0.96	0.97
Observations	180	180	176	176	176
Time Effects	Yes	Yes	Yes	Yes	Yes
Number of country	26	26	26	26	26

* $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Table 4: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015					
Variables	(Gmm-System) (Model 1)	(Gmm-System) (Model 2)	(Gmm-System) (Model 3)	(Gmm-System) (Model 4)	(Gmm-System) (Model 5)
Marginal Taxation	0.2045	0.1258	0.2517	-0.5179	-0.5808
	[0.1797]	[0.2296]	[0.2383]	[0.5130]	[0.4828]
Lagged Dependent Variable	0.8322***	0.8308***	0.7459***	0.8939***	0.8900***
	[0.1354]	[0.1368]	[0.1351]	[0.1284]	[0.1381]
Investments	0.4388*	0.6279***	0.6619*	0.2574*	0.5436**
	[0.2446]	[0.1744]	[0.3910]	[0.1458]	[0.2533]
Population Growth Rate	0.0367	-0.0845	0.0326	-0.0793	-0.2513
	[0.0707]	[0.0746]	[0.0963]	[0.1237]	[0.2430]
Human Capital	0.5012	0.073	-0.189	-0.8683	0.0942
	[0.6257]	[0.9408]	[0.5676]	[1.1625]	[0.8343]
Istitutional Quality		0.3283	0.5386		
		[0.3328]	[0.3857]		
Average Tax Rate			0.2019	0.088	-0.1549
			[0.2403]	[0.4729]	[0.4246]
Tax Decentralization			0.0693	-0.0513*	-0.0478
			[0.0777]	[0.0305]	[0.1233]
Net Inequality			0.1939	0.136	-0.1649
			[0.5090]	[0.2853]	[0.2956]
Interaction marginal taxation * Dinst				0.7716	0.9037**
				[0.5716]	[0.4502]
Interaction investments * Dinst					-1.0359
					[0.7275]
Interaction population growth rate * Dinst					0.0379
					[0.2754]
Interaction human capital * Dinst					-1.0354
					[1.6918]
Interaction average tax rate * Dinst					-0.0576
					[1.3769]
Interaction tax decentralization * Dinst					0.2406
					[0.5828]
Interaction net inequality * Dinst					1.355
					[0.9572]
P-values Arellano-Bond test (AR1)	0.11	0.39	0.87	0.65	0.58
P-values Arellano-Bond test (AR2)	0.32	0.23	0.11	0.32	0.4
P-values Arellano-Bond test (AR3)	0.83	0.14	0.69	0.87	0.37
P-values Hansen test	0.66	0.87	0.74	0.49	0.61
Time Effects	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.97	0.96	0.93	0.90	0.88
Observations	180	180	176	176	176
Instruments	22	22	21	21	25
Number of country	26	26	26	26	26

* $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Appendix: Robustness checks

Table 5: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015					
Variables	(Fe) (Model 1)	(Fe) (Model 2)	(Fe) (Model 3)	(Fe) (Model 4)	(Fe) (Model 5)
Marginal Taxation	-0.0021	-0.0035	-0.0101	-0.1434*	-0.1731**
	[0.0179]	[0.0184]	[0.0186]	[0.0783]	[0.0764]
Lagged Dependent Variable	0.7672***	0.7371***	0.7047***	0.8089***	0.7569***
	[0.0545]	[0.0445]	[0.0261]	[0.0473]	[0.0394]
Investments	0.2480***	0.2347***	0.2531***	0.1958***	0.1800***
	[0.0320]	[0.0362]	[0.0312]	[0.0333]	[0.0335]
Population Growth Rate	-0.005	0.0028	0.0044	-0.0005	0.002
	[0.0123]	[0.0121]	[0.0135]	[0.0176]	[0.0235]
Human Capital	-0.0052	-0.0374	-0.181	0.1361	0.2006
	[0.1753]	[0.1718]	[0.1191]	[0.1952]	[0.2314]
Istitutional Quality		0.1585**	0.0767		
		[0.0672]	[0.0629]		
Average Tax Rate			0.1051	0.0002	0.0116
			[0.0630]	[0.0710]	[0.0664]
Tax Decentralization			-0.0284*	-0.0392**	-0.0520***
			[0.0140]	[0.0142]	[0.0094]
Net Inequality			-0.0332	0.0945	-0.3462***
			[0.0790]	[0.0778]	[0.1095]
Interaction marginal taxation * Dinst				0.1565*	0.2038**
				[0.0782]	[0.0770]
Interaction investments * Dinst					-0.0326
					[0.0564]
Interaction population growth rate * Dinst					0.0234
					[0.0311]
Interaction human capital * Dinst					0.0569
					[0.1537]
Interaction average tax rate * Dinst					-0.1146
					[0.1108]
Interaction tax decentralization * Dinst					0.0709***
					[0.0225]
Interaction net inequality * Dinst					0.6210***
					[0.1261]
Constant	1.7187***	2.1412***	2.2464***	0.8102*	1.7781***
	[0.4417]	[0.4089]	[0.2752]	[0.4639]	[0.4560]
F statistic	489.54	473.7	1,742.34	389.27	281,972.69
Adjusted R-squared	0.96	0.97	0.97	0.96	0.97
Observations	180	180	176	176	176
Time Effects	Yes	Yes	Yes	Yes	Yes
Number of country	26	26	26	26	26

Note: estimations performed by using an alternative indicator of institutional quality. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Table 6: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1975 - 2015					
Variables	(Gmm-System) (Model 1)	(Gmm-System) (Model 2)	(Gmm-System) (Model 3)	(Gmm-System) (Model 4)	(Gmm-System) (Model 5)
Marginal Taxation	0.2045	0.1258	0.2517	-0.5179	-0.5808
	[0.1797]	[0.2296]	[0.2383]	[0.5130]	[0.4828]
Lagged Dependent Variable	0.8322***	0.8308***	0.7459***	0.8939***	0.8900***
	[0.1354]	[0.1368]	[0.1351]	[0.1284]	[0.1381]
Investments	0.4388*	0.6279***	0.6619*	0.2574*	0.5436**
	[0.2446]	[0.1744]	[0.3910]	[0.1458]	[0.2533]
Population Growth Rate	0.0367	-0.0845	0.0326	-0.0793	-0.2513
	[0.0707]	[0.0746]	[0.0963]	[0.1237]	[0.2430]
Human Capital	0.5012	0.073	-0.189	-0.8683	0.0942
	[0.6257]	[0.9408]	[0.5676]	[1.1625]	[0.8343]
Istitutional Quality		0.3283	0.5386		
		[0.3328]	[0.3857]		
Average Tax Rate			0.2019	0.088	-0.1549
			[0.2403]	[0.4729]	[0.4246]
Tax Decentralization			0.0693	-0.0513*	-0.0478
			[0.0777]	[0.0305]	[0.1233]
Net Inequality			0.1939	0.136	-0.1649
			[0.5090]	[0.2853]	[0.2956]
Interaction marginal taxation * Dinst				0.7716	0.9037**
				[0.5716]	[0.4502]
Interaction investments * Dinst					-1.0359
					[0.7275]
Interaction population growth rate * Dinst					0.0379
					[0.2754]
Interaction human capital * Dinst					-1.0354
					[1.6918]
Interaction average tax rate * Dinst					-0.0576
					[1.3769]
Interaction tax decentralization * Dinst					0.2406
					[0.5828]
Interaction net inequality * Dinst					1.355
					[0.9572]
P-values Arellano-Bond test (AR1)	0.11	0.39	0.87	0.65	0.58
P-values Arellano-Bond test (AR2)	0.32	0.23	0.11	0.32	0.4
P-values Arellano-Bond test (AR3)	0.83	0.14	0.69	0.87	0.37
P-values Hansen test	0.66	0.87	0.74	0.49	0.61
Time Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.97	0.96	0.93	0.90	0.88
Observations	180	180	176	176	176
Instruments	22	22	21	21	25
Number of country	26	26	26	26	26

Note: estimations performed by using an alternative indicator of institutional quality. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Table 7: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015		
Variables	(Fe) (Model 4)	(Fe) (Model 5)
Marginal Taxation	-0.0002	0.0096
	[0.0202]	[0.0193]
Lagged Dependent Variable	0.7957***	0.7806***
	[0.0452]	[0.0433]
Investments	0.2162***	0.1856***
	[0.0311]	[0.0494]
Population Growth Rate	0.0027	0.0026
	[0.0172]	[0.0273]
Human Capital	0.1797	0.1378
	[0.1870]	[0.2050]
Average Tax Rate	0.0189	0.01
	[0.0707]	[0.0939]
Tax Decentralization	-0.0339**	-0.0128
	[0.0141]	[0.0148]
Net Inequality	0.0658	0.1825**
	[0.0786]	[0.0712]
Interaction marginal taxation * Dinst	-0.0945*	-0.1006
	[0.0549]	[0.0712]
Interaction investments * Dinst		0.0047
		[0.0536]
Interaction population growth rate * Dinst		0.0134
		[0.0314]
Interaction human capital * Dinst		0.2074
		[0.1533]
Interaction average tax rate * Dinst		-0.0436
		[0.0844]
Interaction tax decentralization * Dinst		-0.0267*
		[0.0150]
Interaction net inequality * Dinst		-0.2942**
		[0.1111]
Constant	0.9009*	0.9146*
	[0.4475]	[0.5154]
F statistic	418.2	477.82
Adjusted R-squared	0.96	0.96
Observations	176	176
Time Effects	Yes	Yes
Number of country	26	26

Note: estimations performed by constructing institutions dummy by quartiles. Only models 4 and 5 which include institutions dummy are estimated. * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in bracket.

Table 8: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015		
Variables	(Gmm-System) (Model 4)	(Gmm-System) (Model 5)
Marginal Taxation	0.1293	-0.1949
	[0.1187]	[0.4808]
Lagged Dependent Variable	0.8157***	0.8560***
	[0.1073]	[0.2800]
Investments	0.4349***	0.5155
	[0.0850]	[0.7060]
Population Growth Rate	-0.0587	0.1905
	[0.0682]	[0.2955]
Human Capital	0.146	-0.2776
	[0.6441]	[1.1808]
Average Tax Rate	0.1358	0.5198
	[0.3467]	[0.8562]
Tax Decentralization	-0.0422*	-0.2342
	[0.0251]	[0.3980]
Net Inequality	0.0032	-0.7557
	[0.1183]	[1.0051]
Interaction marginal taxation * Dinst	-0.4045**	0.1284
	[0.1616]	[0.5097]
Interaction investments * Dinst		0.0391
		[0.6134]
Interaction population growth rate * Dinst		-0.3092
		[0.3564]
Interaction human capital * Dinst		1.4608
		[1.6536]
Interaction average tax rate * Dinst		-0.8691
		[0.8219]
Interaction tax decentralization * Dinst		0.1502
		[0.3790]
Interaction net inequality * Dinst		0.828
		[1.5391]
P-values Arellano-Bond test (AR1)	0.87	0.4
P-values Arellano-Bond test (AR2)	0.33	0.21
P-values Arellano-Bond test (AR3)	0.95	0.66
P-values Hansen test	0.45	0.58
Time Effects	Yes	Yes
R-squared	0.97	0.91
Observations	176	176
Instruments	24	24
Number of country	26	26

Note: estimations performed by constructing institutions dummy by quartiles. Only models 4 and 5 which include institutions dummy are estimated. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Table 9: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015		
Variables	(Fe) (Model 4)	(Fe) (Model 5)
Marginal Taxation	-0.1236**	-0.2616***
	[0.0583]	[0.0767]
Lagged Dependent Variable	0.8738***	0.8605***
	[0.0422]	[0.0461]
Investments	0.1664***	0.2209***
	[0.0475]	[0.0649]
Population Growth Rate	-0.0013	-0.0454
	[0.0204]	[0.0328]
Human Capital	0.0193	-0.2658
	[0.1518]	[0.2233]
Average Tax Rate	0.0129	0.0559
	[0.0636]	[0.0720]
Tax Decentralization	-0.0410***	-0.0510***
	[0.0107]	[0.0123]
Net Inequality	0.0824	-0.0332
	[0.0515]	[0.0974]
Interaction marginal taxation * Dinst	0.1304**	0.2719***
	[0.0633]	[0.0786]
Interaction investments * Dinst		-0.11
		[0.0739]
Interaction population growth rate * Dinst		0.0655
		[0.0385]
Interaction human capital * Dinst		0.3589**
		[0.1425]
Interaction average tax rate * Dinst		0.0033
		[0.1015]
Interaction tax decentralization * Dinst		0.0718***
		[0.0165]
Interaction net inequality * Dinst		0.1388
		[0.1049]
Constant	0.3489	0.5598
	[0.4254]	[0.4654]
F statistic	827.99	994.46
Adjusted R-squared	0.97	0.98
Observations	213	213
Time Effects	Yes	Yes
Number of country	26	26

Note: estimations performed by using a different time period (1975 – 2015). Only models 4 and 5 which include institutions dummy are estimated because of missing information for proxy of institutional quality over the period 1975 – 1980. * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in bracket.

Table 10: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015		
Variables	(Gmm-System) (Model 4)	(Gmm-System) (Model 5)
Marginal Taxation	0.12	0.0695
	[0.2724]	[0.3087]
Lagged Dependent Variable	0.9536***	0.8092***
	[0.2351]	[0.1720]
Investments	0.3684**	0.7630*
	[0.1447]	[0.4160]
Population Growth Rate	-0.1012*	-0.2082
	[0.0601]	[0.2218]
Human Capital	0.087	0.8566
	[0.6137]	[1.2981]
Average Tax Rate	0.0074	-0.2687
	[0.4449]	[0.5663]
Tax Decentralization	0.1488	0.0807
	[0.1554]	[0.0873]
Net Inequality	-0.0644	0.0229
	[0.4060]	[0.2730]
Interaction marginal taxation * Dinst	0.1369	0.0917
	[0.2026]	[0.3702]
Interaction investments * Dinst		-0.7432
		[0.5505]
Interaction population growth rate * Dinst		0.0859
		[0.3963]
Interaction human capital * Dinst		-2.6145
		[2.2716]
Interaction average tax rate * Dinst		0.1555
		[1.6554]
Interaction tax decentralization * Dinst		-0.4023
		[0.3620]
Interaction net inequality * Dinst		0.7728
		[0.9146]
P-values Arellano-Bond test (AR1)	0.9	0.47
P-values Arellano-Bond test (AR2)	0.32	0.23
P-values Arellano-Bond test (AR3)	0.53	0.44
P-values Hansen test	0.89	1
Time Effects	Yes	Yes
R-squared	0.91	0.75
Observations	213	213
Instruments	25	27
Number of country	26	26

Note: estimations performed by using a different time period (1975 – 2015). Only models 4 and 5 which include institutions dummy are estimated because of missing information for proxy of institutional quality over the period 1975 – 1980. * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in bracket.

Table 11: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015					
Variables	(Fe) (Model 1)	(Fe) (Model 2)	(Fe) (Model 3)	(Fe) (Model 4)	(Fe) (Model 5)
Marginal Taxation	0.0228	0.0184	0.0195	-0.0661	-0.2146**
	[0.0143]	[0.0152]	[0.0159]	[0.0777]	[0.0849]
Lagged Dependent Variable	0.8683***	0.7792***	0.6948***	0.8822***	0.8396***
	[0.0757]	[0.0834]	[0.0615]	[0.0454]	[0.0640]
Investments	0.1944***	0.1802***	0.2024***	0.1625***	0.0592
	[0.0356]	[0.0321]	[0.0418]	[0.0420]	[0.0607]
Population Growth Rate	-0.0085	-0.0005	0.0036	-0.0048	0.0177
	[0.0169]	[0.0147]	[0.0144]	[0.0188]	[0.0149]
Human Capital	-0.3106**	-0.3164***	-0.1951	-0.1375	-0.2621
	[0.1416]	[0.1042]	[0.1553]	[0.1597]	[0.3047]
Istitutional Quality		0.3137***	0.2479**		
		[0.0847]	[0.0866]		
Average Tax Rate			-0.0748	-0.0714	0.1297
			[0.0800]	[0.1128]	[0.1852]
Tax Decentralization			-0.0017	-0.0213*	-0.0498***
			[0.0085]	[0.0121]	[0.0127]
Net Inequality			0.0226	0.1582**	-0.2386**
			[0.0603]	[0.0555]	[0.0976]
Interaction marginal taxation * Dinst				0.1012	0.2491***
				[0.0805]	[0.0856]
Interaction investments * Dinst					0.0869
					[0.0746]
Interaction population growth rate * Dinst					-0.0034
					[0.0293]
Interaction human capital * Dinst					0.2409
					[0.1542]
Interaction average tax rate * Dinst					-0.1842
					[0.1649]
Interaction tax decentralization * Dinst					0.0789**
					[0.0279]
Interaction net inequality * Dinst					0.4791***
					[0.1167]
Constant	1.2041	2.2392**	3.0939***	0.5971	1.0511
	[0.7092]	[0.8427]	[0.6373]	[0.6109]	[0.7801]
F statistic	604.5	407.63	6,421.41	1,895.43	.
Adjusted R-squared	0.97	0.97	0.97	0.97	0.97
Observations	126	126	123	123	123
Time Effects	Yes	Yes	Yes	Yes	Yes
Number of country	26	26	26	26	26

Note: estimations performed by using an alternative country sample. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Table 12: Effects of marginal taxation on economic growth

Dependent Variable: Real Gdp per capita Years: 1985 - 2015					
Variables	(Gmm-System) (Model 1)	(Gmm-System) (Model 2)	(Gmm-System) (Model 3)	(Gmm-System) (Model 4)	(Gmm-System) (Model 5)
Marginal Taxation	0.2722**	0.6182	0.0736	-0.1275	-0.4815
	[0.1299]	[0.5165]	[1.0807]	[2.4637]	[0.3965]
Lagged Dependent Variable	0.9142***	0.8612***	0.8575**	0.8512**	0.8188***
	[0.0520]	[0.1749]	[0.3528]	[0.3546]	[0.2596]
Investments	0.3320*	0.3779	0.7376	0.7989	-0.0076
	[0.1894]	[0.4380]	[1.2791]	[1.3402]	[0.5468]
Population Growth Rate	0.0011	0.171	-0.228	-0.2839	-0.0376
	[0.0831]	[0.2302]	[0.6643]	[0.3922]	[0.2658]
Human Capital	0.0647	0.4477	-0.9722	-2.0039	-0.4929
	[0.2136]	[0.8158]	[5.1069]	[13.7638]	[1.4785]
Istitutional Quality		-0.6807	0.1939		
		[1.0788]	[2.6120]		
Average Tax Rate			0.051	0.4146	0.4198
			[2.0001]	[3.4360]	[0.9103]
Tax Decentralization			0.6433	0.6119	0.1114
			[1.2614]	[2.0139]	[0.3478]
Net Inequality			0.5567	0.3886	0.2826
			[0.7369]	[1.2961]	[0.8003]
Interaction marginal taxation * Dinst				0.0969	0.5904
				[3.3148]	[0.4369]
Interaction investments * Dinst					0.3327
					[0.7433]
Interaction population growth rate * Dinst					0.0464
					[0.2864]
Interaction human capital * Dinst					-0.6877
					[1.5040]
Interaction average tax rate * Dinst					0.0824
					[1.3236]
Interaction tax decentralization * Dinst					0.0052
					[0.4821]
Interaction net inequality * Dinst					0.6906
					[0.8683]
P-values Arellano-Bond test (AR1)	0.85	0.93	0.51	0.67	0.82
P-values Arellano-Bond test (AR2)	0.67	0.15	0.86	0.93	0.46
P-values Arellano-Bond test (AR3)	0.57	0.32	0.94	1	0.5
P-values Hansen test	0.22	0.9	0.59	0.76	1
Time Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.94	0.68	0.41	0.35	0.83
Observations	126	126	123	123	123
Instruments	17	17	17	17	25
Number of country	26	26	26	26	26

Note: estimations performed by using an alternative country sample. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Robust standard errors in bracket.

Chapter 3

Female labour decision and joint taxation: evidence from Germany

1. Introduction

In recent years, there has been increased interest in the effect of gender labour discrimination against women on economic growth. In particular, a growing body of literature strongly suggests that gender inequality in education and employment not only disadvantage women considered, but has negative growth effects (Drèze and Sen, 1989; Pritchett and Summers, 1996; Klasen, 2000; 2002; 2006).

A number of contributions have reported that gender bias in education reduces the average amount of human capital in the society, and hence affects negatively economic growth (Galor and Weil, 1996; Dollar and Gatti, 1999; Knowles, Lorgelly, and Owen, 2002; Klasen, 2002; Abu-Ghaida and Klasen, 2004). Lagerlöf (2003) argues that reducing gender gaps in education lowers fertility rates and promotes the education of the next generation, and thus in turn enhances economic development.

On the other hand, some scholars have provided theoretical and empirical evidence that gender inequality in employment reduces economic growth as do gender gaps in education (Seguino, 2000; Klasen and Wink, 2002; Seguino and Floro, 2003; Esteve-Volart, 2004; Blackden, Canagarajah, Klasen, and Lawson, 2007; Cavalcanti and Tavares, 2007; Klasen and Lamanna, 2009). The results reported by Klasen and Lamanna (2009), who analyses the link between gender inequality in employment and growth when applying OLS cross-country and fixed effects panel estimations for 131 countries during the period 1960 to 2000, show that gender gaps considerably reduce economic growth.

Severe explanations for this negative link have been proposed by the literature. A first argument is proposed by Esteve-Volart (2004), who develops a theoretical model which captures gender discrimination as the exclusion of women from labor market or managerial positions. Her model predicts that gender inequality in employment causes lower growth rates because of a misallocation of talent and a

reduction of average productivity of human capital. In this study, she also provides empirical evidence supporting her theoretical insights. Yet, Seguino and Floro (2003) explain the negative link between gender inequality in employment and growth through the importance of female work participation for their household bargaining power. Since women and men have different propensities to save, with women saving rates higher than men ones (Hinz, McCarthy, and Turner, 1996; Bajtelsmit and Bernasek, 1996; Bajtelsmit and Van Derhei, 1997; Hungerford, 1999), and given that female participation in labour market increases their household bargaining power (Klasen and Wink, 2002), Seguino and Floro (2003) suggest that increasing in female employment, and consequent increasing in female bargaining power, is beneficial for entire society because changes in household saving rates have implications for aggregate saving rates. Using a panel data set of 20 semi-industrialised economies for the period 1975-1995, they find a positive link between women bargaining power and gross domestic saving rates.

Gender discrimination may take different forms. Many social, cultural or religious practices may leave women out of labour market. On the other hand, fiscal policies such as joint taxation might bring about this gender discrimination as well. Indeed, a large literature argues that joint taxation schedule discourages female participation in the labour market (Blundell, 1995). To the extent that joint taxation has implications for gender discrimination in employment by leaving women out of the market, a joint tax system may thus have negative growth consequences.

Many efforts have been put into the attempt to stimulating the entrance of women in the labour market. However, countries show substantial heterogeneity in the women labour force participation rates: in some countries the female labour force participation rate tends to converge to the men level; whereas in other countries this rate has not changed remarkably. Table 1 shows male and female labour force participation rates in 2016 for selected OECD countries according to estimations by the International Labour Organization (ILO). It is evident that the female labour force participation rates in Nordic countries are high: Iceland (69.2%), Norway (61.9%), Sweden (61.4%), and Denmark (57.3%). Other countries exhibit an intermediate level in female employment rates, like Austria (54.6%), Czech Republic (54.3%), France

(53.2%), Germany (55.4%), and Spain (54.3%). The lowest female labour force participation rate is reported by Italy (41.2%).

[TABLE 1 ABOUT HERE]

Differences in income tax regimes are proposed as an explanation of these divergent behaviors. Steiner and Wrohlich (2008 p.116) sustains that *"[ ] low labor-force participation rate of married women in Germany is closely related to the negative labor-supply incentives for secondary earners implied by the tax system."*

Many countries apply tax systems which treat secondary earners[30] in couples differently to single taxpayers (Feldstein and Feenberg, 1996; Smith, Dex, Vlasblom, and Callan, 2003). This is the case of so-called joint taxation systems, which allow joint tax filing for married couples. Joint taxation has the advantage of satisfying the rule of equal treatment of households with the same total income, namely the tax burden of the married couples depends exclusively on the total income, and does not depend on the distribution of income among the spouses. However, joint tax treatments have the important drawback of equalizing the marginal tax rates between primary and secondary earners – i.e. joint taxation leads to lower marginal tax rates for primary earner, and higher marginal tax rates for secondary earner. Since the second earner is often the female partner (OECD, 2012), when a married women enters in the labour market, the first euro of her earned income is taxed at her husband's current marginal rate. Therefore, this potentially affects the labour decisions of married women.

Given the wide and long-lasting interest in the relationship between women's labour decision and tax treatment of married couples, many empirical analyses have been devoted to understanding this connection. Several scholars have investigated whether the joint tax treatments are a disincentive for married women as suggested by economic theory (see e.g. Leuthold, 1984; Chiappori, 1988; Blundell, 1995; Crossley and Jeon, 2007; Kleven, Kreiner, and Saez, 2009; Gugl, 2009; Immervoll, Kleven, Kreiner, and Verdelin, 2011). Not surprisingly, the vast majority of these

[30] In a household the primary earner is the partner who earns the highest salary, while the secondary earner is the partner who earns the lowest salary of the household income.

studies show a negative correlation between joint taxation and female labour behaviour, and interestingly they find that policy reforms aimed at rising jointness of tax schedule affect negatively the female employment (see e.g. Boskin and Sheshinski, 1983; Blundell, 1995; Blundell, Duncan, and Meghir, 1998; Steiner and Wrohlich, 2004, 2008; Callan, van Soest, and Walsh 2009; Immervoll, Kleven, Kreiner, and Verdelin, 2009). An extensive survey of this literature has been proposed by Keane (2011).

We propose an empirical model to test the impact of joint taxation on female labour decisions, and thus ask whether the exposure to a joint tax treatment plays an important role on the labour responses of women, and whether this effect is negative as suggested by economic theory. A natural starting point is to compare women from a country eligible for joint taxation with women from a country ineligible for. In this spirit, as already performed by some scholars (e.g. Steiner and Wrohlich, 2004; Dearing, Hofer, Lietz, Winter-Ebmer, and Wrohlich, 2007), we investigate the impact of joint taxation on female labour decision through the comparison between Austrian and German tax schedules. To this aim, we use micro data from the European Union Statistics on Income and Living Conditions (EU-SILC) to evaluate female labour responses to two different taxation systems: joint taxation in Germany, whereas individual taxation in Austria. We examine female labour behaviour at the extensive margin as well as at the intensive margin by considering three different participation states (marginal, part-time, or full-time participation).

In this framework, issues of selection biases must be taken into consideration. A joint determination of the causal association between female labour decision and marital status is likely to emerge. On one hand, the marital status is likely to influence female labour behaviour, and indeed a dummy for the marital status is often included in the labour supply equation as explanatory variable. On the other, the participation of women in the labour market affects female prospective of getting married and having children. To deal with this endogeneity problem we first use a two-stages least square (TSLS) estimator, which is advocated to be consistent in estimating the average causal effects of the treatment (Angrist, 1991; Angrist and Pischke, 2009). Second, we re-estimate our model using a set of bivariate probit (BP) models. The identification of both maximum likelihood (ML) and TSLS estimates

relies on the existence of at least one variable, called instrument, which is correlated with the treatment variable but is uncorrelated with our outcome measure. The definition of our instrument will be discussed below.

This study is connected to three strands of literature. First, our work builds on the literature suggesting that joint taxation may create disincentives to female participation in the labour market, and that it may play a fundamental role in explaining cross-country differences in female labour responses (Blundell, 1995; Prescott, 2004; Davis and Henrekson, 2004; Rogerson, 2006; Steiner and Wrohlich, 2008; Olovsson, 2009).

Second, to the extent that joint taxation has implications for gender discrimination in employment by leaving women out of the market, our work relates to recent literature which argues that gender discrimination in employment may affect negatively economic growth. Some examples are Seguino (2000), Klasen and Wink (2002), Seguino and Floro (2003), Esteve-Volart (2004), Klasen and Lamanna (2009).

Third, several scholars evaluate the impact of tax reforms on labour force participation. The most recent and related studies are LaLumia (2008) for the United States, Steiner and Wrohlich (2008) and Bach, Geyer, Haan, and Wrohlich (2011) for Germany, Selin (2009) for the Swedish context, Kalíšková (2014) for the Check Republic case, Jaumotte (2003) and Bick and Fuchs-Schündeln (2017) for a sample of OECD countries which includes both joint and separate taxation systems.

Particularly, our work is connected with the studies of Steiner and Wrohlich (2004) and Dearing et al. (2007) who assess the impact of joint taxation on female labour supply through the comparison between Austrian and German tax schedules. Indeed, as above mentioned, we also conduct the analysis by comparing Austrian and German tax systems. However, our work differs from previous studies for three aspects. First, we use updated micro data for 2012. Second, they carry out the analysis at mothers' labour supply instead of women level. Third, previous studies apply simulation strategies to evaluate the impact of joint taxation on female labour behaviour, whereas our investigation is performed with econometrics methods by exploiting past information. In light of the foregoing, our main contribution to this literature is to examine the economic hypothesis of a negative effect of joint tax treatments on secondary earners by using program evaluation methods based on

instrumental variables estimations which exploit the variation induced by the existence of the counterfactual framework, and importantly which face up to selection biases. Blundell and MaCurdy (1999) develop a review of the estimation approaches that can be used to modeling labour supply.

In answering the questions posed above – does the exposure to a joint tax treatment play an important role on the labour responses of women? If any, is this effect negative as suggested by economic theory? - we find that the nature of the tax treatment seem to matter for women's labour decisions. Our analysis shows that German married women, who are exposed to joint taxation, are more likely to not participate in the labour market than women who are not exposed to. Moreover, considering the labour behaviour at the intensive margin we find that women who are taxed with joint taxation are estimated to be almost 14 percentage points more likely to work marginally (less than 15 hours per week) than women who are taxed with separate tax treatments; yet, women who are exposed to joint taxation are almost 5% (8%) less likely to work part-time (full-time) than women who are not exposed to. Therefore, the results seem to be in line with previous literature establishing that joint taxation discourages female participation in labour force.

The remainder of the chapter is organized as follows. Section 2 provides a review of the empirical literature. Section 3 gives a brief description of the existing family tax systems of married couples. In section 4 we illustrate our identification strategy. In fifth section we describe institutional framework in Austria and Germany, focusing on unemployment rates and fiscal treatment of couple families. Section 6 illustrates data and some descriptive evidence. Section 7 reports our results. Finally, section 8 concludes this chapter.

2. Brief literature background

From a theoretical point of view, joint taxation leads to equal marginal tax rate between primary and secondary earners. Since husbands earn usually more than wives, the definition of household as tax unit involves lower marginal tax rates for husbands and higher marginal tax rates for wives, on average. Thus, joint taxation could constitute a disincentive to work for married women (Blundell, 1995).

Several studies investigate the validity of this economic theory. Some scholars examine the effect of joint taxation on women labour supply through microsimulation of a switch from joint to individual taxation (see e.g. Steiner and Wrohlich, 2004; Dearing et al., 2007; Haan, 2010; Bach et al., 2011). Other scholars use program evaluation methods which exploit the variation induced by policy reforms or in general by the existence of the counterfactual framework (Rubin, 1974). For instance, LaLumia (2008) and Selin (2009) exploit the introduction of policy reforms in estimating the effect of joint taxation on labour supply.

Using a microsimulation approach, Steiner and Wrohlich (2004) examine the effect of a switch from joint to individual taxation on labour participation of married women in Germany. They find that a *"joint versus individual shifting"* reduces substantially the splitting advantage in West Germany but only marginally in East Germany. In particular, they sustain that a switch from joint taxation to individual taxation would raise female labour participation by approximately 430 thousand housewives, whose about 95% lives in West Germany. They suggest that the effect on women labour participation is more prevalent in West Germany because of the labour market differences existing between West and East Germany - East German women have a participation rate in work larger than West German women - and the fact that a change in tax transfer from joint to individual scheme would affect especially single-earner household, that is, West German households.

Dearing et al. (2007) exploit similarities and differences existing between Germany and Austria to investigate whether variation in labour supply response of German and Austrian mothers depends on differences in both taxation schemes and childcare systems. While simulating the labour supply of German mothers in a hypothetical tax regime based on individual taxation and parental leave benefit schemes applied in Austria, they show that the introduction of Austrian parental leave benefit, which is more generous than German childcare system, would decrease labour supply of German mothers, whereas the introduction of individual taxation would increase labour supply of mothers in Germany. Further, the simultaneous introduction of both individual tax and childcare schemes would positively affect labour supply of German mothers because, they argue, the labour supply increase derived by *joint versus individual shifting* would be quantitatively

larger than its decrease due to the introduction of Austrian parental leave benefit scheme.

Moreover, Bach et al. (2011) use a tax-benefit microsimulation model to examine the effect of a tax reform which entails a shift from joint to individual taxation on labour supply of German married women. They sustain that joint taxation generates a strong disincentive for married women, and there would be a significant rising in the labour supply of married women if the simulated tax reform were realized. In terms of effort, they show that married women would increase their average working hours by about 7.4%.

Focusing on the United States joint tax system, Guner, Kaygusuz, and Ventura (2012) evaluate the introduction of a revenue-neutral tax reform in which individual taxation is applied for married household. They find that female labour supply increases by almost 4%, which mainly derives from the rise in worked hours by married women. While simulating an inverse tax reform, that is the introduction of joint taxation in an individual-base tax system, Colonna and Marcassa (2015) who use Italian micro data from EU-SILC (2007-2011) show that joint taxation brings about a decrease in married women labour participation rates, and that female labour supply and husband' income are correlated negatively.

Other scholars exploit program evaluation methods. For instance, Crossley and Jeon (2007) examine the labour supply of married women in Canada by using a difference-in-difference (DID) approach. In the 1988 the Canadian government applied a federal tax reform which involved a decrease in wife's effective marginal tax rate when husband's marginal income tax rate was relatively high, but the wife's effective marginal tax rate remained almost unchanged when husband's marginal income tax rate was relatively low. Hence, this reform implied a reduction of "jointness" by current tax regime. They applied the DID method by comparing the change in labour supply of women with husband's income higher than a specific threshold (treated group) and change in labour supply of women with husband's income lower than the same threshold (untreated group). The results show that the reduction of jointness on family taxation implies an increase in labour supply of low educated married women. Therefore, Crossley and Jeon' (2007) work supports the

idea that a switch from joint to separate taxation has an incentive power in increase the labour force participation of married women.

LaLumia (2008) tries to examine how married couples' labour participation rates responds to introduction of joint taxation by exploiting the existence of cross-state variation in tax regimes in the United States before 1948. In fact, in 1948 the government of the United States introduced the joint taxation as legally tax regime; however, regions with community property laws already used this tax scheme before 1948. Considering married women exposed to joint taxation before 1948 as treated group and married women not exposed to joint taxation before 1948 as untreated group, she uses a DID approach which estimates the change in potential outcome of treated group between 1940 and 1950 and the variation in potential outcome of untreated group from 1940 to 1950. The empirical findings show that the transition from individual to joint taxation determined a decline in labour participation rate by 2 percentage points among married women in highly-educated couples. Selin (2009) reaches a similar result but of a higher magnitude than LaLumia (2008) by studying the abolition of joint taxation in Sweden in 1971. Indeed, he finds that the employment rates of married women increased by 10 percentage points.

Moreover, Carbonnier (2008) who exploits French income tax returns for 2005 shows that joint taxation have a negative impact on the probability of the secondary earner to participate in the labour market. Also Kalíšková (2014) evaluates the labour responses of married women by using a difference-in-difference method. For this purpose, she exploits the introduction of joint taxation in Czech Republic in 2005. This policy reform prescribed the application of joint taxation for married couples raising at least one child. This condition allows Kalíšková to compare labour participation rates of married women bringing at least one child up (treated group) with labour participation rates of married women without children (untreated group). She finds that the introduction of joint taxation in Check Republic implied a decrease in employment probability of married women with children by 2.9 percentage points relative to unmarried or childless women.

3. Theoretical background

Across Europe different tax treatments are applied for married couples. Some countries apply individual taxation systems, which tax each spouse separately for their own income without regards on their marital status (e.g. Austria, Denmark, Finland, and Sweden). Other countries use joint taxation rules, which either levy the tax on the total income of household as a whole or levy the tax on each spouse individually for half of the total income (e.g. Czech Republic, France, and Germany).

Comparing the labour responses of women from a country eligible for joint taxation with women from a country ineligible for might be a natural starting point in order to investigate the impact of joint taxation on labour decision of women. To set the idea behind our analysis we briefly describe the different income tax regimes for married couples, and the implied effects on women labour decisions.

Married couples might be taxed individually or jointly.

Individual taxation function: $$T(Y_1, Y_2) = T(Y_1) + T(Y_2)$$

Joint taxation function: $$T(Y_1, Y_2) = T(Y_1 + Y_2)$$

Assuming that the household income is distributed unequally within the married couple, with $Y_2 > Y_1$, and the income tax schedule is progressive, individual taxation has the disadvantage of treating differently households with same total income but with differences in income distribution among spouses, which involves a discrimination of income distribution among partners. Otherwise, joint taxation has the drawback of causing a marriage penalty, namely the tax burden increases when the individuals get married; moreover, the more equally distributed the household income, the bigger this marriage penalty. Assuming that husbands earn on average more than wives, joint taxation leads to disadvantage dual-earners couples when they get married, and disproportionately discourages women labour-force participation (Steiner and Wrohlich, 2004).

To avoid both discrimination of the income distribution among spouses and marriage penalty, Germany applies a joint tax regime with income splitting. In this case, couples are allowed for joint tax filing, which overcomes income distribution discriminations, and then their jointed income can be divided and taxed separately.

Joint taxation function with income splitting: $T(Y_1, Y_2) = 2T\left(\frac{Y_1 + Y_2}{2}\right)$

Therefore, under income splitting tax regime, on one hand, the tax burden of the married couples depends exclusively on the sum of incomes, instead of the income distribution between partners; on the other hand, this tax treatment saves married couples from marriage penalty, that arises in progressive tax system, because the partners' tax burden does not increase when they get married.

However, income splitting implies no-tax neutrality towards marital status. Assuming a progressive tax system, an unmarried couple will pay a larger amount of income tax than a married couple with the same household income. This difference in tax burden between unmarried and married couples is defined splitting advantage, and is illustrated in Figure 1. Point *A* indicates the tax liability of an unmarried couple, while point *B* is the tax liability of a married couple under income splitting. The vertical difference is the splitting advantage. Therefore, the joint taxation with income splitting translates into a fiscal advantage of married couples over unmarried couples with the same household income.

[FIGURE 1 ABOUT HERE]

Figure 2 illustrates the income splitting advantage for married couples under the joint taxation system in Germany in 2015/2016. The first (second) number indicates the primary's (secondary's) percentage share in the spouses' joint household pre-tax income per year (in euro). As can be seen, the larger the gap in the percentage share in the spouses' joint household pre-tax income per year, the bigger the splitting advantage. Therefore, the income splitting tax system discourages married women to participate in the labour market as well.

[FIGURE 2 ABOUT HERE]

4. Identification strategy

In this work we look at the impact of joint taxation on the labour responses of women, which are assumed to be the secondary earners. To employ this analysis we focus on the German context, where a joint taxation system with income splitting for married couples is applied. The common approach used in many studies on policy evaluation is to apply a treatment effects analysis, and compare the pre – post treatment data. However, this approach cannot be used to examine the impact of joint taxation on women's labour behaviour in Germany because the shifting towards a system of joint taxation happened since the 1980s. Given the lack of information on women labour force participation before - after the German tax reform, the investigation of the impact of joint taxation on women's labour decision might be carried out comparing the labour response for a group treated with a joint tax system (married women living in Germany) to the labour response for a group not treated with (unmarried women living in Germany and married/unmarried women living in Austria). Indeed, Austria uses an individual fiscal schedule for single taxpayers as well as for married couples. Importantly, to estimate the average causal effect of joint tax system on female labour responses, a fundamental assumption is that the decision of getting married reflects the voluntary exposure to family fiscal treatment, that is, the joint taxation with income splitting in the German case.

In defining the estimation strategy, we have paid particular attention to the issue of selection bias. When women with low-skills are more likely to get married, then single-equation models would misestimate the effects of a joint taxation. Therefore, it is necessary to take this self-selection into consideration – potential sources of self-selection will be better discussed below. To deal with this issue, it is advocated the employment of TSLS technique as a consistent estimator of the average casual effects of the treatment (Angrist, 1990; Angrist and Pischke, 2009). Therefore, we first start to perform two-stage least square estimations.

However, Angrist (1991) himself suggest that TSLS estimator might be inferior to a maximum likelihood (ML) specification. For this reason, our second step is to re-estimate our model using maximum likelihood regressions. Particularly, since the outcome variable and the treatment variable (dummy for German married women)

are dichotomous, we examine the women's decision to working by estimating a set of bivariate probit models. A similar approach has been used for evaluating both labour supply and income wage impacts of various policy reforms (see e.g. Evans and Schwab, 1995; Angrist and Evans, 1998; Angrist, 1998; Abadie, 2003).

The model is formulated as a discrete choice model evaluating the female labour response at both extensive and intensive margins(marginal, part-time, and full-time participation). Thus, female labour decision can be characterized as follows:

Participation decision: $Y_i = 1$ if women i works, else $Y_i = 0$ (non-participation).

Marginal participation: $Y_i = 1$ if women i weekly working time does not exceed 15 hours, and $Y_i = 0$ otherwise.

Part-time participation: $Y_i = 1$ if women i weekly working time varies from 16 to 35 hours, and $Y_i = 0$ otherwise.

Full-time participation: $Y_i = 1$ if women i weekly working time is strictly larger than 35 hours, and $Y_i = 0$ otherwise.

The choice problem is described by the latent variable model (IV second stage regression):

$$Y_i = X_i\zeta + W_i\beta + \epsilon_i \quad (1)$$

where Y_i indicates the four alternative outcome variables, defined as above, W_i is our treatment variable that equals 1 for German married women and 0 otherwise. The X_i vector of controls used here includes socio-demographic characteristics that affect labour preferences, and finally ϵ_i is a normally distributed random error with zero mean and constant variance.

We observe individual i as participant, marginal participant, part-time participant, or full-time participant whether the expected net benefit is positive, and thus the probability of observing individual i as participant, marginal participant, part-time participant, or full-time participant is given by

$$prob[Y_i = 1] = prob[X_i\zeta + W_i\beta + \epsilon_i > 0] = \Phi[X_i\zeta + W_i\beta] \quad (2)$$

where $\Phi[\ ]$ expresses the conditional distribution function.

As above said, results from single-equation might be not robust because the equation in this study is likely to suffer from self-selection biases. A first concern comes from the possibility of having omitted relevant (observable/unobservable) characteristics of the women that are correlated both with the marital status and with the probability of working and that, consequently, cause the overestimation of the benefits of the exposure to the joint taxation. This is the case of female unobservable-skills which might influence both the probability of working and the probability of getting married, and then are a potential source of endogeneity. Second, we could expect that the decision of getting married and bringing children up affects the propensity of working, and simultaneously we could expect that the propensity of working might affect the decision of getting married in order to exploit fiscal advantage from joint tax treatment when is provided.

Both ML and TSLS specifications we estimate can address these omitted variables bias and reverse causality issues, which cannot be dealt with standard ordinary least square (OLS) estimations. These two-equation estimation strategies have the advantage of providing consistent estimations when the treatment variable is endogenous, though their credibility relies on the existence of at least one variable, called instrument, which fulfills two important conditions. First, it is required that the instrument is correlated with the endogenous variable (the relevance condition). On the other hand, it is necessary no-correlation between the instrument and the outcome variable except through the endogenous variable (the exclusion restriction). While the first condition can be tested by conducting some analysis, the second condition, which entails the exogeneity of our instrument, is not verifiable. Indeed, unless the model is overidentified (i.e., there are more instruments than endogenous variables) testing for the failure of the exclusion restriction is not possible because we cannot observe the error term, and thus we need to rely on economic theory or economic insights to find instrumental variables that fulfill this condition. In short, identifying a *"convincing instrument"* which is relevant (i.e. it is not a weak instrument) and exogenous (i.e. it satisfies the exclusion restriction) is one of the biggest challenges for economists.

Talking of instrumental variables and program evaluation methods, Wooldridge (2010) affirms that *"Actual participation is almost always voluntary, and it may be endogenous because it can depend on unobserved factors that affect the response. However, it is often reasonable to assume that eligibility is exogenous. Because participation and eligibility are correlated, the latter can be used as an IV for the former"* (Wooldridge, 2010, p.94). Thus, since participation into the treatment is likely to suffer from endogeneity because of selection bias, Wooldridge (2010) suggests that eligibility for the treatment can be used as an exogenous instrument because *often* satisfies the condition of being correlated with the treatment variable and uncorrelated with the outcome measure. In the spirit of this and of what previous literature has already done (Angrist, 1991; Evans and Schwab, 1995; Angrist, Imbens, and Rubin, 1996; Schultz, 2004; Leòn and Younger, 2007; Ponce and Bedi, 2010; Gundersen, Kreider, and Pepper, 2012), we choose eligibility for the treatment as our instrument. In detail, our instrument consists in a dummy variable that equals 1 when the woman is from a country eligible for the fiscal treatment of joint taxation (German women) and 0 when the woman is from a country ineligible for (Austrian women).

After employing IV analysis a set of tests should be performed to control for the main problems that can arise in this framework: heteroskedasticity, weak instrument, endogeneity, and exogeneity of regressors. We check for the presence of these problems by performing a battery of tests for each specification. First of all, we control for heteroskedasticity by implementing the Pagan and Hall' (1983) test. Indeed, instrumental variables analysis is consistent in the presence of heteroskedastic errors; their IV estimations are inefficient, though.

To test for the fulfillment of the relevance condition (weak instruments) different tests are available; most of them are easily based on the F-statistics of the first stage regressions. Staiger and Stock (1997) and Stock and Yogo (2005) proposed widely used tests which belong to this class of tests using F-statistic. We first include the Staiger-Stock test, which relies on the assumption of homoskedasticity and autocorrelation. The rejection of the null hypothesis occurs on the basis of a "rule of thumb" which implies to reject that the instrument is weak if F-statistic in the first stage is larger than ten. Stock and Yogo (2005) formalize Staiger and Stock' (1997)

rule of thumb; this test rejects the null hypothesis of weak instruments when the Cragg and Donald (1993) statistic is larger than given critical values.

Moreover, we also report confidence intervals of the conditional likelihood ratio (CLR) test (Moreira, 2003), which is fully robust to weak instruments when only one endogenous variable is included in the model. CLR confidence intervals are reported with Limited Information Maximum Likelihood (LIML) estimates, as these are more robust to weak instrument than TSLS estimates.[31] The test rejects the null hypothesis of weak instruments when the dependent variable coefficient falls within the range of CLR (LIML) confidence interval.

Since all these procedures rely on the assumption of conditionally homoskedastic and serially uncorrelated model errors, we finally perform Montiel-Pflueger test (see Pflueger and Wang, 2015), which is robust to weak instruments under heteroskedasticity, autocorrelation, and clustering. This test, which is an extension of Stock and Yogo (2005), compare "Nagar bias" estimator relative to "worst case" benchmark (when instruments are completely uninformative and when first- and second- stage errors are perfectly correlated) TSLS and LIML estimators with a single endogenous variable, *treatment dummy* in this case. The rejection of the null hypothesis occurs when the effective F-statistic exceeds a certain critical value.

Lastly, we test for endogeneity of instrument. Indeed, before implementing IV analysis, it is important to check whether it is necessary to use an instrumental variables technique, that is, whether a set of IV estimates is consistent. Baum, Schaffer, and Stilllman (2007) suggest the Durbin-Wu-Hausman (DWH) test to check for the presence of endogeneity. The null hypothesis states that an ordinary least squares (OLS) estimator of the same equation would yield consistent estimates: that is, any endogeneity among the regressors would not have deleterious effects on OLS estimates. A rejection of the null hypothesis indicates that endogenous regressors' effects on the estimates are meaningful, and instrumental variables techniques are required.

[31] Differently from Staiger and Stock (1997) and Stock and Yogo (2005), Moreira' (2003) CLR approach, like Anderson-Rubin (AR) and Lagrange Multiplier (LM) tests, has the advantage to use a statistic whose distribution does not depend on concentration parameter μ. Though, CLR is showed in simulations to outperform AR and LM tests (Moreira and Stock, 2006).

As already mentioned, testing for the failure of the exclusion restriction (exogeneity) is not possible in case of exactly identified models, but strong assumptions are needed. Particularly, the credibility of the exogeneity of our instrument relies on the assumption that German women are no more likely to find a job than Austrian women, that is the eligibility criteria must be uncorrelated with our outcome variable. As we argue below, it appears that either being German or being Austrian does not imply significant differences in probability of entering in the labour market. To control for the reliability of this assumption next section reports statistics information on the probability of finding a job in Austria and Germany by focusing on unemployment rates for 2012, which is the year used for our empirical analysis. Moreover, exogeneity of our instrument could fail with regards to immigrants, who could make their decision about the receiving country by pondering multiple factors - unemployment rates and taxation rules are among the factors they might consider. For this reason, we include immigrant status dummies as control variable. We also estimate the model by using a sample of only native women, and find that estimated coefficients are similar.[32] This would suggest that the results are not very sensitive to the sample used.

After detecting the importance and of IV estimates we now present the reduced-form equation (IV first stage regression) as described below:

$$W_i = X_i\zeta + Z_i\gamma + \mu_i \tag{3}$$

where X_i is the vector of observable characteristics, included also in equation (1), and Z_i is our instrument, which is not contained in equation (1).

Imbens and Angrist (1994) show that β_{2SLS} estimator, named Wald estimand, can be interpreted as a local average treatment effect (LATE). The LATE is the average treatment effect for those *"who can be induced to change status by a change in the instrument"* (Imbens and Angrist, 1994, p.470). Namely, TSLS regressions estimate the average effect of $W_i$on Y_i for those who change status in response to a change in instrument. Different instruments might identify different average effects. In other words, β_{2SLS} estimator can be interpreted as the average causal effect for those who comply with the particular instrument, named compliers. It is hence local because it

[32] Resuts are shown in Tables 14 through 21 in Appendix.

applies to special group of individuals and depends on the specific used instrument (Imbens and Angrist, 1994). Formally, the LATE can be expressed as

$$LATE = \frac{E[Y\,|Z=1]-E[Y\,|Z=0]}{Prob[W=1\,|Z=1]-Prob[W=1\,|\,Z=0]} = E[(Y_1 - Y_0)|\,(W_1 - W_0) = 1] \quad (4)$$

Angrist et al. (1996) provide a classification of population according to the link with the treatment and the instrument. This classification can be exemplified as follows:

Angrist, Imbens, and Rubin' (1996) population classification

Z = 0	W = 0	*Compliers*	W = 1	Z = 1
	W = 0	*Never takers*	W = 0	
	W = 1	*Always takers*	W = 1	
	W = 1	*Defiers*	W = 0	

Sources: Angrist et al.' (1996) classification (p. 448).

In the special case in which there are not always takers, as our case analysis is, the denominator term of Wald estimand becomes $Prob[W = 1\,|Z = 1]$ because when individuals are ineligible for the treatment $(Z = 0)$ they are always not treated $(W = 0)$, hence, the $Prob[W = 1\,|\,Z = 0] = 0$. Consequently, the Wald estimand can be interpreted as the average treatment effect on the treated (ATT):

$$LATE = \frac{E[Y\,|Z=1]-E[Y\,|Z=0]}{Prob[W=1\,|Z=1]} = E[(Y_1 - Y_0)|\,W = 1] = ATT \quad (5)$$

Therefore, TSLS model estimates the average treatment effect on the treated. In other words, the Wald estimand we estimate can be interpreted as the average effect of joint taxation on women labour responses for those women who get married and live in Germany (treated group).

In contrast, the β_{ML} estimator has the advantage of estimating the average treatment effects for the population (ATE). Angrist and Pischke (2009, p.201) stated that *"Bivariate probit and other models of this sort can be used to estimate unconditional average causal effects and/or effects on the treated"*. Hence, using the bivariate probit approach the coefficient we estimate can be interpreted as the average causal effect of joint taxation on female labour decisions.

5. Institutional background

5.1 Unemployment rates evidence

The credibility of our identification strategy depends on the assumption that German women are no more likely to find a job than Austrian women. For this reason, in this section we summarize statistics information of unemployment rates in these countries for 2012, which is the year used for our empirical analysis, in order to assess whether there are different probabilities of entering in the labour market, especially for women.

Figure 3 reports the unemployment rates in 2012. Austria and Germany report unemployment rates greatly lower than the OECD mean; the position of Austria is very close to the position of Germany, though the former is slightly inferior to the latter. Specifically, the unemployment rate of Austria is 4.9%, while the value reported by Germany is 5.4%. Moreover, focusing on only women, Figure 4 confirms that the unemployment rates of Austria and Germany are very close. The female unemployment rate in Austria is 4.8%, while the value reported in Germany is 5.2%. Therefore, it seems reasonable to assume that being German women or being Austrian women does not imply significant differences in probability of entering in the labour market.

[FIGURES 3 TO 4 ABOUT HERE]

5.2 Austrian and German system for couple families

To examine the impact of joint taxation on women labour decision we compare the labour responses for a group eligible for the joint taxation (women living in Germany) to the labour responses for a group ineligible for (women living in Austria). For the sake of comparability, it is fundamental to verify that factors influencing female labour decisions are similar in German and Austrian systems. Because of this, we compare Austria and Germany in the elements which might significantly affect

women labour decision: child benefit rules, parental leave benefits, and tax treatments.

Considering the child benefit rules, Austria and Germany share some similarities. In Germany, the child benefit – "Kindergeld" – consists of a money transfer from the government to all families who have children, regardless with both parents' employment status and their income level. This financial support is provided for each family until child reaches the age of majority – until their 25th birthday if they are still enrolled in school. It ranges from a minimum of 184.00 Euro per month to a maximum of 215.00 Euro according to children number. Families may be alternatively entitled to a child tax allowance whether the amount of the tax relief exceeds the child benefit (Dearing et al., 2007).

Similarly, in Austria all families are entitled to the child benefit – "Familienbeihilfe" - independently on both parents' employment status and their income level. The amount is determined relative to age of children, varying from 109.70 Euro per month from birth to 158.90 Euro per month from 19 years old and over – child benefit is received since children reach the age of majority, but in some cases may be prolonged until their 24th birthday. It is still increased in value when the number of children rises, and the supplement varies from 6.70 Euro per month for two children to 50.00 Euro per month for seven and more children. In Austria each family who receives child benefit is also entitled to child tax credit – "Kinderabsetzbetrag" - which amounts to 58.40 Euro per month per child (Dearing et al., 2007).

On the other hand, they diverge for two important aspects. First, they apply different tax treatment for married couples. As above mentioned, Austria applies separate taxation for single tax payers as well as married couples; whereas Germany applies separate taxation only for single tax payers, and joint taxation with income splitting for married couples. This is the aspect that allows implementing this analysis, and that leads to expect that German women work less than Austrian women.

Second, they differ significantly for the parental leave benefits. In Austria, parental leave period is from at least three months to at most 24 months, and the entitlement to the parental leave is individual based. It may be taken by one parent only or by both parents on an alternating basis. Parental leave is unpaid, though

parent who receives parental leave could be entitled to a child-care allowance in case his/her gross income does not exceed 14,600.00 Euro per year. This child-care allowance is granted for 30 months, or 36 if both parents apply for the payment, and it amounts to almost 436.00 Euro per month – almost 618.00 Euro for those who earn less than 1,000.00 Euro per month (Dearing et al., 2007). By contrast, in Germany the parental leave period is until 3th birthday of child. Parents who are entitled to parental leave receive a parental benefit – "Elterngeld" – which is granted for 12 months, or 14 if both parents apply for. The amount is 67% of the applying parent's net income, and ranges from 300.00 Euro to 1,800 Euro – parents who are in unemployment status are eligible for the minimum of 300.00 Euro.

Therefore, differences in parental leave benefit concern mostly the grant period, which is shorter in Germany (30 months in Austria but 12 months in Germany), the income threshold, which is larger in Germany (14,600.00 Euro per year in Austria but 250,000.00 Euro per year in Germany), and the amount of grant, which is larger in Austria when total benefit period and low income earners are considered (Cygan-Rehm, 2016). Since Austria has a more generous parental leave benefit than Germany, a negative sign in our analysis, which suggests that German women work less than Austrian women, would be caused only from joint taxation.

6. Data and descriptive evidence

The estimation is based on micro data from EU-SILC (2008-2012) which includes information about personal characteristics (age, education, marital status), the economic activity during the reference week (employed, unemployed, retired), and the number of hours worked in the reference week (if employed). Cross-section data of women from Austria and Germany for the year 2012 are used. Only women who are aged 20-50 are included in order to avoid biased issues deriving from differences in both early retirement rules and education system duration between the two countries.

Aiming at measuring the extensive and intensive female labour response four different dependent variables are employed in this study. "Participation decision", which is basic labour information on current activity status, is defined as a dummy

which takes value 1 when woman is at work.[33] We also investigate the intensive labour response by splitting the weekly labour supply into marginal, part-time, or full-time jobs. To do this, it is necessary to define intervals for working time in the three states. A worker is categorized as a marginal participant, part-time participant, or full-time participant if her weekly labour supply does not exceed 15 hours, varies from 16 to 35, or is strictly larger than 35, respectively.

The explanatory variables used here include socio-demographic characteristics which might affect female labour decision (e.g. age and level of education). The number of preschool children is not included because of missing information into the EU-SILC (2008-2012), and might be a cause of omitted variables bias, which is dealt with IV strategy. The age controls we use are age, age square, and a dummy that takes value 1 for women aged 30 years old and over. Education, which is measured by the International Standard Classification of Education (ISCED), is defined as the highest level of an educational programme the person has successfully completed. This information is included for each women and relative spouse (if any). Activity status of spouse is also included (where occurs) to account for couples with women as primary earner when spouse is in unemployment status. This control, named "Unemployment status of spouse", is measured by a dummy that equals to one when spouse is in unemployment status, and zero otherwise. Finally, we control for immigrant status by including a dummy which takes value 1 if the individual *i* is immigrant, and value 0 otherwise. Similar control variable have been included in labour supply equations by Crossley and Jeon (2007), LaLumia (2008), Selin (2009), and Kalíšková (2014).

Table 2 reports basic summary statistics for the variables included in the estimations. This table shows that the participation rates do not vary considerably across groups. However, Austria and Germany differ with respect to weekly hours. German women work marginally to a much larger degree than Austrian women. On the other hand, the part-time and full-time employment rates in Austria are higher than in Germany. Further personal characteristics appear pretty stable across groups. We observe that

[33] The status should be self-defined. EU-SILC Guide Lines (2009) fix that the concept of "current" activity status rules out the possibility of any concept of averaging over any specific reference period, the status as of the time of the interview should be reported instead.

the average education level, based on ISCED classification, is slightly smaller in Austria than in Germany.

[TABLE 2 ABOUT HERE]

Figure 5 illustrates the distribution of total hours worked per week in the two countries. It is worth noting that the number of Austrian and German women who chooses a full-time job is pretty similar. However, we can observe that a part-time working experience is mainly preferred by Austrian than German women, whereas a marginal working experience is mainly chosen by German than Austrian women.

[FIGURE 5 ABOUT HERE]

These working preferences also emerge in Table 3. This table shows that the share of women who work part-time is larger in Austria than in Germany across age cohorts. Full-time jobs are chosen similarly in Austria and Germany: in younger age cohort the share of women who work full-time is pretty large in both countries, and then a decreasing trend emerges across age cohorts. However, focusing on married women a different behaviour occurs – German married women exhibit employment rates in full-time jobs pretty lower than married women in Austria. As highlighted in previous Figure 5, marginal working experiences are preferred by German than Austrian women. This preference is especially prominent among women in couple families: in younger and older age cohorts the share of German married women who work marginally doubles the respective share of married women in Austria. For instance, we observe that in 20–29 age cohorts the share of German married women who work marginally is almost 24%, the respective share of married women in Austria is almost 11% instead. Particularly, it is noticeable that this working preference emerges only in case of married sub-sample.

[TABLE 3 ABOUT HERE]

Figure 6 illustrates histograms of women education level by ISCED classification. It is worth noting that the share of women with an intermediate education level is larger in Austria than in Germany, whereas the number of women with a tertiary education level is bigger, almost double, in Germany than in Austria.

[FIGURE 6 ABOUT HERE]

In sum, this descriptive analysis reveals differences in labour behaviour between Austrian and German women, which might be due to differences in tax treatment of couples due to the disincentive power of joint taxation on German married women.

7. Results

As illustrated by descriptive analysis there are very large differences in the observed labour behaviour of Austrian and German married women. It might be expected that the female labour responses of women from Austria and Germany varies due to the very different tax schemes that Austrian and German families face. Because of the principle of separate taxation, the disincentives for participation of the second earner are not expected to be large in the Austrian tax system. On the other hand, the German tax system is expected to have large disincentive effects on female labour decision because of the substantial deductions for married couples and the splitting principle. Therefore, we try to test the coefficients equality of the effect of exposure to joint taxation on the extensive and intensive margins of the women labour responses in Austria and Germany, and in this section we present the main estimation results.

The results from estimating the two-stage least square approach illustrated in section 5 are shown in Tables 4 through 7. Each table corresponds to a different measure of labour supply: participation decision, marginal participation, part-time participation, and full-time participation. The first table refers to the investigation of the extensive margin through the proxy for participation decision (dummy equals to one if employed), the other ones consider the intensive margin of the labour supply of women through the choice among marginal participation (Table 5), part-time

participation (Table 6), and full-time participation (Table 7). Tables from 8 to 11 follow the same structure, but employ the bivariate probit estimator. Finally, Table 12 illustrates the average treatment on the treated estimates obtained by TSLS for each specification, while in Table 13 we report the average treatment effects computed using the bivariate probit estimator. We comment results obtained using TSLS estimator in the first instance.

For each specification, we test for heteroskedasticity and for the relevance of the instrument. We detect the presence of heteroskedasticity by implementing the Pagan and Hall' (1983) test, which not surprisingly reveals that the assumption of homoskedastic errors is not fulfilled; in order to face up to this issue, robust standard errors have been used. Weak instruments problems are checked through the implementation of three tests: first stage F-statistics of Staiger and Stock (1997), CLR test, and Montiel-Pflueger test. Results of first stage F-statistics are always higher than 10 while Montiel-Pflueger F-statistics are always statistically different from zero suggesting a significant role for our instrument; besides, in each specifications LIML estimator is within the range of CLR confidence intervals provided by the Moreira test, which is robust whether the instruments are weak or strong, indicating that our estimates do not suffer from weak instruments problems. Finally, we also test for endogeneity. The small p-values of DHW test bring about the rejection of the null hypothesis in almost all specifications (expect in some regressions in Table 5 where marginal participation is used as dependent variable), concluding that the OLS regressions are not consistent, and IV estimates are needed.

[TABLES 5 TO 13 ABOUT HERE]

Model 1 in each table shows the results of the parsimonious specification which includes only age controls. First stage results suggest that the probability of being treated, that is, being German married women, is increasing in age, but not linearly. In the following models we progressively add our set of additional controls. Model 2 and Model 3 introduce the proxy for women education and partner education, respectively; while Model 4 further includes a measure of spouse unemployment status. In first stage regressions the coefficients related to education level of both

spouses are not significant, while the inactivity status of partner is significant and negative. Finally, a control for immigrant status is included in Model 5, where first stage coefficients are significant and positive, as expected.

Second stage results suggest that the effect of the exposure to joint taxation on female labour decision is significant and negative, as expected. Considering the summary Table 12, the coefficient of -0.0188 in column 1 of Panel E, which is our preferred because has the larger number of controls, indicates that married women in Germany experience an almost 1.9 percentage point decline in the probability of participating in the labour market relative to German unmarried women and married/unmarried women from Austria. This significant and negative sign associated to the current participation in the labour market persists in almost all the regressions, independently on what controls are introduced.

In column 2 of Table 12 the analysis is repeated with marginal participation measure as the dependent variable. In this case women taxed with joint taxation, namely German married women, are estimated to be approximately 14 percentage point more likely to choose marginal employment than women taxed with individual tax systems. In column 3 and 4, part-time and full-time participation are respectively used as the dependent variable. Significant and negative sign associated to the treatment effect coefficients emerges in all specifications suggesting that the exposure to joint taxation is linked to a lower probability of working part-time or full-time for women in Germany. These TSLS estimates of the average treatment effect on the treated (ATT) are pretty stable to different covariates.

Overall, TSLS estimates confirm the expectations concerning the effect of the exposure to the joint tax schemes. Namely, this analysis suggests a negative impact of joint taxation on female participation in the labour market. Focusing on the intensive margin, the German tax treatment of spouses seems responsible for the female labour preferences implying a large number of German married women to choose marginal participation, rather than part-time or full-time participation.

The findings we have illustrated here are in line with the bulk of studies on this topic including the work of LaLumia (2008), who applies a difference-in-difference (DID) strategy to examine the effect of introduction of joint taxation in the United States

after the tax reform in 1948, and of Kalíšková (2014), who applies the same econometric approach (DID) to the Czech Republic case.

Interestingly, these results are also consistent with those of Dearing et al. (2007) who simulate a simultaneous shift from both parental leave benefits and joint tax treatment applied in Germany to corresponding parental leave benefits and individual taxation applied in Austria, though they apply the analysis at mothers' labour supply instead of widely women level. In this scenario where married couples in Germany are taxed with separate taxation and take advantage from the more generous Austrian parental leave benefit scheme, while married couples in Austria are subject to joint tax regime with income splitting and German parental benefits, they show that Austrian mothers would decrease their labour force participation rates in both part-time and full-time participation, whereas German mothers would increase their labour force participation rates. Though we cannot discern whether it is a mother with youngest child 0 to 3 years old or not – because of missing information on EU-SILC database - Dearing et al.' (2007) evidence seems in line with the results that we find here.

Albeit the estimates of the average treatment effect on the treated using TSLS estimator might be consistent (Angrist, 1991; Angrist et al., 1996; Angrist and Pischke, 2009), the TSLS approach adopted so far may be subject to criticisms because of producing biased estimations as the familiar ordinary least square estimator. Therefore, we strengthen our study by introducing a further array of estimates using bivariate probit technique.

In Table 8 the analysis is carried out by using participation decision as the dependent variable. Considering the second stage results we observe negative and significant coefficients in all regressions. This suggests that the exposure to joint taxation reduce the participation of women in the labour market. Considering the Table 13, the coefficient of -0.0151 in column 1 of Panel E indicates that women in Germany have a 1.5 percentage point decline in the probability of being employed, relative to women in Austria. This captures labour responses on the extensive margin. The control variables in the most parsimonious specification (first model in Table 8) indicate that the age of the wife does not in general seem to be important for the participation decision. One reason that the age variable does not play an important

role in this setting is probably that the sample is restricted to women aged less than 50 years. In line with the literature in this field (Kalíšková, 2014) higher spouses (wife and husband) education leads to higher employment probability of women while inactivity of the partner decreases the employment probability of a woman.

Table 9 repeats the analysis with marginal participation as the dependent variable. The BP estimates are positive and significant across all specifications. In column 2 of Panel E (Table 13), the coefficient of the treatment variable is 0.1378. This result suggests that a woman taxed with joint taxation is estimated to be almost 13 percentage points more likely to work marginally, namely less than 15 hours per week, than a woman taxed with individual taxation. Considering the coefficients of control variables in Table 9, the female age does not seem to play a part in the making decision process of marginal employment. Marginal participation is decreasing in education of both spouses (wife and husband) and increasing in unemployment status of the partner instead. One interpretation of the positive coefficient associated with the spouse inactivity status may be that these families might be supported financially by government subsidies, which might be removed in the case in which household income exceeds a given income threshold.

In Table 10, the part-time participation is used as the dependent variable. Negative and significant coefficients are found in each BP estimates. Considering the result in column 3 of Panel E, estimates suggest that women in Germany have a 4.4 percentage point decline in the probability of participating in the labour market, relative to women in Austria. In all specifications of Table 10 age controls are significant. Particularly, part-time participation is increasing in age, but not linearly, and women aged 30 years and over are less likely to prefer part-time jobs. Moreover, the probability of female part-time participation decreases as women education increases, while higher education of partner implies higher probability of being employed part-time. Instead the inactivity of the partner is negatively associated with the part-time participation of wives.

The dependent variable in Table 11 is a dummy equal to one if a woman worked more than 35 hours per week. As marginal participation and part-time participation, this measure captures labour responses on the intensive margin. Considering the last row result of column 4 in Table 13, we observe that the probability of working full-

time for a German woman decreases by almost 9.6 percentage point, relative to a woman in Austria. It is also noticeable that this negative sign resists across specifications. In each specification of Table 11, full-time participation is no-linearly increasing in age. In columns 2 through 5, we observe that women with higher attained education level are more likely to work full-time while the husband education appears to be insignificant. On the contrary, unemployment status of the partner implies a negative effect on full-time labour supply of wives.

Overall, all these estimates seem to confirm previous results obtained with TSLS estimator, and especially are indicative that the treatment of spouses by the joint tax regime can have important consequences for the labour behaviour of married women. Particularly, our analysis would suggest that the joint tax systems impose excessive distortions on the labour decisions of married women, relative to those of women in separate tax regimes.

8. Conclusions

This study have investigated whether joint tax systems impose excessive distortions on the labour decisions of married women, relative to those of men and single women, discouraging the participation of married women in the labour market.

In absence of real experiments, unlike previous studies that usually focused on microsimulation strategies, we have exploited differences in tax system between Austria and Germany to perform our natural quasi-experiment analysis. In detail, we have compared the labour respponses of a group treated with joint taxation (married women living in Germany) with the labour responses of a group not-treated with joint tax regime (unmarried women living in Germany and married/unmarried women living in Austria).

Dealing with potential selection biases, we have preceded applying instrumental variables methods to examine the effect of joint taxation on female labour behaviour. Two-stage least square and bivariate probit models have been estimated by exploiting the eligibility for the evaluated program to define a convincing instrument. Namely our instrument is a binary variable which takes value 1 when women are from a country eligible for joint taxation (German women) and 0 when women are from a country ineligible for (Austrian women).

We find robust evidence of a negative effect of joint tax regimes on the female labour force participation rate. This result emerges when the female participation decision is considered at the extensive margin as well as when their intensive margin is explored through the splitting of the labour supply in three different states: marginal participation, part-time participation, and full-time participation. The evidence from the investigation of the intensive margin of the labour supply indicates that the married women taxed with joint taxation are more likely to choose marginal employment than women taxed with separate tax regimes, while more participative job experiences – either part-time or full-time participation – are more likely to being undertaken by the latter than the married women in Germany.

These findings are confirmed when considering the two-stage least square estimator as well as the maximum likelihood specifications. Moreover, they are robust to the inclusion of our controls, namely the age controls, education of spouses, inactivity of partner, and immigrant status.

In sum, our analysis suggests that the German tax system seems mainly responsible for these female labour responses indicating that the nature of the tax treatment seems to matter for female labour decisions, and importantly that the OECD countries that apply joint tax regimes might encourage female in leaving out of labour market instead of stimulating their participation. Therefore, this study seems to indicate that a move towards more neutral tax systems relative to female labour decisions – e.g. a shift from joint to separate taxation - should be considered by policy makers who have an interest in fostering the female participation in the labour market in order to increase gender equality, and then enhance economic growth.

References

Abadie A. (2003) Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113: 231-263.

Abu-Ghaida D., Klasen S. (2004) The Costs of Missing the Millennium Development Goal on Gender Equity. *World Development*. 32: 1075-1107.

Angrist J.D. (1990) Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence From Social Security Administrative Records. *American Economic Review*, 80: 313-335.

Angrist J.D. (1991) Instrumental Variables Estimation of Average Treatment Effects in Econometrics and Epidemiology. NBER Working Papers.

Angrist J.D. (1998) Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants. *Econometrica*, 66: 249-288.

Angrist J.D., Imbens G., Rubin D. (1996) Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91: 444-455.

Angrist J.D., Evans W.N. (1998) Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, 88: 450-477.

Angrist J.D., Pischke J.-S. (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Bach S., Geyer J., Haan P., Wrohlich K. (2011) Reform of Income Splitting for Married Couples: Only Individual Taxation Significantly Increase Working Incentives. *DIW Economic Bulletin*, 1: 13-19.

Bajtelsmit V., Bernasek A. (1996) Why do women invest differently than men? *Financial Counseling and Planning*. 7: 1-10.

Bajtelsmit V., Van Derhei J. (1997) Risk aversion and pension investment choices. In: Gordon M. Mitchell O., Twinney M., (Eds) *Positioning Pensions for the Twenty-First Century*. Philadelphia: University of Pennsylvania Press. 45-66.

Baum C.F., Schaffer M.E., Stilllman S. (2007) Enhanced routines for instrumental variables, GMM estimation and testing. *The Stata Journal*. 7: 465-506.

Bick A., Fuchs-Schündeln N. (2017) Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis. IZA Discussion Papers.

Blackden M., Canagarajah S., Klasen S., Lawson D. (2006) Gender and growth in Sub-Saharan Africa: Issues and evidence. UNU-WIDER Research Papers.

Boskin M., Sheshinski E. (1983) Optimal Tax Treatment Of The Family: Married Couples. *Journal of Public Economics*, 20: 281-297.

Blundell R. (1995) The Impact of Taxation on Labour Force Participation and Labour Supply. *OECD Jobs Study: Taxation, Employment, and Unemployment*, Chapter 3, Paris: OECD.

Blundell R., Duncan A., Meghir C. (1998) Estimating Labor Supply Responses Using Tax Reforms. *Econometrica*, 66: 827-861.

Blundell R., MaCurdy T. (1999) *Labour Supply: A Review of Alternative Approaches*. Handbook of Labor Economics, 3A.

Callan T., van Soest A., Walsh J.R. (2009) Tax Structure and Female Labour Supply: Evidence from Ireland. *Labour*, 23: 1-35.

Carbonnier C. (2008) Spouse labor supply: fiscal incentive and income effect, evidence from French fully joint income tax system. Working Paper.

Cavalcanti T., Tavares J. (2008) The Output Costs of Gender Discrimination: A Model-Based Macroeconomic Estimate. CEPR Discussion Papers.

Chiappori P.A. (1988) Rational Household Labor Supply. *Econometrica*, 56: 63-90.

Colonna F., Marcassa S. (2015) Taxation and female labor supply in Italy. IZA Working Papers.

Cragg J.G., Donald S.G. (1993) Testing Identiafibility and Specification in Instrumental Variable Models. *Econometric Theory*. 9: 222-240.

Crossley T.F., Jeon S.H. (2007) Joint Taxation and the Labour Supply of Married Women: Evidence from the Canadian Tax Reform of 1988. *Fiscal Studies*, 28: 343-365.

Cygan-Rehm K. (2016) Parental leave benefit and differential fertility responses: evidence from German reform. *Journal of Population Economics*, 29: 73-103.

Davis S.J., Henrekson M. (2004) Tax Effects on Work Activity, Industry Mix and Shadow Economy Size: Evidence from Rich-Country Comparisons. NBER Working Papers.

Dearing H., Hofer H., Lietz C., Winter-Ebmer R., Wrohlich K. (2007) Why Are Mothers Working Longer Hours in Austria than in Germany? A Comparative Micro Simulation Analysis. *Fiscal Studies*, 28: 463-495.

Dollar D., Gatti R. (1999) Gender Inequality, Income, and Growth: Are Good Times Good for Women? Policy Research Report, Engendering Development, Working Papers.

Drèze J., Sen A. (1989) *Hunger and Public Action*. New York: Oxford University Press.

Esteve-Volart B. (2004) Gender Discrimination and Growth: Theory and Evidence from India. STICERD Discussion Papers.

Evans W.N., Schwab R.M. (1995) Finishing High School and Starting College: Do Catholic Schools Make a Difference?. *The Quarterly Journal of Economics*, 110: 941-974.

Feldstein M., Feenberg D.R. (1996) The Taxation of Two-Earner Families. In: Feldstein M., Poterba J. (Eds) *Empirical Foundations of Household Taxation*. University of Chicago Press. 39-75.

Galor O., Weil D.N. (1996) The Gender Gap, Fertility, and Growth. *American Economic Review*. 86: 374-387.

Guner N., Kaygusuz R., Ventura G. (2012) Taxation and Household Labour Supply. *Review of Economic Studies*. 79: 1113-1149.

Gugl E. (2009) Income Splitting, Specialization, and Intra-Family Distribution. *The Canadian Journal of Economics*, 42: 1050-1071.

Gundersen C., Kreider B., Pepper J. (2012) The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics*. 166: 79-91.

Haan P. (2010) A Multi-state Model of State Dependence in Labor Supply: Intertemporal Labor Supply Effects of a Shift from Joint to Individual Taxation. *Labour Economics*, 17: 323-335.

Hinz R.P., McCarthy D.D., Turner J.A. (1996) Are women conservative investors? gender differences in participant-directed pension investments. In: Gordon M. Mitchell O., Twinney M., (Eds) *Positioning Pensions for the Twenty-First Century*. Philadelphia: University of Pennsylvania Press. 145-156.

Hungerford K. (1999) Saving for a rainy day: does pre-retirement access to retirement savings increase retirement saving? Mimeo, Washington, DC, Social Security Administration.

Imbens G.W., Angrist J. (1994) Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62: 467-476.

Immervoll H., Kleven H.J., Kreiner C.T., Verdelin N. (2009) An Evaluation of the Tax-Transfer Treatment of Married Couples in European Countries. IZA Discussion Papers 3965.

Immervoll H., Kleven H.J., Kreiner C.T., Verdelin N. (2011) Optimal Tax and Transfer Programs for Couples with Extensive Labor Supply Responses. *Journal of Public Economics*, 95: 1485-1500.

Jaumotte F. (2003) Female Labour Force Participation: Past Trends and Main Determinants in OECD Countries. OECD Economics Department Working Papers.

Kalíšková K. (2014) Labour supply consequences of family taxation: Evidence from the Czech Republic. *Labour Economics*, 30: 234-244.

Keane M.P. (2011) Labor Supply and Taxes: A Survey. *Journal of Economic Literature*, 49: 961-1075. LMU Working Papers.

Klasen S. (2000) Does Gender Inequality Reduce Growth and Development? Evidence from Cross-country Regressions. Policy Research Report, Engendering Development, Working Papers, World Bank.

Klasen S. (2002) Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development. *World Bank Economic Review*. 16: 345-373.

Klasen S. (2006) Pro-Poor Growth and Gender Inequality. *Ibero America Institute for Economic Research Discussion Papers*.

Klasen S., Lamanna F. (2009) The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries. *Feminist Economics*. 15: 91-132.

Klasen S., Wink C. (2002) A Turning-Point in Gender Bias in Mortality? An Update on the Number of Missing Women. *Population and Development Review*. 28: 285-312.

Kleven H.J., Kreiner C.T., Saez E. (2009) The Optimal Income Taxation of Couples. *Econometrica*, 77: 537-560.

Knowles S., Lorgelly P.K., Owen P.D. (2002) Are Educational Gender Gaps a Brake on Economic Development? Some Cross-Country Empirical Evidence. *Oxford Economic Papers*. 54: 118-149.

Lagerlöf NP (2003) Gender Equality and Long-Run Growth. *Journal of Economic Growth*. 8: 227-258.

LaLumia S. (2008) The effects of joint taxation of married couples on labor supply and non-wage income. *Journal of Public Economics*, 92: 1698-1719.

León M., Younger S.D. (2007) Transfer payments, mothers' income and child health in Ecuador. *The Journal of Development Studies*. 43: 1126-1143.

Leuthold J.H. (1984) Income Splitting and Women's Labor-Force Participation. *Industrial and Labor Relations Review*, 38: 98-105.

Moreira M.J. (2003) A Conditional Likelihood ratio Test for Structural Models. *Econometrica*. 71: 1027-1048.

OECD (2012) Taxing Wages 2011.

Olovsson C. (2009) Why do Europeans work so little? *International Economic Review*. 50: 40-61.

Pagan A.R., Hall A.D. (1983) Diagnostic tests as residual analysis. *Econometric Review*. 2: 159-218.

Pflueger C.E., Wang S. (2015) WEAKIVTEST: Stata module to perform weak instrument test for a single endogenous regressor in TSLS and LIML. Working Paper.

Ponce J., Bedi A.S. (2010) The impact of a cash transfer program on cognitive achievement: The *Bono de Desarrollo Humano* of Ecuador. *Economics of Education Review*. 29: 116-125.

Prescott E.C. (2004) Why do Americans Works so Much More than Europeans? NBER Working Papers.

Pritchett L., Summers L. (1996) Wealthier is healthier. *Journal of Human Resources*, 31: 844-868.

Rogerson R. (2006) Understanding differences in hours worked. *Review of Economic Dynamics*. 9: 365-409.

Rubin D.B. (1974) Estimating Causal Effects of Treatment in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66: 688-701.

Schultz T.P. (2004) School subsides for the poor: evaluating the Mexican Progresa poverty program. *Journal of Development Economics*. 74: 199-250.

Seguino S. (2000) Gender Inequality and Economic Growth: A Cross-Country Analysis. *World Development*. 28: 1211-1230.

Seguino S., Floro M.S. (2003) Does Gender have Effect on Aggregate Saving? An empirical analysis. *International Review of Applied Economics*. 17: 147-166.

Selin H.K. (2009) The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971. *International Tax and Public Finance*, 21: 894-922.

Smith N., Dex S., Vlasblom J.D., Callan T. (2003) The Effects of Taxation on married Women's Labour Supply across Four Countries. *Oxford Economic Papers*. 55: 417-439.

Staiger D., Stock J.H. (1997) Instrumental variables regression with weak instruments. *Econometrica*. 65: 557-586.

Steiner V., Wrohlich K. (2004) Household Taxation, Income Splitting and Labour Supply Incentives – A Microsimulation Study for Germany. *CESifo Economic Studies*, 50: 541-568.

Steiner V., Wrohlich K. (2008) Introducing Family Tax Splitting in Germany: How Would It Affect the Income Distribution, Work Incentives and Household Welfare?. *FinanzArchiv*, 64: 115-142.

Stock J.H., Yogo M. (2005) Testing for weak instruments in linear IV regression. In: Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. *Cambridge University Press*. 80-108.

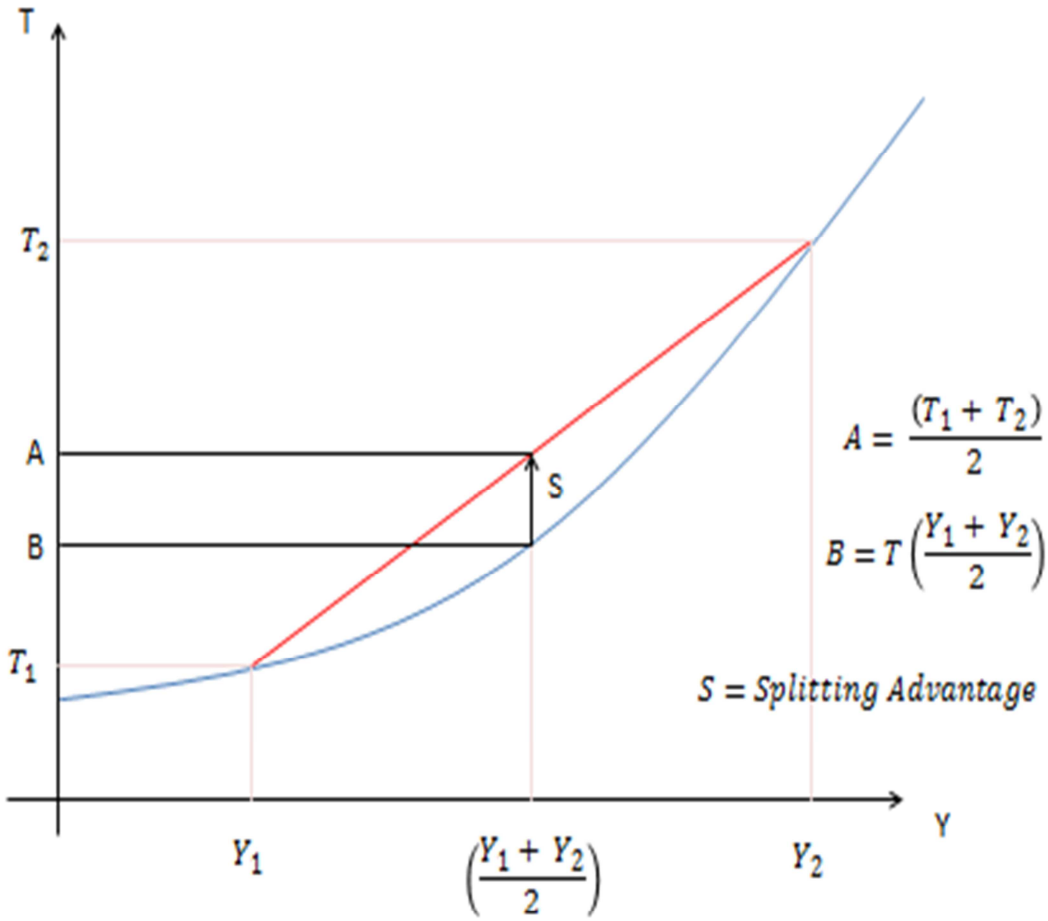
Wooldridge J.M. (2010) *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Massachusetts: MIT Press, second edition.

Table 1: Labour force participation rates in selected countries

	Total	Male	Female
Australia	64.2	71.6	57.1
Austria	60.5	66.7	54.6
Belgium	57.2	63.5	51.2
Canada	65.5	71.2	60.2
Czech Republic	63.0	72.2	54.3
Denmark	62.0	67.0	57.3
Finland	59.1	63.6	55.0
France	58.0	63.2	53.2
Germany	61.7	68.3	55.4
Greece	54.7	64.1	46.1
Hungary	57.6	67.2	49.2
Iceland	73.8	78.5	69.2
Italy	50.9	61.6	41.2
Japan	61.0	73.5	49.6
Netherlands	62.8	70.3	55.6
New Zealand	69.1	75.6	63.2
Norway	66.5	71.0	61.9
Poland	60.1	69.3	51.9
Portugal	61.7	68.3	56.0
Spain	60.9	67.9	54.3
Sweden	65.8	70.3	61.4
United Kingdom	63.3	70.0	56.9
United States	64.1	71.5	57.0
Average	61.89	68.97	55.30

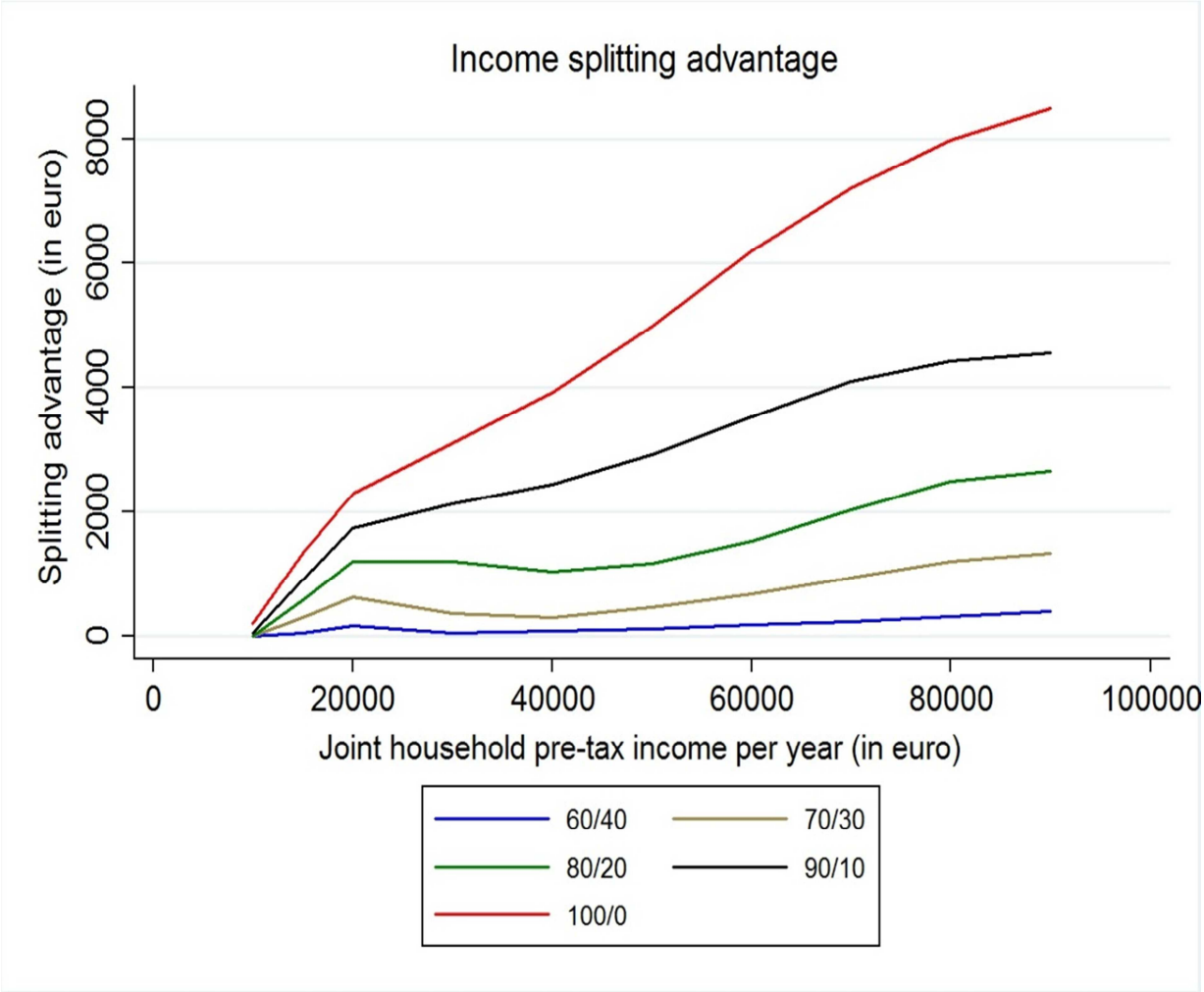
Sources: ILO estimates (%); age:25+.

Figure 1: Splitting advantage for married couples



Sources: Our computations.

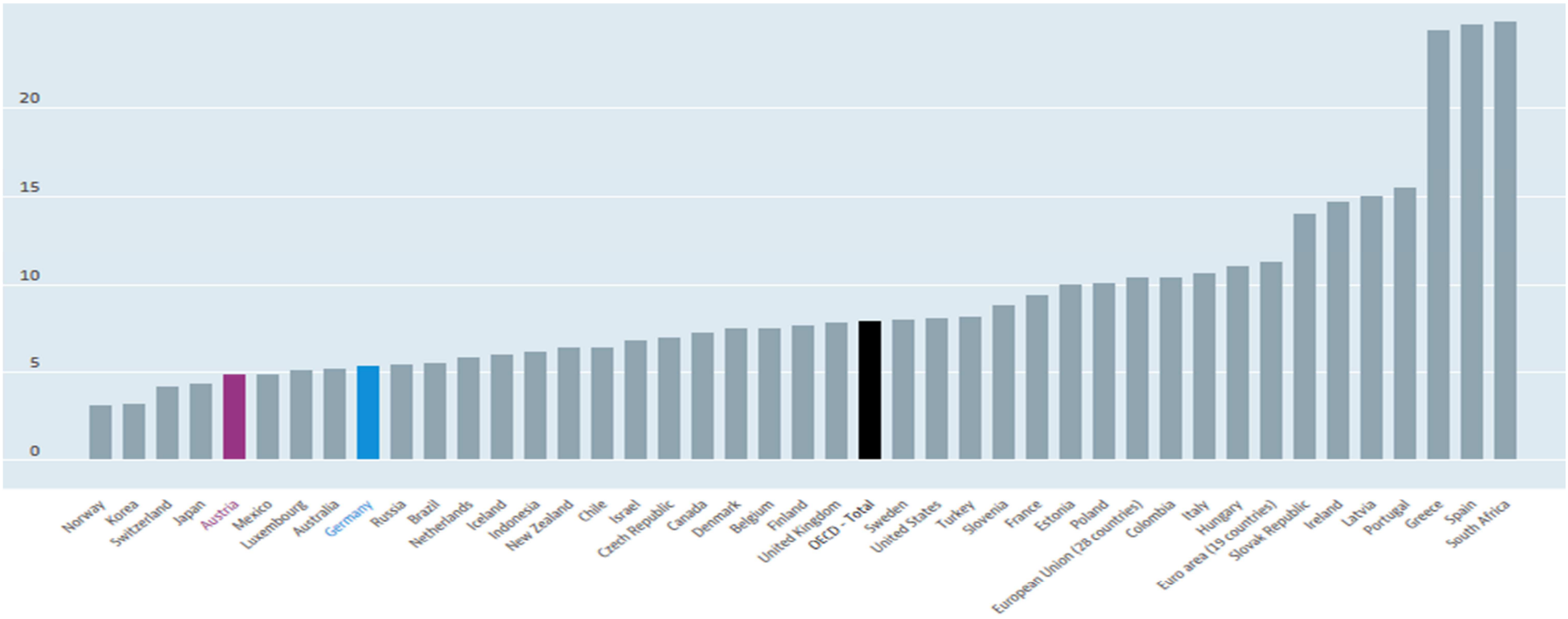
Figure 2: Splitting advantage for married couples under the current joint taxation system in Germany (2015/16)



Sources: Our computations.

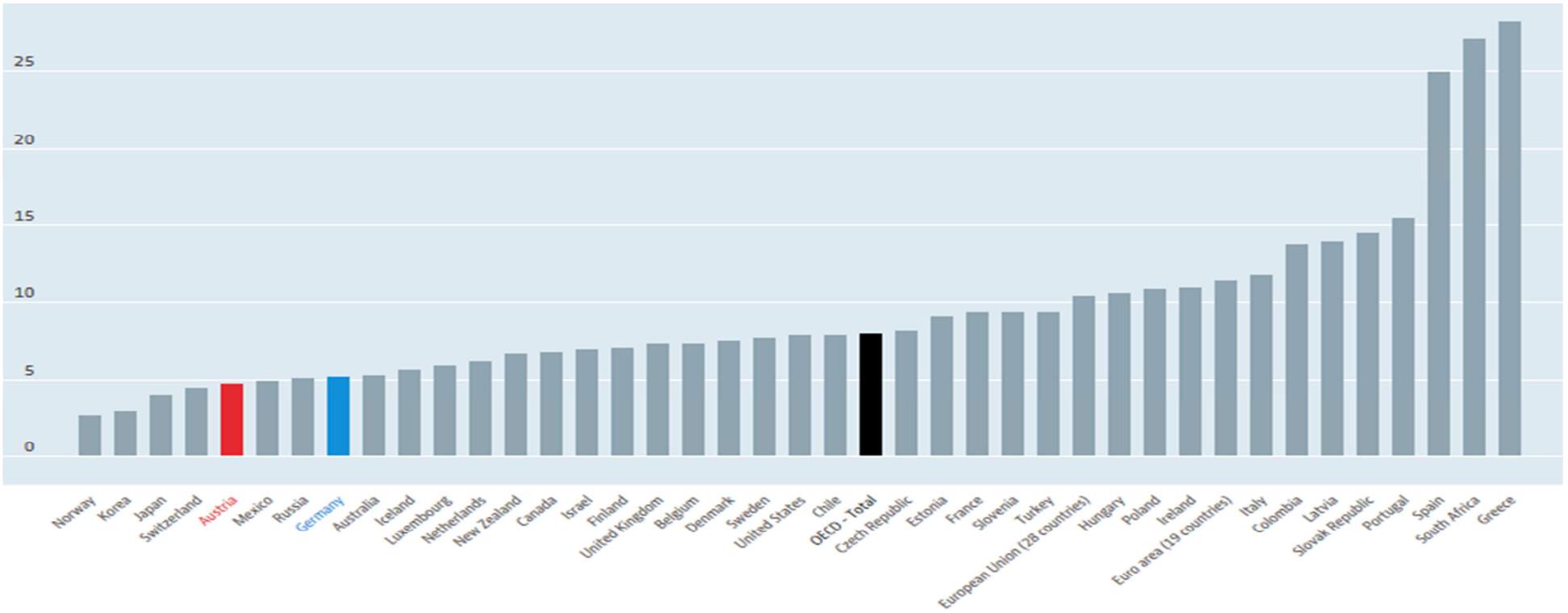
Note: The first (second) number indicates the primary's (secondary's) percentage share in the spouses' joint household pre-tax income per year (in euro).

Figure 3: Unemployment rate (Total, % of labour force, 2012)



Sources: OECD Labour Force Statistics

Figure 4: Unemployment rate (Women, % of labour force, 2012)



Sources: OECD Labour Force Statistics

Table 2: Statistics and source

	Obs	Mean	Standard Deviation	Obs	Mean	Standard Deviation	Source
	Austria			*Germany*			
Variables							
Participation decision (dummy = 1 if employed)	2321	0.9345	0.2474	4310	0.9297	0.2557	Eu-Silc (2008-2012)
Marginal participation (dummy= 1 if hours worked <= 15)	2321	0.1146	0.3186	4310	0.1794	0.3837	Eu-Silc (2008-2012)
Part-time participation (dummy = 1 if 15 < hours worked < 35)	2321	0.3512	0.4774	4310	0.3244	0.4682	Eu-Silc (2008-2012)
Full-time participation (dummy = 1 if hours worked >= 35)	2321	0.5343	0.4989	4310	0.4963	0.5000	Eu-Silc (2008-2012)
Highest women education level attained (ISCED)	2321	3.4580	0.9627	4310	3.7712	1.0357	Eu-Silc (2008-2012)
Highest spouse education level attained (ISCED)	1428	3.6071	0.9685	2726	3.9358	1.0482	Eu-Silc (2008-2012)
Unemployment status of spouse	1428	0.0364	0.1874	2726	0.0341	0.1816	Eu-Silc (2008-2012)
Age	2321	37.6032	8.5973	4310	38.3566	8.3847	Eu-Silc (2008-2012)

Sources: Our elaborations based on Eu-silc data.

Table 3: Share of participation in marginal/ part-time/ full-time jobs by age cohort

	Unemployment ratio	*Marginal participation*	*Part-time participation*	*Full-time participation*	*Unemployment ratio*	*Marginal participation*	*Part-time participation*	*Full-time participation*
Age Cohort	Complete women sample (%)				Only married women (%)			
Austria								
20 - 29	0.0805	0.1046	0.1670	0.7284	0.0714	0.1071	0.4107	0.7321
30 - 39	0.0820	0.1406	0.3675	0.4919	0.0593	0.1469	0.4605	0.3927
40 - 50	0.0491	0.1034	0.4216	0.4750	0.0317	0.1033	0.4725	0.4242
Germany								
20 - 29	0.0870	0.1292	0.1478	0.7230	0.1395	0.2403	0.2248	0.5349
30 - 39	0.0750	0.1746	0.3078	0.5175	0.0430	0.2034	0.3854	0.3682
40 - 50	0.0618	0.1999	0.3967	0.4034	0.0349	0.2305	0.4549	0.3146

Sources: Our elaborations based on Eu-silc data.

Figure 5: Female total hours worked per week in Austria and Germany

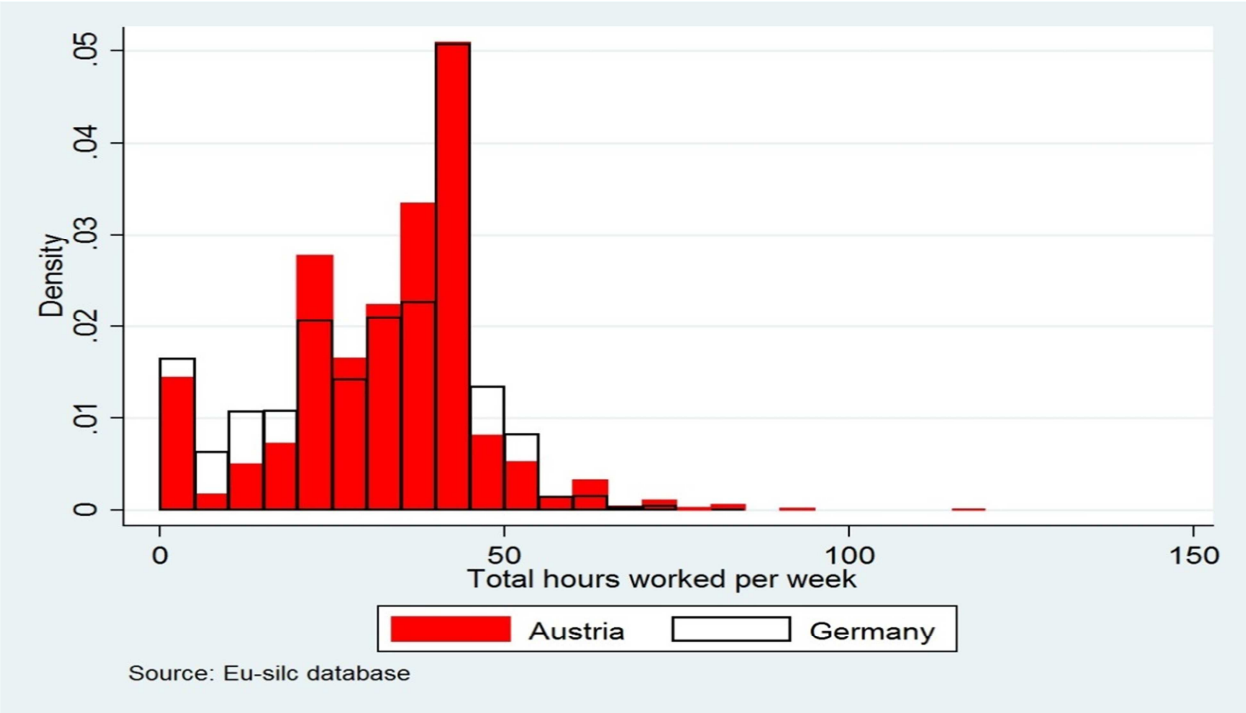


Figure 6: Female highest ISCED level attained in Austria and Germany

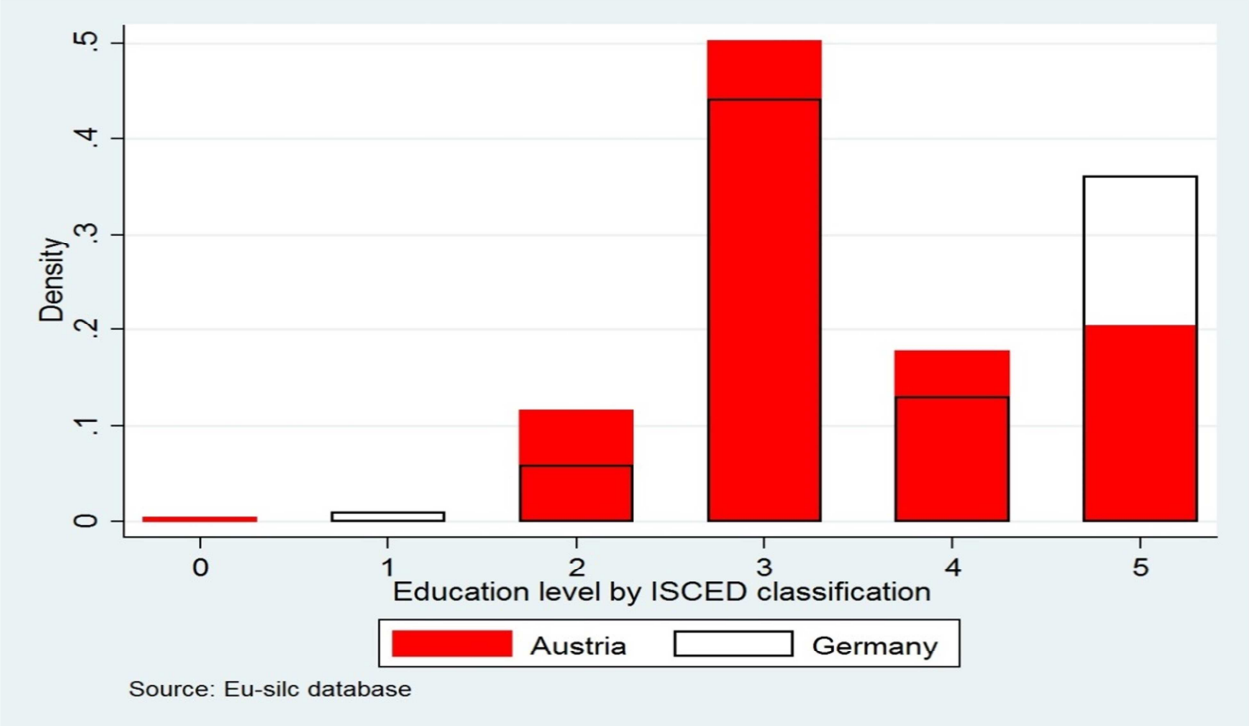


Table 4: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
participation decision						
	exposure to joint taxation	-0.0112	-0.0387***	-0.0204**	-0.0167**	-0.0188**
		[0.0123]	[0.0125]	[0.0087]	[0.0083]	[0.0081]
	age	0.007	0	0.0157*	0.0106	0.0111
		[0.0056]	[0.0055]	[0.0084]	[0.0081]	[0.0081]
	age square	-0.0001	0	-0.0002*	-0.0001	-0.0001
		[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
	age30	-0.0127	-0.0025	-0.0059	-0.0039	-0.0038
		[0.0191]	[0.0187]	[0.0233]	[0.0221]	[0.0221]
	women education		0.0462***	0.0300***	0.0257***	0.0255***
			[0.0036]	[0.0040]	[0.0038]	[0.0038]
	spouse education			0.0127***	0.0070**	0.0068**
				[0.0034]	[0.0033]	[0.0033]
	spouse in unemployment				-0.3143***	-0.3127***
					[0.0395]	[0.0397]
	immigrant status					-0.0159
						[0.0129]
	constant	0.7863***	0.7403***	0.4660***	0.6079***	0.6036***
		[0.0961]	[0.0949]	[0.1544]	[0.1472]	[0.1470]
exposure to joint taxation						
	german women (GE)	0.5208***	0.5192***	0.7982***	0.7987***	0.8033***
		[0.0075]	[0.0077]	[0.0079]	[0.0079]	[0.1171]
	age	0.0573***	0.0563***	0.0435***	0.0426***	0.0415***
		[0.0072]	[0.0073]	[0.0118]	[0.0118]	[0.0118]
	age square	-0.0007***	-0.0006***	-0.0005***	-0.0005***	-0.0004***
		[0.0001]	[0.0001]	[0.0001]	[0.0002]	[0.0002]
	age30	0.0706***	0.0715***	0.1044***	0.1048***	0.1053***
		[0.0255]	[0.0255]	[0.0364]	[0.0364]	[0.0364]
	women education		0.0051	-0.0053	-0.0060	-0.0055
			[0.0049]	[0.0057]	[0.0057]	[0.0057]
	spouse education			0.0030	0.0020	0.0029
				[0.0054]	[0.0054]	[0.0054]
	spouse in unemployment				-0.0548*	-0.0597**
					[0.0303]	[0.0304]
	immigrant status					0.0444***
						[0.0117]
	constant	-1.2439***	-1,2452***	-1.0539***	-1.0299***	-1.0239***
		[0.1213]	[0.1211]	[0.2101]	[0.2105]	[0.2106]
Observations		6631	6631	4154	4154	4154
Adjusted R-squared		.	0.0224	0.0374	0.1110	0.1114
Staiger-Stock test: First-stage F-statistics		1252.23***	1002.41***	2124.26***	1823.76***	1621.68***
CLR (LIML) confidence intervals		[-.036;.013]	[-.063;-.014]	[-.037;-.004]	[-.033;-.005]	[-.035;-.002]
Montiel-Pflueger Effective F-statistics (5%)		2776.81	2701.31	6125.92	6135.21	6065.62
Durbin-Wu-Hausman (chi-sq)		19.95***	38.98***	5.54**	2.90*	3.66*
Pagan-Hall (Chi-sq)		13.37**	262.76***	185.28***	334.30***	339.94***

Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 5: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
marginal participation						
	exposure to joint taxation	0.1216***	0.1646***	0.1432***	0.1406***	0.1428***
		[0.0168]	[0.0171]	[0.0146]	[0.0145]	[0.0147]
	age	-0.0063	0.0047	0.0018	0.0054	0.0049
		[0.0071]	[0.0070]	[0.0117]	[0.0116]	[0.0116]
	age square	0.0001	-0.0001	-0.0000	-0.0001	-0.0001
		[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
	age30	0.0238	0.0077	0.0063	0.0049	0.0049
		[0.0228]	[0.0222]	[0.0303]	[0.0296]	[0.0296]
	women education		-0.0723***	-0.0653***	-0.0623***	-0.0621***
			[0.0047]	[0.0066]	[0.0066]	[0.0066]
	spouse education			-0.0099	-0.0059	-0.0056
				[0.0063]	[0.0063]	[0.0063]
	spouse in unemployment				0.2210***	0.2192***
					[0.0412]	[0.0415]
	immigrant status					0.0169
						[0.0194]
	constant	0.2144*	0.2863**	0.3367	0.2370	0.2415
		[0.1210]	[0.1189]	[0.2099]	[0.2062]	[0.2062]
exposure to joint taxation						
	german women (GE)	0.5208***	0.5192***	0.7982***	0.7987***	0.8033***
		[0.0075]	[0.0077]	[0.0079]	[0.0079]	[0.1171]
	age	0.0573***	0.0563***	0.0435***	0.0426***	0.0415***
		[0.0072]	[0.0073]	[0.0118]	[0.0118]	[0.0118]
	age square	-0.0007***	-0.0006***	-0.0005***	-0.0005***	-0.0004***
		[0.0001]	[0.0001]	[0.0001]	[0.0002]	[0.0002]
	age30	0.0706***	0.0715***	0.1044***	0.1048***	0.1053***
		[0.0255]	[0.0255]	[0.0364]	[0.0364]	[0.0364]
	women education		0.0051	-0.0053	-0.0060	-0.0055
			[0.0049]	[0.0057]	[0.0057]	[0.0057]
	spouse education			0.0030	0.0020	0.0029
				[0.0054]	[0.0054]	[0.0054]
	spouse in unemployment				-0.0548*	-0.0597**
					[0.0303]	[0.0304]
	immigrant status					0.0444***
						[0.0117]
	constant	-1.2439***	-1,2452***	-1.0539***	-1.0299***	-1.0239***
		[0.1213]	[0.1211]	[0.2101]	[0.2105]	[0.2106]
Observations		6631	6631	4154	4154	4154
Adjusted R-squared		0.0169	0.0513	0.0583	0.0698	0.0699
Staiger-Stock test: First-stage F-statistics		1252.23***	1002.41***	2124.26***	1823.76***	1621.68***
CLR (LIML) confidence intervals		[.087;.157]	[.129;.199]	[.113;.173]	[.111;.170]	[.113;.173]
Montiel-Pflueger Effective F-statistics (5%)		2776.81	2701.31	6125.92	6135.21	6065.62
Durbin-Wu-Hausman (chi-sq)		2.52	12.43***	0.90	0.45	0.63
Pagan-Hall (Chi-sq)		60.43***	353.11***	234.08***	238.31***	240.96***

Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 6: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
part-time participation						
	exposure to joint taxation	-0.0686***	-0.0633***	-0.0525***	-0.0510**	-0.0542***
		[0.0230]	[0.0232]	[0.0202]	[0.0202]	[0.0202]
	age	0.0675***	0.0689***	0.1086***	0.1067***	0.1074***
		[0.0083]	[0.0084]	[0.0145]	[0.0145]	[0.0146]
	age square	-0.0008***	-0.0008***	-0.0013***	-0.0012***	-0.0012***
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	[0.0002]
	age30	-0.0157	-0.0177	-0.0881**	-0.0873**	-0.0873**
		[0.0271]	[0.0271]	[0.0399]	[0.0399]	[0.0399]
	women education		-0.0090	-0.0117	-0.0133	-0.0136*
			[0.0057]	[0.0082]	[0.0082]	[0.0082]
	spouse education			0.0186**	0.0164**	0.0160**
				[0.0080]	[0.0081]	[0.0081]
	spouse in unemployment				-0.1218***	-0.1193***
					[0.0378]	[0.0378]
	immigrant status					-0.0242
						[0.0242]
	constant	-1.0441***	-1.0351***	-1.7779***	-1.7229***	-1.7293***
		[0.1396]	[0.1395]	[0.2546]	[0.2554]	[0.2556]
exposure to joint taxation						
	german women (GE)	0.5208***	0.5192***	0.7982***	0.7987***	0.8033***
		[0.0075]	[0.0077]	[0.0079]	[0.0079]	[0.1171]
	age	0.0573***	0.0563***	0.0435***	0.0426***	0.0415***
		[0.0072]	[0.0073]	[0.0118]	[0.0118]	[0.0118]
	age square	-0.0007***	-0.0006***	-0.0005***	-0.0005***	-0.0004***
		[0.0001]	[0.0001]	[0.0001]	[0.0002]	[0.0002]
	age30	0.0706***	0.0715***	0.1044***	0.1048***	0.1053***
		[0.0255]	[0.0255]	[0.0364]	[0.0364]	[0.0364]
	women education		0.0051	-0.0053	-0.0060	-0.0055
			[0.0049]	[0.0057]	[0.0057]	[0.0057]
	spouse education			0.0030	0.0020	0.0029
				[0.0054]	[0.0054]	[0.0054]
	spouse in unemployment				-0.0548*	-0.0597**
					[0.0303]	[0.0304]
	immigrant status					0.0444***
						[0.0117]
	Constant	-1.2439***	-1,2452***	-1.0539***	-1.0299***	-1.0239***
		[0.1213]	[0.1211]	[0.2101]	[0.2105]	[0.2106]
Observations		6631	6631	4154	4154	4154
Adjusted R-squared		0.0320	0.0339	0.0311	0.0333	0.0331
Staiger-Stock test: First-stage F-statistics		1252.23***	1002.41***	2124.26***	1823.76***	1621.68***
CLR (LIML) confidence intervals		[-.114;-.024]	[-.109;-.018]	[-.092;-.013]	[-.091;-.012]	[-.094;-.015]
Montiel-Pflueger Effective F-statistics (5%)		2776.81	2701.31	6125.92	6135.21	6065.62
Durbin-Wu-Hausman (chi-sq)		61.15***	56.75***	32.32***	31.08***	32.50***
Pagan-Hall (Chi-sq)		662.71***	661.48***	346.15***	361.29***	351.15***

Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 7: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
full-time participation						
	exposure to joint taxation	-0.0530**	-0.1013***	-0.0907***	-0.0896***	-0.0886***
		[0.0240]	[0.0239]	[0.0201]	[0.0200]	[0.0202]
	age	-0.0612***	-0.0736***	-0.1105***	-0.1121***	-0.1123***
		[0.0095]	[0.0093]	[0.0155]	[0.0154]	[0.0155]
	age square	0.0007***	0.0008***	0.0013***	0.0013***	0.0013***
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	[0.0002]
	age30	-0.0080	0.0101	0.0817*	0.0824*	0.0824*
		[0.0314]	[0.0306]	[0.0433]	[0.0431]	[0.0432]
	women education		0.0813***	0.0770***	0.0756***	0.0757***
			[0.0058]	[0.0082]	[0.0082]	[0.0082]
	spouse education			-0.0087	-0.0105	-0.0103
				[0.0079]	[0.0079]	[0.0079]
	spouse in unemployment				-0.0991**	-0.0999**
					[0.0402]	[0.0402]
	immigrant status					0.0073
						[0.0251]
	constant	1.8297***	1.7488***	2.4411***	2.4859***	2.4878***
		[0.1615]	[0.1584]	[0.2750]	[0.2741]	[0.2743]
exposure to joint taxation						
	german women (GE)	0.5208***	0.5192***	0.7982***	0.7987***	0.8033***
		[0.0075]	[0.0077]	[0.0079]	[0.0079]	[0.1171]
	age	0.0573***	0.0563***	0.0435***	0.0426***	0.0415***
		[0.0072]	[0.0073]	[0.0118]	[0.0118]	[0.0118]
	age square	-0.0007***	-0.0006***	-0.0005***	-0.0005***	-0.0004***
		[0.0001]	[0.0001]	[0.0001]	[0.0002]	[0.0002]
	age30	0.0706***	0.0715***	0.1044***	0.1048***	0.1053***
		[0.0255]	[0.0255]	[0.0364]	[0.0364]	[0.0364]
	women education		0.0051	-0.0053	-0.0060	-0.0055
			[0.0049]	[0.0057]	[0.0057]	[0.0057]
	spouse education			0.0030	0.0020	0.0029
				[0.0054]	[0.0054]	[0.0054]
	spouse in unemployment				-0.0548*	-0.0597**
					[0.0303]	[0.0304]
	immigrant status					0.0444***
						[0.0117]
	constant	-1.2439***	-1,2452***	-1.0539***	-1.0299***	-1.0239***
		[0.1213]	[0.1211]	[0.2101]	[0.2105]	[0.2106]
Observations		6631	6631	4154	4154	4154
Adjusted R-squared		0.0707	0.1059	0.0845	0.0857	0.0856
Staiger-Stock test: First-stage F-statistics		1252.23***	1002.41***	2124.26***	1823.76***	1621.68***
CLR (LIML) confidence intervals		[-.099;-.006]	[-.148;-.055]	[-.129;-.052]	[-.128;-.051]	[-.128;-.049]
Montiel-Pflueger Effective F-statistics (5%)		2776.81	2701.31	6125.92	6135.21	6065.62
Durbin-Wu-Hausman (chi-sq)		39.57***	22.09***	25.49***	26.47***	26.82***
Pagan-Hall (Chi-sq)		170.49***	162.11***	46.99***	49.05***	58.83***

Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 8: The impact of joint taxation on female labor decision

Variables	Bivariate Probit	(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
participation decision						
	exposure to joint taxation			-0.1982**	-0.1544	-0.1852*
				[0.0989]	[0.1011]	[0.1040]
	age			0.0922	0.0568	0.0656
				[0.0675]	[0.0717]	[0.0717]
	age square			-0.0009	-0.0005	-0.0006
				[0.0009]	[0.0009]	[0.0009]
	age30			0.0193	0.0337	0.0352
				[0.1939]	[0.2020]	[0.2018]
	women education			0.3380***	0.3106***	0.3043***
				[0.0463]	[0.0468]	[0.0462]
	spouse education			0.1434***	0.0884**	0.0860**
				[0.0387]	[0.0397]	[0.0397]
	spouse in unemployment				-1.3118***	-1.2998***
					[0.1175]	[0.1184]
	immigrant status					-0.1629
						[0.1086]
	constant			-2.0220*	-1.0118	-1.1009
				[1.1875]	[1.2582]	[1.2545]
exposure to joint taxation						
	german women (GE)			7.6286***	7.7613***	7.8816***
				[0.0423]	[0.0442]	[0.0872]
	age			0.3542***	0.3504***	0.3478***
				[0.0640]	[0.0642]	[0.0645]
	age square			-0.0039***	-0.0039***	-0.0038***
				[0.0008]	[0.0008]	[0.0008]
	age30			0.1555	0.1608	0.1509
				[0.1499]	[0.1501]	[0.1508]
	women education			-0.0333	-0.0388	-0.0412
				[0.0326]	[0.0328]	[0.0329]
	spouse education			0.0159	0.0085	0.0153
				[0.0310]	[0.0311]	[0.0314]
	spouse in unemployment				-0.3338**	-0.3335**
					[0.1570]	[0.1579]
	immigrant status					0.3680***
						[0.1231]
	constant			-14.4368***	-14.4444***	-14.5666***
				[1.1397]	[1.1449]	[1.1527]
Observations				4,154	4,154	4,154
Log pseudolikelihood				-1867.82	-1809.20	-1803.52
Likelihood-ratio (LR) test P-value				0.0431	0.2887	0.2040
Wald test (joint significance) P-value				0.0000	0.0000	0.0000

Note: Model 1 and Model 2 are not reported because the convergence is not achieved. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 9: The impact of joint taxation on female labor decision

Variables	Bivariate Probit (Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
marginal participation					
	exposure to joint taxation	0.6903***	0.5892***	0.5726***	0.5905***
		[0.0836]	[0.0699]	[0.0697]	[0.0719]
	age	0.0183	0.0144	0.0309	0.0282
		[0.0327]	[0.0522]	[0.0524]	[0.0525]
	age square	-0.0003	-0.0002	-0.0004	-0.0004
		[0.0004]	[0.0006]	[0.0007]	[0.0007]
	age30	0.0422	0.0369	0.0356	0.033
		[0.1100]	[0.1444]	[0.1445]	[0.1442]
	women education	-0.3093***	-0.2680***	-0.2572***	-0.2560***
		[0.0206]	[0.0276]	[0.0279]	[0.0278]
	spouse education		-0.0462*	-0.0284	-0.0273
			[0.0256]	[0.0258]	[0.0258]
	spouse in unemployment			0.6935***	0.6864***
				[0.1157]	[0.1164]
	immigrant status				0.0822
					[0.0783]
	constant	-0.4839	-0.4746	-0.9238	-0.9003
		[0.5593]	[0.9351]	[0.9390]	[0.9385]
exposure to joint taxation					
	german women (GE)	6.7567***	7.4242***	7.7487***	7.5564***
		[0.0258]	[0.0424]	[0.0444]	[0.0863]
	age	0.4232***	0.3559***	0.3510***	0.3484***
		[0.0380]	[0.0641]	[0.0643]	[0.0646]
	age square	-0.0050***	-0.0039***	-0.0039***	-0.0038***
		[0.0005]	[0.0008]	[0.0008]	[0.0008]
	age30	0.065	0.1613	0.1652	0.1555
		[0.1128]	[0.1501]	[0.1502]	[0.1509]
	women education	-0.0029	-0.0338	-0.0387	-0.0414
		[0.0198]	[0.0326]	[0.0328]	[0.0329]
	spouse education		0.0165	0.0086	0.0154
			[0.0310]	[0.0311]	[0.0315]
	spouse in unemployment			-0.3442**	-0.3451**
				[0.1560]	[0.1568]
	immigrant status				0.3709***
					[0.1239]
	Constant	-15.3204***	-14.2697***	-14.4443***	-14.2572***
		[0.6718]	[1.1419]	[1.1468]	[1.1545]
Observations		6,631	4,154	4,154	4,154
Log pseudolikelihood		-5299.45	-2933.40	-2912.24	2907.22
Likelihood-ratio (LR) test P-value		0.0007	0.5629	0.9627	0.8115
Wald test (joint significance) P-value		0.0000	0.0000	0.0000	0.0000

Note: Model 1 is not reported because the convergence is not achieved. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table10: The impact of joint taxation on female labor decision

Variables	Bivariate Probit					
		(Model1)	(Model2)	(Model 3)	(Model 4)	(Model 5)
part time participation						
	exposure to joint taxation			-0.1165**	-0.1120**	-0.1194**
				[0.0506]	[0.0507]	[0.0509]
	age			0.3290***	0.3251***	0.3264***
				[0.0462]	[0.0463]	[0.0463]
	age square			-0.0038***	-0.0038***	-0.0038***
				[0.0006]	[0.0006]	[0.0006]
	age30			-0.2664**	-0.2659**	-0.2648**
				[0.1207]	[0.1207]	[0.1208]
	women education			-0.0313	-0.0361*	-0.0367*
				[0.0219]	[0.0219]	[0.0219]
	spouse education			0.0488**	0.0429**	0.0418*
				[0.0213]	[0.0214]	[0.0215]
	spouse in unemployment				-0.3617***	-0.3549***
					[0.1182]	[0.1183]
	immigrant status					-0.0567
						[0.0658]
	constant			-6.8276***	-6.7063***	-6.7132***
				[0.8336]	[0.8365]	[0.8359]
exposure to joint taxation						
	german women (GE)			7.5922***	7.6051***	7.7082***
				[0.0421]	[0.0438]	[0.0824]
	age			0.3511***	0.3468***	0.3437***
				[0.0640]	[0.0642]	[0.0644]
	age square			-0.0039***	-0.0038***	-0.0038***
				[0.0008]	[0.0008]	[0.0008]
	age30			0.1658	0.1673	0.159
				[0.1490]	[0.1492]	[0.1498]
	women education			-0.0275	-0.0331	-0.0356
				[0.0321]	[0.0323]	[0.0324]
	spouse education			0.0108	0.0034	0.0101
				[0.0307]	[0.0308]	[0.0311]
	spouse in unemployment				-0.3258**	-0.3288**
					[0.1551]	[0.1556]
	immigrant status					0.3737***
						[0.1234]
	constant			-14.3378***	-14.2115***	-14.3093***
				[1.1400]	[1.1444]	[1.1523]
Observations				4,154	4,154	4,154
Log pseudolikelihood				-3865.62	-3859.09	-3853.84
Likelihood-ratio (LR) test P-value				0.0000	0.0000	0.0000
Wald test (joint significance) P-value				0.0000	0.0000	0.0000

Note: Model 1 and Model 2 are not reported because the convergence is not achieved. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 11: The impact of joint taxation on female labour decision

Variables	Bivariate Probit					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
full time participation						
	exposure to joint taxation	-0.1793***	-0.3113***	-0.2672***	-0.2611***	-0.2610***
		[0.0598]	[0.0631]	[0.0541]	[0.0542]	[0.0545]
	age	-0.1744***	-0.2063***	-0.3011***	-0.3062***	-0.3066***
		[0.0284]	[0.0283]	[0.0451]	[0.0450]	[0.0450]
	age square	0.0020***	0.0024***	0.0035***	0.0036***	0.0036***
		[0.0004]	[0.0004]	[0.0006]	[0.0006]	[0.0006]
	age30	0.0006	0.0399	0.2296*	0.2312*	0.2311*
		[0.0865]	[0.0874]	[0.1182]	[0.1180]	[0.1181]
	women education		0.2197***	0.2083***	0.2048***	0.2052***
			[0.0163]	[0.0226]	[0.0227]	[0.0227]
	spouse education			-0.0235	-0.0287	-0.0283
				[0.0217]	[0.0219]	[0.0219]
	spouse in unemployment				-0.2797**	-0.2824**
					[0.1197]	[0.1198]
	immigrant status					0.0205
						[0.0672]
	constant	3.7472***	3.5182***	5.2950***	5.4298***	5.4309***
		[0.4917]	[0.4877]	[0.8065]	[0.8054]	[0.8056]
exposure to joint taxation						
	german women (GE)	6.7655***	6.8749***	7.5969***	7.5927***	7.6807***
		[0.0239]	[0.0250]	[0.0408]	[0.0427]	[0.0773]
	age	0.4152***	0.4194***	0.3503***	0.3447***	0.3415***
		[0.0376]	[0.0381]	[0.0642]	[0.0644]	[0.0646]
	age square	-0.0049***	-0.0050***	-0.0039***	-0.0038***	-0.0038***
		[0.0005]	[0.0005]	[0.0008]	[0.0008]	[0.0008]
	age30	0.1016	0.0933	0.1771	0.1795	0.173
		[0.1128]	[0.1131]	[0.1495]	[0.1494]	[0.1501]
	women education		0.0046	-0.0235	-0.0294	-0.0312
			[0.0198]	[0.0324]	[0.0326]	[0.0327]
	spouse education			0.015	0.0056	0.012
				[0.0310]	[0.0310]	[0.0313]
	spouse in unemployment				-0.3589**	-0.3630**
					[0.1522]	[0.1525]
	immigrant status					0.3530***
						[0.1221]
	constant	-15.1779***	-15.3869***	-14.3567***	-14.1804***	-14.2611***
		[0.6677]	[0.6743]	[1.1441]	[1.1482]	[1.1555]
Observations		6,631	6,631	4,154	4,154	4,154
Log pseudolikelihood		-6900.87	-6803.83	-3805.89	-3799.37	-3795.12
Likelihood-ratio (LR) test P-value		0.0000	0.0000	0.0000	0.0000	0.0000
Wald test (joint significance) P-value		0.0000	0.0000	0.0000	0.0000	0.0000

Note: Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 12: Estimated effect of joint taxation on female labour decision

Two-Stage Least Square estimations				
Dependent variable				
	Participation decision	Marginal participation	Part-time participation	Full-time participation
	(1)	(2)	(3)	(4)
	Panel A. Covariates: age/age square/dummy age>30			
Treatment effect	-0.0112	0.1216***	-0.0686***	-0.0530**
	[0.0123]	[0.0168]	[0.0230]	[0.0240]
	Panel B. Covariates: Panel A covariates plus women education			
Treatment effect	-0.0387***	0.1646***	-0.0633***	-0.1013***
	[0.0125]	[0.0171]	[0.0232]	[0.0239]
	Panel C. Covariates: Panel B covariates plus spouse education			
Treatment effect	-0.0204**	0.1432***	-0.0525***	-0.0907***
	[0.0087]	[0.0146]	[0.0202]	[0.0201]
	Panel D. Covariates: Panel C covariates plus spouse unemployment dummy			
Treatment effect	-0.0167**	0.1406***	-0.0510**	-0.0896***
	[0.0083]	[0.0145]	[0.0202]	[0.0200]
	Panel E. Covariates: Panel D covariates plus immigrant status			
Treatment effect	-0.0188**	0.1428***	-0.0542***	-0.0886***
	[0.0081]	[0.0147]	[0.0202]	[0.0202]

Robust standard errors in bracket. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 13: Estimated effect of joint taxation on female labour decision

Bivariate Probit estimations				
Dependent variable				
	Participation decision	Marginal participation	Part-time participation	Full-time participation
	(1)	(2)	(3)	(4)
	Panel A. Covariates: age/age square/dummy age>30			
Treatment effect	_____	_____	_____	-0.0679***
	_____	_____	_____	[0.0225]
	Panel B. Covariates: Panel A covariates plus women education			
Treatment effect	_____	0.1557***	_____	-0.1147***
	_____	[0.0190]	_____	[0.0228]
	Panel C. Covariates: Panel B covariates plus spouse education			
Treatment effect	-0.0178**	0.1393***	-0.0436**	-0.0985***
	[0.0091]	[0.0162]	[0.0189]	[0.0197]
	Panel D. Covariates: Panel C covariates plus spouse unemployment dummy			
Treatment effect	-0.0126	0.1337***	-0.0419**	-0.0962***
	[0.0084]	[0.0159]	[0.0189]	[0.0197]
	Panel E. Covariates: Panel D covariates plus immigrant status			
Treatment effect	-0.0151*	0.1378***	-0.0446**	-0.0961***
	[0.0087]	[0.0165]	[0.0190]	[0.0198]

Robust standard errors in bracket. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Appendix: Robustness checks

Table 14: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
participation decision						
	exposure to joint taxation	-0.0359***	-0.0584***	-0.0299***	-0.0204**	
		[0.0126]	[0.0129]	[0.0089]	[0.0083]	
	age	0.0075	0.0002	0.0136	0.0055	
		[0.0057]	[0.0056]	[0.0086]	[0.0081]	
	age square	-0.0001	0.0000	-0.0001	-0.0000	
		[0.0001]	[0.0001]	[0.0001]	[0.0001]	
	age30	-0.0069	0.0022	0.0050	0.0112	
		[0.0198]	[0.0194]	[0.0246]	[0.0228]	
	women education		0.0443***	0.0264***	0.0211***	
			[0.0037]	[0.0040]	[0.0037]	
	spouse education			0.0139***	0.0074**	
				[0.0036]	[0.0033]	
	spouse in unemployment				-0.3693***	
					[0.0462]	
	immigrant status					
	constant	0.7870***	0.7504***	0.5202***	0.7221***	
		[0.0979]	[0.0968]	[0.1560]	[0.1448]	
exposure to joint taxation						
	german women (GE)	0.5102***	0.5082***	0.7943***	0.7956***	
		[0.0078]	[0.0079]	[0.0084]	[0.0083]	
	age	0.0566***	0.0552***	0.0448***	0.0435***	
		[0.0076]	[0.0077]	[0.0129]	[0.0129]	
	age square	-0.0006***	-0.0006***	-0.0005**	-0.0005**	
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	
	age30	0.0703**	0.0715**	0.1114**	0.1127**	
		[0.0274]	[0.0274]	[0.0403]	[0.0403]	
	women education		0.0075	-0.0033	-0.0042	
			[0.0053]	[0.0063]	[0.0063]	
	spouse education			0.0031	0.0020	
				[0.0059]	[0.0059]	
	spouse in unemployment				-0.0619	
					[0.0382]	
	immigrant status					
	constant	-1.2414***	-1,2429***	-1.1060***	-1.0739***	
		[0.1276]	[0.1273]	[0.2290]	[0.2304]	
Observations		5977	5977	3697	3697	
Adjusted R-squared		.	0.0096	0.0329	0.1298	
Staiger-Stock test: First-stage F-statistics		739.07***	591.80***	924.31***	793.49***	
CLR (LIML) confidence intervals		[-.062;-.009]	[-.084;-.032]	[-.047;-.012]	[-.037;-.004]	
Montiel-Pflueger Effective F-statistics (5%)		2218.09	2168.77	4926.26	4930.71	
Durbin-Wu-Hausman (chi-sq)		34.94***	52.94***	9.57***	4.09**	
Pagan-Hall (Chi-sq)		17.37***	226.46***	153.68***	285.25***	

Note: Sample is restricted to native individuals. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 15: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
marginal participation						
	exposure to joint taxation	0.1320***	0.1690***	0.1397***	0.1328***	
		[0.0179]	[0.0183]	[0.0156]	[0.0154]	
	age	-0.0033	0.0086	0.0101	0.0159	
		[0.0072]	[0.0071]	[0.0121]	[0.0118]	
	age square	0.0000	-0.0001	-0.0001	-0.0002	
		[0.0001]	[0.0001]	[0.0002]	[0.0001]	
	age30	0.0095	-0.0056	-0.0124	-0.0169	
		[0.0235]	[0.0229]	[0.0316]	[0.0305]	
	women education		-0.0727***	-0.0683***	-0.0645***	
			[0.0049]	[0.0069]	[0.0069]	
	spouse education			-0.0106	-0.0059	
				[0.0067]	[0.0067]	
	spouse in unemployment				0.2668***	
					[0.0482]	
	immigrant status					
	constant	0.1495	0.2095*	0.1887	0.0429	
		[0.1223]	[0.1206]	[0.2153]	[0.2094]	
exposure to joint taxation						
	german women (GE)	0.5102***	0.5082***	0.7943***	0.7956***	
		[0.0078]	[0.0079]	[0.0084]	[0.0083]	
	age	0.0566***	0.0552***	0.0448***	0.0435***	
		[0.0076]	[0.0077]	[0.0129]	[0.0129]	
	age square	-0.0006***	-0.0006***	-0.0005**	-0.0005**	
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	
	age30	0.0703**	0.0715**	0.1114**	0.1127**	
		[0.0274]	[0.0274]	[0.0403]	[0.0403]	
	women education		0.0075	-0.0033	-0.0042	
			[0.0053]	[0.0063]	[0.0063]	
	spouse education			0.0031	0.0020	
				[0.0059]	[0.0059]	
	spouse in unemployment				-0.0619	
					[0.0382]	
	immigrant status					
	constant	-1.2414***	-1,2429***	-1.1060***	-1.0739***	
		[0.1276]	[0.1273]	[0.2290]	[0.2304]	
Observations		5977	5977	3697	3697	
Adjusted R-squared		0.0169	0.0510	0.0609	0.0755	
Staiger-Stock test: First-stage F-statistics		739.07***	591.80***	924.31***	793.49***	
CLR (LIML) confidence intervals		[.094;.169]	[.132;.207]	[.108;.172]	[.101;.165]	
Montiel-Pflueger Effective F-statistics (5%)		2218.09	2168.77	4926.26	4930.71	
Durbin-Wu-Hausman (chi-sq)		4.68**	13.65***	1.18	0.39	
Pagan-Hall (Chi-sq)		66.05***	331.48***	226.28***	222.21***	

Note: Sample is restricted to native individuals. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 16: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
part-time employment						
	exposure to joint taxation	-0.0744***	-0.0690***	-0.0508**	-0.0478**	
		[0.0252]	[0.0254]	[0.0218]	[0.0218]	
	age	0.0684***	0.0701***	0.1135***	0.1108***	
		[0.0087]	[0.0088]	[0.0154]	[0.0155]	
	age square	-0.0008***	-0.0008***	-0.0013***	-0.0013***	
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	
	age30	-0.0331	-0.0353	-0.1178***	-0.1158***	
		[0.0286]	[0.0286]	[0.0428]	[0.0427]	
	women education		-0.0105*	-0.0156*	-0.0173**	
			[0.0061]	[0.0088]	[0.0088]	
	spouse education			0.0276***	0.0255***	
				[0.0086]	[0.0086]	
	spouse in unemployment				-0.1190***	
					[0.0421]	
	immigrant status					
	constant	-1.0667***	-1.0580***	-1.8926***	-1.8275***	
		[0.1460]	[0.1458]	[0.2702]	[0.2716]	
exposure to joint taxation						
	german women (GE)	0.5102***	0.5082***	0.7943***	0.7956***	
		[0.0078]	[0.0079]	[0.0084]	[0.0083]	
	age	0.0566***	0.0552***	0.0448***	0.0435***	
		[0.0076]	[0.0077]	[0.0129]	[0.0129]	
	age square	-0.0006***	-0.0006***	-0.0005**	-0.0005**	
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	
	age30	0.0703**	0.0715**	0.1114**	0.1127**	
		[0.0274]	[0.0274]	[0.0403]	[0.0403]	
	women education		0.0075	-0.0033	-0.0042	
			[0.0053]	[0.0063]	[0.0063]	
	spouse education			0.0031	0.0020	
				[0.0059]	[0.0059]	
	spouse in unemployment				-0.0619	
					[0.0382]	
	immigrant status					
	constant	-1.2414***	-1,2429***	-1.1060***	-1.0739***	
		[0.1276]	[0.1273]	[0.2290]	[0.2304]	
Observations		5977	5977	3697	3697	
Adjusted R-squared		0.0326	0.0348	0.0357	0.0378	
Staiger-Stock test: First-stage F-statistics		739.07***	591.80***	924.31***	793.49***	
CLR (LIML) confidence intervals		[-.124;-.025]	[-.119;-.019]	[-.094;-.008]	[-.091;-.005]	
Montiel-Pflueger Effective F-statistics (5%)		2218.09	2168.77	4926.26	4930.71	
Durbin-Wu-Hausman (chi-sq)		58.35***	55.47***	29.77***	28.18***	
Pagan-Hall (Chi-sq)		615.51***	615.43***	319.26***	336.72***	

Note: Sample is restricted to native individuals. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 17: The impact of joint taxation on female labour decision

Variables	Two-Stage Least Square					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
full-time employment						
	exposure to joint taxation	-0.0576**	-0.1000***	-0.0889***	-0.0850***	
		[0.0262]	[0.0261]	[0.0217]	[0.0216]	
	age	-0.0650***	-0.0787***	-0.1235***	-0.1268***	
		[0.0098]	[0.0097]	[0.0162]	[0.0161]	
	age square	0.0007***	0.0009***	0.0014***	0.0015***	
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	
	age30	0.0237	0.0409	0.1302***	0.1327***	
		[0.0329]	[0.0321]	[0.0463]	[0.0460]	
	women education		0.0832***	0.0840***	0.0819***	
			[0.0062]	[0.0087]	[0.0087]	
	spouse education			-0.0170**	-0.0196**	
				[0.0084]	[0.0084]	
	spouse in unemployment				-0.1478***	
					[0.0444]	
	immigrant status					
	constant	1.9172***	1.8485***	2.7038***	2.7846***	
		[0.1663]	[0.1636]	[0.2876]	[0.2859]	
exposure to joint taxation						
	german women (GE)	0.5102***	0.5082***	0.7943***	0.7956***	
		[0.0078]	[0.0079]	[0.0084]	[0.0083]	
	age	0.0566***	0.0552***	0.0448***	0.0435***	
		[0.0076]	[0.0077]	[0.0129]	[0.0129]	
	age square	-0.0006***	-0.0006***	-0.0005**	-0.0005**	
		[0.0001]	[0.0001]	[0.0002]	[0.0002]	
	age30	0.0703**	0.0715**	0.1114**	0.1127**	
		[0.0274]	[0.0274]	[0.0403]	[0.0403]	
	women education		0.0075	-0.0033	-0.0042	
			[0.0053]	[0.0063]	[0.0063]	
	spouse education			0.0031	0.0020	
				[0.0059]	[0.0059]	
	spouse in unemployment				-0.0619	
					[0.0382]	
	immigrant status					
	constant	-1.2414***	-1,2429***	-1.1060***	-1.0739***	
		[0.1276]	[0.1273]	[0.2290]	[0.2304]	
Observations		5977	5977	3697	3697	
Adjusted R-squared		0.0801	0.1151	0.0972	0.0993	
Staiger-Stock test: First-stage F-statistics		739.07***	591.80***	924.31***	793.49***	
CLR (LIML) confidence intervals		[-.108;-.006]	[-.150;-.049]	[-.131;-.047]	[-.127;-.043]	
Montiel-Pflueger Effective F-statistics (5%)		2218.09	2168.77	4926.26	4930.71	
Durbin-Wu-Hausman (chi-sq)		33.25***	20.62***	22.39***	24.41**	
Pagan-Hall (Chi-sq)		162.71***	145.11***	43.64***	47.06***	

Note: Sample is restricted to native individuals. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 18: The impact of joint taxation on female labor decision

Variables	Bivariate Probit					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
participation decision						
	exposure to joint taxation			-0.3132***	-0.2197*	
				[0.1191]	[0.1208]	
	age			0.0911	0.0291	
				[0.0740]	[0.0795]	
	age square			-0.0009	-0.0001	
				[0.0009]	[0.0010]	
	age30			0.1107	0.1668	
				[0.2175]	[0.2254]	
	women education			0.3250***	0.2824***	
				[0.0517]	[0.0519]	
	spouse education			0.1652***	0.1043**	
				[0.0428]	[0.0448]	
	spouse in unemployment				-1.4691***	
					[0.1320]	
	immigrant status					
	constant			-1.9892	-0.4262	
				[1.2923]	[1.3856]	
exposure to joint taxation						
	german women (GE)			7.5853***	7.6576***	
				[0.0427]	[0.0445]	
	age			0.3450***	0.3416***	
				[0.0658]	[0.0660]	
	age square			-0.0038***	-0.0037***	
				[0.0008]	[0.0008]	
	age30			0.1338	0.1418	
				[0.1563]	[0.1566]	
	women education			-0.0254	-0.0305	
				[0.0341]	[0.0343]	
	spouse education			0.0157	0.008	
				[0.0323]	[0.0325]	
	spouse in unemployment				-0.3361**	
					[0.1666]	
	immigrant status					
	constant			-14.2706***	-14.2270***	
				[1.1703]	[1.1767]	
Observations				3697	3697	
Log pseudolikelihood				-1682.16	-1624.48	
Likelihood-ratio (LR) test P-value				0.0139	0.1966	
Wald test (joint significance) P-value				0.0000	0.0000	

Note: Sample is restricted to native individuals. Model 1 and Model 2 are not reported because the convergence is not achieved. Since "immigrant status" variable is omitted because of collinearity Model 5 is not reported. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 19: The impact of joint taxation on female labor decision

Variables	Bivariate Probit				
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
marginal participation						
	exposure to joint taxation		0.7343***	0.5762***	0.5432***	
			[0.0983]	[0.0773]	[0.0765]	
	age		0.0372	0.053	0.0807	
			[0.0350]	[0.0568]	[0.0570]	
	age square		-0.0005	-0.0006	-0.001	
			[0.0004]	[0.0007]	[0.0007]	
	age30		-0.0226	-0.0502	-0.0632	
			[0.1182]	[0.1587]	[0.1582]	
	women education		-0.3167***	-0.2841***	-0.2703***	
			[0.0223]	[0.0300]	[0.0303]	
	spouse education			-0.0492*	0.0079	
				[0.0275]	[0.0325]	
	spouse in unemployment				0.8053***	
					[0.1332]	
	immigrant status					
	constant		-0.836	-1.1613	-1.8402*	
			[0.5982]	[1.0181]	[1.0214]	
exposure to joint taxation						
	german women (GE)		6.7710***	7.3992***	7.4749***	
			[0.0259]	[0.0428]	[0.0445]	
	age		0.4172***	0.3475***	0.3431***	
			[0.0392]	[0.0660]	[0.0662]	
	age square		-0.0049***	-0.0038***	-0.0038***	
			[0.0005]	[0.0008]	[0.0008]	
	age30		0.0328	0.1425	0.147	
			[0.1168]	[0.1566]	[0.1567]	
	women education		0.0038	-0.0249	-0.0294	
			[0.0206]	[0.0342]	[0.0344]	
	spouse education			0.0163	-0.0288	
				[0.0323]	[0.0278]	
	spouse in unemployment				-0.3425**	
					[0.1652]	
	immigrant status					
	constant		-15.2864***	-14.1446***	-14.0782***	
			[0.6932]	[1.1741]	[1.1794]	
Observations			5977	3697	3697	
Log pseudolikelihood			-4820.69	-2655.47	-2633.33	
Likelihood-ratio (LR) test P-value			0.0003	0.5292	0.9005	
Wald test (joint significance) P-value			0.0000	0.0000	0.0000	

Note: Sample is restricted to native individuals. Model 1 is not reported because the convergence is not achieved. Since "immigrant status" variable is omitted because of collinearity Model 5 is not reported. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 20: The impact of joint taxation on female labor decision

Variables	Bivariate Probit					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
part-time participation						
	exposure to joint taxation			-0.1176**	-0.1082**	
				[0.0546]	[0.0547]	
	age			0.3474***	0.3417***	
				[0.0496]	[0.0498]	
	age square			-0.0040***	-0.0040***	
				[0.0006]	[0.0006]	
	age30			-0.3560***	-0.3520***	
				[0.1304]	[0.1304]	
	women education			-0.0420*	-0.0469**	
				[0.0236]	[0.0236]	
	spouse education			0.0732***	0.0675***	
				[0.0228]	[0.0230]	
	spouse in unemployment				-0.3694***	
					[0.1351]	
	immigrant status					
	constant			-7.2378***	-7.0887***	
				[0.8980]	[0.9018]	
exposure to joint taxation						
	german women (GE)			7.5769***	7.5902***	
				[0.0426]	[0.0443]	
	age			0.3448***	0.3406***	
				[0.0659]	[0.0661]	
	age square			-0.0038***	-0.0037***	
				[0.0008]	[0.0008]	
	age30			0.1433	0.1453	
				[0.1555]	[0.1557]	
	women education			-0.0199	-0.0255	
				[0.0336]	[0.0338]	
	spouse education			0.0126	0.0047	
				[0.0319]	[0.0320]	
	spouse in unemployment				-0.3274**	
					[0.1652]	
	immigrant status					
	constant			-14.2683***	-14.1424***	
				[1.1724]	[1.1773]	
Observations				3697	3697	
Log pseudolikelihood				-3496.24	-3490.95	
Likelihood-ratio (LR) test P-value				0.0000	0.0000	
Wald test (joint significance) P-value				0.0000	0.0000	

Note: Sample is restricted to native individuals. Model 1 and Model 2 are not reported because the convergence is not achieved. Since "immigrant status" variable is omitted because of collinearity Model 5 is not reported. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * p<0.1; ** p<0.05; *** p<0.01

Table 21: The impact of joint taxation on female labor decision

Variables	Bivariate Probit					
		(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
full-time employment						
	exposure to joint taxation	-0.1859***	-0.3034***	-0.2569***	-0.2443***	
		[0.0654]	[0.0692]	[0.0588]	[0.0588]	
	age	-0.1901***	-0.2252***	-0.3431***	-0.3537***	
		[0.0300]	[0.0300]	[0.0487]	[0.0486]	
	age square	0.0021***	0.0026***	0.0040***	0.0041***	
		[0.0004]	[0.0004]	[0.0006]	[0.0006]	
	age30	0.0897	0.1307	0.3676***	0.3755***	
		[0.0917]	[0.0929]	[0.1280]	[0.1277]	
	women education		0.2256***	-0.0153	-0.0209	
			[0.0175]	[0.0340]	[0.0342]	
	spouse education			-0.0473**	0.006	
				[0.0234]	[0.0323]	
	spouse in unemployment				-0.4373***	
					[0.1436]	
	immigrant status					
	constant	4.0837***	3.8803***	6.1284***	6.3832***	
		[0.5209]	[0.5173]	[0.8729]	[0.8713]	
exposure to joint taxation						
	german women (GE)	6.7672***	6.8123***	7.5578***	7.5749***	
		[0.0246]	[0.0252]	[0.0411]	[0.0431]	
	age	0.4116***	0.4146***	0.3438***	0.3392***	
		[0.0388]	[0.0394]	[0.0660]	[0.0661]	
	age square	-0.0048***	-0.0049***	-0.0038***	-0.0037***	
		[0.0005]	[0.0005]	[0.0008]	[0.0008]	
	age30	0.0687	0.0618	0.1561	0.1575	
		[0.1172]	[0.1174]	[0.1557]	[0.1557]	
	women education		0.0134	0.2293***	0.2243***	
			[0.0205]	[0.0244]	[0.0245]	
	spouse education			0.016	-0.0555**	
				[0.0322]	[0.0235]	
	spouse in unemployment				-0.3465**	
					[0.1607]	
	immigrant status					
	constant	-15.1508***	-15.3035***	-14.2592***	-14.1208***	
		[0.6891]	[0.6967]	[1.1757]	[1.1798]	
Observations		5977	5977	3697	3697	
Log pseudolikelihood		-6287.49	-6197.29	-3422.79	-3413.95	
Likelihood-ratio (LR) test P-value		0.0000	0.0000	0.0000	0.0000	
Wald test (joint significance) P-value		0.0000	0.0000	0.0000	0.0000	

Note: Sample is restricted to native individuals. Since "immigrant status" variable is omitted because of collinearity Model 5 is not reported. Diagnostics of weak instruments and endogeneity are reported for the two-stage least square estimations. Robust standard errors in bracket; asterisks denote significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$