Essays on Inequality and Fiscal Policy

Weijie Luo

PhD

University of York

Economics

March 2018

Abstract

This PhD thesis gathers three essays on income inequality and fiscal policy. Chapter 2, "Inequality and the Size of Government", written with Andrew Pickering and Paulo Santos Monteiro, revisits Meltzer and Richard (1981) but with the twist that income inequality is induced by differences in capital income as well as differences in labor productivity. When capital income is difficult to tax, as often observed, then greater capital income inequality leads to reduced demands for tax as the poor cannot effectuate redistribution. Using OECD data, government size and capital income inequality (proxied by the top 1 percent income share) are found to be negatively related in both fixed effects and instrumental variable regressions.

Chapter 3, "Inequality and Growth in the Twenty-First Century", builds on chapter 2 to investigate how economic growth is affected by inequality in an endogenous growth model. The benchmark is Persson and Tabellini (1994), who argue that productivity-induced income inequality leads to lower growth as distortionary taxes increase and harm capital accumulation. However, if income inequality stems from differences in capital, then labor tax rates fall, leading to higher growth. Based on OECD data, the chapter shows that an increase in capital income inequality has a significant positive relationship with subsequent economic growth.

Chapter 4, "Demography and the Composition of Taxes", analyzes the impact of population aging on the composition of taxes in an overlapping generations model. When the median voter is of working age, then population aging increases the demand for expenditure taxes rather than income taxes in order to increase the tax burden on the retired population. Consistent with the theory, international panel data exhibit a robust negative correlation between the extent of taxes on income relative to expenditure, and the fraction of the retired population.

Contents

List of Tables

List of Figures

This thesis is dedicated to all York graduate students.

Acknowledgements

I would like to express my sincere thanks to my two supervisors: Dr Andrew Pickering and Dr Paulo Santos Monteiro. You are much appreciated for the continuous support of my PhD study. All staff and students at the Department of Economics and Related Studies at the University of York contributed to a fantastic experience and great learning environment.

I owe many thanks to Prof Frank Cowell and Prof Gulcin Ozkan, members of my examination committee, for their valuable comments and stimulating discussions.

I am grateful to my friend Difu Liang, an airport designer and engineer in San Francisco, for his insightful comments on economics and encouragement.

I especially wish to thank Jay Chou, whose songs have consistently given me mental support to finish my PhD.

Finally, I especially thank my family.

Declaration

The contents of Chapter 2 have been presented at: the Public Choice Society Conference, Florida, March 2016; the European Public Choice Society Conference, University of Freiburg, April 2016; the Royal Economic Society PhD Meetings, Westminster Business School, January 2017; the Money Macro and Finance Research Group PhD Workshop, University of Portsmouth, April 2017; the Royal Economic Society Symposium of Junior Researchers, University of Bristol, April 2017; the Royal Economic Society Annual Conference, University of Bristol, April 2017; the Asian Meeting of the Econometric Society, Chinese University of Hong Kong, June 2017; the China Meeting of the Econometric Society, Wuhan University, June 2017; the North American Summer Meeting of the Econometric Society, Washington University in St. Louis, June 2017; and the Meeting of the Society for the Study of Economic Inequality, City University of New York, July 2017.

The contents of Chapter 3 have been presented at: the 49th Money Macro and Finance Research Group Annual Conference, King's College London, September 2017; the Royal Economic Society PhD Meetings, Westminster Business School, December 2017; the Royal Economic Society Annual Conference, University of Sussex, March 2018; the Louis-André Gérard-Varet Conference in Public Economics, Aix-en-Provence, June 2018. A paper version of Chapter 3 has won the first prize award for the best PhD paper at the 49th Money Macro and Finance Research Group Annual Conference.

The contents of Chapter 4 have been presented at: the Centre for Historical Economics and Related Research at York, University of York, November 2016; the European Public Choice

Society Conference, Central European University, April 2017; and the Scottish Economic Society Conference, Perth, April 2017.

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1

Introduction

Rising inequality in many developed countries has received renewed interest among policy-makers, academics, and the general public over the last decade, as shown by the attention generated by an academic book by Piketty (2014). Following up on Kuznets' (1953) pioneering study, a number of authors (Alvaredo et al., 2011-2018) have constructed long-run series of top income shares to measure income inequality. For example, estimates from the World Wealth and Income Database find that income concentration is high and growing in the United States: the richest 1 percent of households earned 20 percent of total income in 2015, up from 11 percent in 1978, while the bottom 50 percent experienced a complete collapse, from 20 percent to 12 percent of total income (Piketty et al., 2018). In contrast, and in spite of a similar trend, the top 1 percent share remains smaller than the bottom 50 percent in China but approaching the US level in 2015 (Piketty et al., 2017), and even less so in France as a typical representative of the West European pattern (Garbinti et al., 2017).

Despite the global resurgence of income inequality, there is still much doubt and debate as to whether such inequality is desirable. This debate is not a new phenomenon, often highly contested, and in recent years has fallen under the category of 'redistribution and growth' debate. However, the seminal paper by Meltzer and Richard (1981), building on earlier research by Romer (1975) and Roberts (1977), offers a sanguine prediction: greater before-tax income inequality implies divergence between mean and median income and so, under universal suffrage, leads to increased redistribution. Democracy thus provides a corrective to increased inequality. However, evidence supporting the Meltzer and Richard (1981) hypothesis is generally weak. Perotti (1996), Benabou (1996), Persson and Tabellini (2003), and Shelton (2007) all find an insignificant or even negative relationship between the size of government and the degree of inequality.

Much of the following literature has focused on the impact of income inequality on economic growth, typically based on Meltzer and Richard (1981). One important benchmark is an endogenous growth model by Persson and Tabellini (1994), who argue that if in a society the political decisions regarding redistribution generate economic policies that tax investment, then inequality should harm growth as it increases redistributive tax pressures. Empirical support for this hypothesis is generally weak. For example, Forbes (2000) finds that an increase in the level of income inequality in a country has a positive and significant relationship with subsequent growth rates in the short and medium term, by controlling for country-specific effects and period effects.

This thesis seeks to contribute to both the theoretical and empirical literature on the sources of income inequality. In the original mechanism (Meltzer and Richard, 1981; Persson and Tabellini, 1994), labor is the sole channel of income and the rich have higher income by means of higher productivity (individual-specific skills, in other words). Nevertheless, in reality, labor is not the sole channel of income for the rich, as widely observed (see Piketty et al., 2018), and moreover, the labor share of income has consistently declined in recent years (see Azmat et al., 2012; Karabarbounis and Neiman, 2013). Indeed, Piketty (2014) links increasing inequality to the declining labor share: if the rate of return on capital is greater than the rate of economic growth, then the capital share increases, and if ownership is highly concentrated within a small number of groups, then inequality inexorably rises. Furthermore, capital income has recently become more unequal as well as more important. Kaymak and Poschke (2016), and Saez and Zucman (2016) document considerable rises in the concentration of wealth in the US over the past 50 years. Hence the current inequality-redistribution and inequality-growth literature has one crucial omission: through focusing on the impacts of inequality and using the often assumed (aggregate) inequality solely induced by differences in labor productivity, the analysis is lacking a comprehensive consideration of capital income inequality.

Chapter 2 of this thesis attempts to address this omission through constructing a median voter model with the distinction that income inequality is engendered from differences in capital income as well as differences in labor productivity. The chapter, therefore, asks how inequality stemming from capital income affects government size. The key issue is that labor income is taxable, whilst income from capital is harder to tax. Evidence abounds of tax evasion or avoidance in the case of the latter (perhaps also due to capital mobility and international tax competition). It is harder to escape from PAYE. Like labor income, individuals differ in their endowment of capital, with a right skewed distribution of capital income. The majority of individuals, endowed with limited (or zero) assets or wealth, are compelled to supply labor for their income, which is taxed. On the other hand, those paid in capital income are to a meaningful extent able to avoid the same tax obligation, then the capital-rich are relatively less exposed to taxation. When income differences are induced by capital income, the capacity of the median voter to redistribute through the tax system is restricted as the capital-rich supply less (taxable) labor. If capital income inequality rises such that the capital-rich supply less labor, then the demand for tax on labor declines as the capital-poor (median voter) cannot expropriate the rich. The original Meltzer and Richard (1981) hypothesis therefore gets reversed: increased inequality in capital income leads to smaller government.

The relationship between the size of government and inequality is investigated empirically using a panel of OECD countries in Chapter 2, including a measure of capital income inequality as an additional explanatory variable. Due to limited availability of direct measures of capital income inequality, in the empirical analysis this is proxied by the top 1 percent total income share, taken from the World Wealth and Income Database. A theoretical justification for this proxy is Piketty (2014), in which capital is disproportionately owned by a small number of groups. In this analysis rising capital income with fixed capital ownership results in larger top income share. Certainly as in Piketty et al. (2017), capital income is an important component of the income of the top 1 percent. The empirical work also separately employs specific measures of productivity-induced labor income inequality as distinct from capital income inequality. These two measures are both empirically and conceptually distinct from one another. Consistent with the theory, the size of government is negatively associated with capital income inequality. The negative relationship holds up when the lagged dependent variable is controlled for, and also when capital income inequality is instrumented with measures of technological progress and capital market access. Moreover, controlling for capital income inequality yields a positive and significant relationship between government size and labor income inequality, consistent with Meltzer and Richard (1981) and in contrast to the voluminous empirical work testing their hypothesis.

Chapter 3, building on Chapter 2, analyzes how economic growth is affected by inequality in an endogenous growth model. The benchmark is Persson and Tabellini (1994), who argue that productivity-induced income inequality leads to lower growth as distortionary taxes increase and harm capital accumulation. When income differences are generated by capital income, the ability of the median voter to redistribute through taxation is constrained, while such redistributive policies are financed by distortionary taxes, in principle, affecting capital accumulation and growth-promoting activities. If capital income inequality increases (and it is the rich who enjoy capital income) such that labor tax rates fall, then the subsequent rate of economic growth increases because distortionary taxes fall and investment is facilitated. In direct contrast to Persson and Tabellini (1994), increased inequality in capital income leads to higher economic growth.

The relationship between inequality and growth is tested in a panel of OECD countries in Chapter 3, augmenting the work of Forbes (2000) to include capital income inequality as an additional explanatory variable. Consistent with the theory, growth is found to be positively associated with capital income inequality in the short and medium term. The estimated relationship is sizable: a one standard deviation increase in capital income inequality is statistically correlated with a 0.9 percent increase in average annual growth over the next five years. The positive relationship survives in a variety of econometric specifications, including when difference and system generalized method of moments techniques are used to address potential endogeneity problems. Moreover, once capital income inequality is controlled for, the impact of labor income inequality becomes negative, consistent with Persson and Tabellini (1994) and in contrast to the empirical work of Forbes (2000) for instance.

A separate strand of recent research instead focuses on the effect of population aging on fiscal policy, again related to Meltzer and Richard (1981). Razin et al. (2002) theorize that increases in the dependency ratio lead to lower labor tax rates and a fall in the generosity of social transfers in democracies. Government redistributes funds emanating from labor income taxes to both workers and retirees, and under democracy the equilibrium tax rate is that preferred by the median voter. An increasing dependency ratio implies a decline in the population growth rate, and lowers income taxes and transfers because the median voter is a worker who increasingly dislikes redistributing as the retired population increases. Empirical evidence generally has not supported this hypothesis that focuses only on income taxes. Disney (2007), Sanz and Velazquez (2007), and Shelton (2008) all find a positive relationship between population aging and the size of the welfare state.

Hence, this thesis also seeks to contribute to both the theoretical and empirical political economy literature on the composition of taxes. This idea is motivated by recent rises in taxes on goods and services and concurrent population aging, whilst income taxes are the only source of revenue in the original Razin et al. (2002) hypothesis. Revenue sources outside of income taxes are thus empirically becoming increasingly important components of total revenue (e.g. expenditure taxes account for approximately 30 percent as a share of the total tax revenue in the United Kingdom in recent decades). This is potentially a paradox as it might have been expected that countries with larger numbers of retirees would have higher income taxes and lower expenditure taxes, reflecting the increased political clout of the retired population who presumably would prefer income taxes.

Chapter 4 of this thesis analyzes the effect of population aging on the composition of taxes in a political-economy model. In an overlapping generations model, taxes are levied on both income and expenditure, which finance redistribution to both the working and retired population, with a balanced budget period by period. Income taxes are paid solely by workers, whist expenditure taxes are paid by both generations. As in Razin et al. (2002) the choice of the median voter is the unique Condorcet winner. When the median voter is of working age (as widely observed), then population aging increases the demand for expenditure rather than income taxes in order to increase the tax burden on the retired population. An unambiguous finding in this chapter is that the composition of taxes, defined as the extent to which taxes are levied on income relative to expenditure, theoretically always declines with the share of retirees.

The relationship between population aging and the extent of taxes on income relative to taxes on expenditure is investigated empirically using international panel data in Chapter 4. Following Pickering and Rajput (2018) the dependent variable is constructed by the ratio of taxes on income, profits and capital gains to taxes on goods and services, and the key demography measure is the percentage of the population over the age of 65, taken from the World Development Indicators database. Consistent with the theory, the extent of taxes on income relative to expenditure is found to be negatively associated with the fraction of the retired population. The estimated relationship is sizable: a one standard deviation rise in the fraction of the retired population is statistically associated with a fall of 0.63 in the ratio of income to expenditure taxes, holding all else equal. This relationship holds more strongly in stronger democracies.

In summary, Chapter 2 provides a theoretical and empirical investigation on how the size of government changes with capital income inequality. Chapter 3, building on Chapter 2, theoretically and empirically analyzes how inequality stemming from capital income affects growth. Chapter 4 investigates how population aging changes the demand for expenditure rather than income taxes, both theoretically and empirically. Chapter 5 concludes.

Chapter 2

Inequality and the Size of Government

2.1 Introduction

Neoclassical models of democracy, as articulated by Meltzer and Richard (1981),[1] offer a sanguine prediction: greater before-tax income inequality implies divergence between mean and median income and so, under universal suffrage, raises redistribution. Democracy, in principle, thus provides a corrective to increased inequality, and we should expect increased ex-ante inequality to lead to an increase in redistribution.

However, evidence supporting the Meltzer and Richard (1981) hypothesis is generally weak. For example, the United States and other Anglo-Saxon countries have greater income inequality but lower public sector spending as a share of total GDP, while Scandinavian countries have relatively equal income distributions and a larger government spending share. Perotti (1996), Benabou (1996), Bassett et al. (1999) and Persson and Tabellini (2003) all find an insignificant or even negative link between the size of government and the degree of inequality.[2]

In response to this puzzle, new theoretical work has proposed mechanisms through which greater inequality levels can coexist with smaller government under democracy. For instance,

[1] Romer (1975) and Roberts (1977) are important antecedents.

[2] More recent empirical literature (Mello and Tiongson, 2006; Shelton, 2007; and Muinelo-Gallo and Roca-Sagales, 2013) is also unsupportive.

Benabou (2000) identifies a functional role for the government to provide insurance (which implies redistribution) under capital market imperfections. The capacity for society to reach consensus on this role increases as the income distribution becomes more equal and risks become aligned and so government grows with equality. However, this type of mechanism also implies that government size should be positively correlated with economic growth and the evidence relating to the so-called 'Armey curve' surveyed by Bergh and Henrekson (2011) if anything points to a negative relationship, at least for high income countries.[3]

The approach taken in this chapter instead revisits Meltzer and Richard (1981) more closely. In the original mechanism, labor is the only source of income and the rich earn more by dint of higher productivity. However, labor is not the only source of income for the rich and, moreover, the labor share of income has declined in recent years (see Azmat et al., 2012; Karabarbounis and Neiman, 2013). Indeed, Piketty (2014) links rising inequality to the declining labor share: if the return to capital exceeds the rate of economic growth, then the capital share grows, and if ownership is concentrated within a small number of dynasties, then inequality inexorably increases. Furthermore, capital income has become more unequal as well as more important. Kaymak and Poschke (2016) document considerable increases in the concentration of wealth in the US over the past 50 years.

Hence we instead ask how inequality stemming from capital income affects government size. Individuals differ in their capital endowment, with a right skewed capital income distribution. The majority of individuals are endowed with limited (or zero) assets or wealth and so are compelled to supply labor for their income, which is taxed. In contrast, if capital-income is not taxed then the capital-rich are relatively less exposed to taxation. In direct contrast to Meltzer and Richard (1981), the key result is that increased inequality in capital income leads to smaller government. When income differences are driven by capital income, the capacity

[3]Other mechanisms are proposed by Persson (1995) and Rodriguez (2004). In the former, utility depends on relative consumption. In this model there is increasingly a problem of excessive labor supply in more equal societies and taxes work to increase utility by reducing labor. As in Benabou (2000), greater equality increases the capacity for agreement to tax, which again solves a market failure. Taxes work to eliminate the negative externalities associated with individual labor supply. Rodriguez (2004) instead models the political power of the rich as increasing with inequality, thereby reducing their obligation to pay tax. The democratic constraint is therefore undermined.

of the median voter to redistribute through the tax system is reduced because the capital-rich supply less (taxable) labor. If capital income inequality increases such that the capital-rich supply less labor, then the preferred labor income tax rate *falls* because the (capital-poor) median voter cannot effectuate redistribution. Our work is related to Krusell and Rios-Rull (1999), who study a version of Meltzer and Richard's model that includes inequality not only in labor income but also in wealth. However, we differ from Krusell and Rios-Rull (1999) as we assume capital income cannot be taxed, for the reasons explained below.

The relationship between the size of government and inequality is investigated empirically using a panel of fifteen OECD countries, including a measure of capital income inequality as an additional explanatory variable. Direct measures of capital income inequality are not widely available.[4] In the empirical work this is proxied by the top 1% total income share, taken from the World Wealth and Income Database (WID).[5] A theoretical justification for this approach is Piketty (2014), wherein capital is disproportionately owned by a small number of dynasties. In this analysis the larger top income share stems from increasing capital income with fixed capital ownership. Certainly capital income represents an important component of the income of the top 1%. Frydman and Saks (2010) document the increasing importance of stock options and long-term bonuses (also in the form of capital payments) in the remuneration of executives in large publicly traded corporations in the US.

Examination of disaggregated capital income data for a subset of countries provides empirical justification for this proxy. The WID contains non-wage (i.e. capital) income data for the top 1% and the top 10% for Australia, Canada, France and the United States. We posit that the higher the ratio of the share of non-wage income going to the top 1% relative to the top 10% the more unequal the capital income distribution. Ideally given our theory we would require that the numerator and denominator would respectively be the mean and 50th percentile non-wage income, but such data are not available. Nonetheless it seems

[4]Limited capital income data indicates that the rich do hide their income from capital, and in other words, it is difficult to tax their capital income. This in turn implies the lack of capital income tax (and data) that the median voter can effectuate, consistent with OECD evidence (extremely small size of capital income taxation), as well as the lack of capital income inequality data.

[5]The 0.1% income share could alternatively be used, though the results are very similiar because the correlation between the 0.1% and 1% income shares is around 0.98.

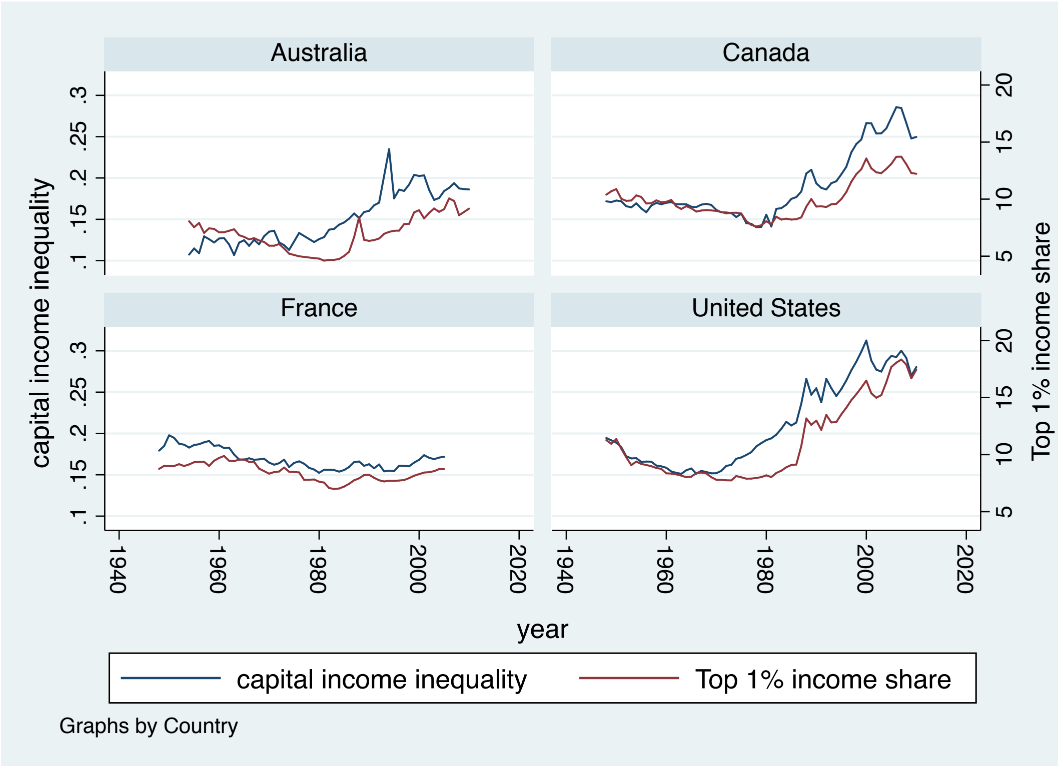


Fig. 2.1 Capital Income Inequality Versus Top 1% Income Share

plausible that inequality between the top 1% and the top 10% would be correlated with the theoretical ideal. Figure 2.1 plots this measure of capital income inequality together with the top 1% income share for these countries. In all four cases there is a strong correspondence between the direct measure of capital income inequality and the top income share, giving some credence to using the latter to proxy for the former for the wider sample of countries.

The empirical analysis below also separately employs specific measures of productivity-induced labor income inequality as distinct from capital income inequality. As we discuss below the two measures are empirically as well as conceptually distinct from one another. Consistent with our theory, the size of government is negatively associated with capital income inequality. A one standard deviation increase in capital income inequality leads to a reduction in the size of government of around 2.6% of GDP. The negative relationship holds up when the lagged dependent variable is controlled for, and also when capital income inequality is instrumented with measures of technological progress and capital market access.

We also find that once capital income inequality is controlled for, then the impact of labor income inequality becomes positive, consistent with Meltzer and Richard (1981) and in contrast to the voluminous empirical work testing their hypothesis.

The next section theoretically analyzes how the size of government changes with capital income inequality. Section 2.3 contains the empirical work, and section 2.4 concludes.

2.2 The Model

As in Meltzer and Richard (1981), individuals, indexed by i, have preferences defined over consumption c_i and leisure l_i, represented by a strictly concave, continuous and twice-differentiable utility function, $u_i(c_i, l_i)$. Consumption and leisure are both normal goods. Following the original, we first analyze the equilibrium behavior conditional on a given tax policy and then address the tax policy choice itself.[6]

2.2.1 Economic Environment

Income may be derived from both labor and capital. All individuals possess a unit of time to allocate to labor n_i, or leisure $l_i = 1 - n_i$. Individual labor income $y_i = x_i n_i$ depends on productivity, x_i, as well as hours worked, and is taxed at a linear rate t. Capital income varies exogenously across individuals and is denoted by R_i.[7] Following Meltzer and Richard (1981), consumption is also financed by lump-sum redistribution, r, common to all individuals, hence:

$$c_i = (1-t)x_i n_i + R_i + r. \tag{2.1}$$

[6] In order to compare with the Meltzer and Richard (1981) model, we start with a static model and focus on the tax policy choice generated by different sources of income inequality, rather than over-generation pension wealth decision.

[7] Capital income analyzed throughout this chapter is the income with zero opportunity cost, such as rental income. Housing price in large cities has consistently increased in recent years and has been accumulated as high levels of housing wealth. Higher levels of housing wealth do lead to larger inequality in wealth, while it cannot be taxed until it is sold (in the case of rental income, it can be claimed as smaller size or other items to avoid tax).

To clarify the argument, capital income is assumed to be untaxed.[8] In practice it is often more difficult to raise taxes on capital than on labor. Capital is often highly mobile internationally, whilst labor is not, and given this Diamond and Mirrlees (1971) show that small open economies should not tax capital income. Whilst in practice capital income taxation rates are positive, Gordon et al. (2004) observe they are lower than average labor income taxes in most countries. Moreover, capital income taxation liability can be reduced or even avoided altogether due to various loopholes including differential rates for different types of capital income (thereby enabling arbitrage opportunities) and the fact that interest payments are often tax-deductible. Indeed Gordon and Slemrod (1988), using US tax return data from 1983, estimated that the tax revenue loss from eliminating capital income taxation completely would be zero, hence that the tax burden on capital was effectively non-existent. Conceivably the perceived deadweight and/or capital flight losses from increasing capital income taxation nullify it as an instrument.[9] Thus we focus on the choice of the labor income tax.

Each individual chooses labor supply so as to maximize:

$$u_i(c_i, l_i) = u_i\Big[(1-t)\, x_i n_i + R_i + r, 1 - n_i\Big]. \tag{2.2}$$

The first-order condition is:

$$(1-t)\, x u_c - u_l = 0, \tag{2.3}$$

which determines the labor supply, $n[(1-t)x, R, r]$, for those who wish to work.[10] Since leisure is a normal good, we have that:

$$\frac{\partial n}{\partial R} = -\frac{(1-t)\, x u_{cc} - u_{cl}}{D} < 0, \tag{2.4}$$

[8]The results would all still stand if we instead modeled capital income taxation as fixed (and unresponsive to inequality), as observed from OECD data. The difficulty to collect capital income tax also underpins this argument. The rich are anti-tax: it could be easily observed that large companies always try to reclassify their labor-capital income in order to find tax haven. In rich economies with low self-employment, tax evasion is small on aggregate but high at the top, strong gradient within top 1% (Alstadsæter et al., 2017).

[9]Deadweight loss as well as capital flight loss leads to a loss of function of capital income taxation, which constraines the ability of median voter to influence over capital income tax rates.

[10]For simplicity (but without loss of generality) we henceforth assume that the joint distribution of x and R is such that $n_i > 0$ for all i, so that everyone supplies a strictly positive amount of market work.

with $D = [(1-t)x]^2 u_{cc} - 2(1-t)xu_{cl} + u_{ll} < 0$, given the assumption that u is strictly concave. Similarly, since consumption is a normal good we have that:

$$\frac{\partial c}{\partial R} = 1 + \frac{\partial n}{\partial R}(1-t)x = -\frac{u_{cl}x(1-t) - u_{ll}}{D} > 0, \tag{2.5}$$

a condition which imposes additional restrictions on u_{cl}. Hence, all else equal, people who are relatively capital-rich supply less labor and enjoy higher consumption.

There are two sources of heterogeneity that determine differences in before-tax labor income. Firstly productivity, as analyzed by Meltzer and Richard (1981), and secondly capital income endowments. At the individual level increases in productivity will all else equal increase labor income.[11] On the other hand increases in capital income will all else equal reduce the labor supply and, therefore, labor income. This underpins their proclivity towards taxation of labor income.

Average labor income can thus be written by integrating:

$$\bar{y} = \int_0^\infty \int_0^\infty xn\Big[R, r, (1-t)x\Big] f(x,R)\,\mathrm{d}x\mathrm{d}R. \tag{2.8}$$

where $f(x,R)$ is the joint density function of x and R. Individual productivity and capital endowments conceivably are correlated with each other to some extent: if, for example, high productivity individuals simultaneously enjoy high capital income. Finally, the government's

[11] Notice that, as in Meltzer and Richard (1981), the sign of

$$\frac{\partial n}{\partial x} = -\frac{(1-t)u_c + (1-t)^2 xnu_{cc} - (1-t)nu_{cl}}{D} \tag{2.6}$$

is indeterminate. Hence, the labor supply could be backward bending as productivity increases. Still, pre-tax labor income may never decline following an increase in productivity. To see this notice that, for any individual earning positive labor income, we have

$$\begin{aligned} \frac{\partial y}{\partial x} &= n + x\frac{\partial n}{\partial x} \\ &= -\frac{(1-t)xu_c + n[u_{cl}(1-t)x - u_{ll}]}{D} > 0, \end{aligned} \tag{2.7}$$

which must be positive given condition (2.5).

balanced budget requirement (in per capita terms) is given by:

$$t\bar{y} = r. \tag{2.9}$$

Note that analogous to (2.4), we have:

$$\frac{\partial n}{\partial r} = -\frac{(1-t)\,xu_{cc} - u_{cl}}{D} < 0. \tag{2.10}$$

Hence for given productivity and capital income endowment, individual labor supply falls with increased redistribution. Therefore:

$$\frac{\partial \bar{y}}{\partial r} = \int_0^\infty \int_0^\infty x \frac{\partial n}{\partial r} f(x,R)\,\mathrm{d}x\mathrm{d}R < 0. \tag{2.11}$$

This establishes that the left-hand side of (2.9) is strictly decreasing with r. Moreover, $t\bar{y}$ is non-negative and bounded above by $t\bar{x}$, where $\bar{x}$ is average productivity. In turn, the right-hand side of (2.9) is strictly increasing with r. Thus, there is a unique value of r to satisfy (2.9) for any t.

2.2.2 The Median Voter's Choice of Tax Policy

We now turn to the policy-setting decision. Crucially, the median voter is still a Condorcet winner even though the electorate is heterogeneous on two dimensions. The logic of this is that the preferred tax rate remains a monotonic function of the labor income alone, regardless of the underlying determinants of that labor income. Hence high labor income (whether induced by either high productivity or low capital income) will engender aversion to taxes, whilst low labor income (whether induced by low productivity or a generous capital income inheritance) will engender support for tax-financed redistribution. Formally, the median voter, m, is denoted as the owner of the median labor income. She sets taxes to maximize utility subject to the budget constraint (2.2), the government budget constraint (2.9), and a rational anticipation of how taxation will affect the incentives to supply labor in the economy.

The first-order condition for the median voter with respect to the tax rate is:

$$\bar{y} - y^m + t\left(\frac{d\bar{y}}{dt}\right) = 0, \tag{2.12}$$

where y^m is the labor income of the median voter. Condition (2.12) yields the following solution for the tax rate chosen by the median voter

$$t = \frac{m - 1 + \eta_r}{m - 1 + \eta_r + m\eta_\tau}, \tag{2.13}$$

with $\eta_r < 0$ and $\eta_\tau > 0$ the partial elasticities of average income (assumed constant, as in Meltzer and Richard, 1981), and $m = \bar{y}/y^m$.[12]

The key insight of Meltzer and Richard (1981) is that an increase in labor income inequality raises taxation, since an increase in income inequality raises m and from (2.13) we have that

$$\frac{dt}{dm} > 0. \tag{2.14}$$

Finally, although we impose almost no restrictions on the joint distribution $f(x,R)$, we wish to guarantee that: i) the chosen tax rate is positive; and that ii) the individuals that are in the top of the capital income distribution are never the decisive voter. Thus, in the sequel we make the following two (empirically supported) assumptions:

Assumption 2.1. *The joint distribution $f(x,R)$ is such that the labor income distribution is right-skewed. Thus, $y^m < \bar{y}$ and the chosen tax rate is positive.*

From (2.12) we see that Assumption 2.1 guarantees that the chosen tax rate is positive.

Assumption 2.2. *The joint distribution $f(x,R)$ is such that the set of individuals $i \in \mathscr{K}$ with capital income R_i (the top $\mathscr{K}$% of the capital income distribution) has productivity x_i which is sufficiently high so that $y_i = x_i n_i > y^m$ for all $i \in \mathscr{K}$.*

[12]Details are available in the Appendix A.1.

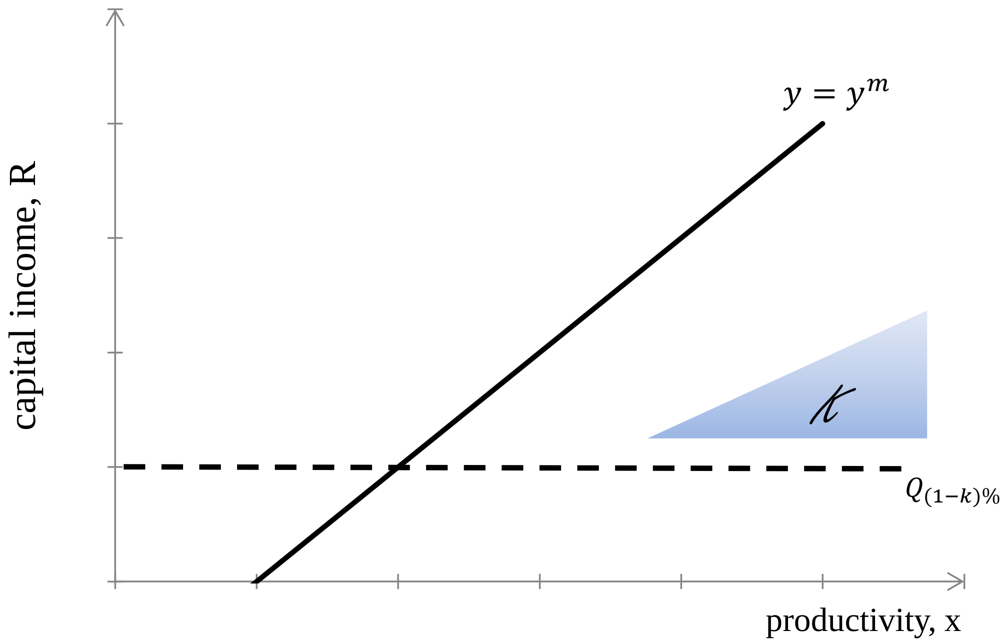


Fig. 2.2 Capital Income and Productivity Joint Distribution (Assumption 2.2)

Figure 2.2 illustrates the condition imposed by Assumption 2.2. The locus denoted $y = y^m$ represents productivity and capital income pairs, (x, R), for which labor income y is equal to the median voter's labor income, y^m. To the right of this locus, $y > y^m$, since $\frac{\partial y}{\partial x} > 0$ and $\frac{\partial y}{\partial R} < 0$. The dashed line denoted $\mathcal{Q}_{(1-\mathcal{K})\%}$ represents the $(1-\mathcal{K})\%$–quantile of the capital income marginal density function.[13] Assumption 2.2 is a condition requiring that the set $\mathcal{K}$ of all individuals with capital income above $\mathcal{Q}_{(1-\mathcal{K})\%}$ is located to the right of the locus $y = y^m$, as shown in Figure 2.2.[14]

[13]The size of the shaded area depends on the size of the set $\mathcal{K}$ (if we choose a larger set $\mathcal{K}$, then the shaded area will be larger). The position of this shaded area indicates initial levels of capital income that individuals in the set $\mathcal{K}$ have (and we will discuss below regarding to the consequence of increased capital income inequality). We focus on the 99% percentile because in the empirical section that follows we use the income share of the top 1% as our measure of capital income inequality.

[14]Consistent with Assumption 2.2, Atkinson and Lakner (2013) found that in the United States the tax units at the top of the labor income distribution are more likely to also be at the top of the capital income distribution.

2.2.3 Capital Income Inequality and Redistribution

We are interested in the consequences of higher capital income inequality. To study this issue we consider an increase in the capital income earned by the individuals in the set $\mathscr{K}$ of all individuals with capital income above $\mathscr{Q}_{(1-\mathscr{K})\%}$. This is represented in Figure 2.3: the individuals in the set $\mathscr{K}$ that correspond to the original individuals in the top $\mathscr{K}\%$ of the capital income distribution receive an exogenous increase in capital income; thus, the set $\mathscr{K}$ shifts upwards in the space (x,R), but still satisfying the restriction imposed by Assumption 2.2, that guarantees that the median voter does not belong to the members of the set $\mathscr{K}$ (the new set is represented by the triangle above, in Figure 2.3). Notice that this experiment constitutes an increase in capital income inequality, since we maintain the capital income of all the other individuals unchanged and, hence, the capital income share of the top $\mathscr{K}\%$ is increased.[15] Under a right-skewed labor income distribution $y^m < \bar{y}$, and given (2.13) above then $t > 0$. As with Meltzer and Richard (1981) demand for redistribution stems from changes in the labor income distribution. However, the labor income distribution may now change depending on the distribution of capital income as well as the productivity distribution.

To see the consequences of higher capital income inequality, notice that all the individuals in the set $\mathscr{K}$ will choose to work less, because they enjoy an increase in their capital income and leisure is a normal good. This will tend to lower the average labor income $\bar{y}$, since we have that

$$\bar{y} = p(\mathscr{K})\,\bar{y}(\mathscr{K}) + (1 - p(\mathscr{K}))\,\bar{y}(\sim\mathscr{K}), \tag{2.15}$$

where $\bar{y}(\mathscr{K})$ denotes the average labor income of the individuals in the set $\mathscr{K}$, $\bar{y}(\sim\mathscr{K})$ denotes the average labor income of the individuals not in the set $\mathscr{K}$, and $p(\mathscr{K})$ is the probability measure of the set of individuals $\mathscr{K}$. Notice that Assumption 2.2 guarantees that $\bar{y}(\mathscr{K}) > y^m$.

[15]It is not, however, a mean preserving spread in capital income. But lowering the capital income of the bottom $\mathscr{Q}_{(1-\mathscr{K})\%}$ capital income earners in order to preserve the mean capital income would only reinforce our results.

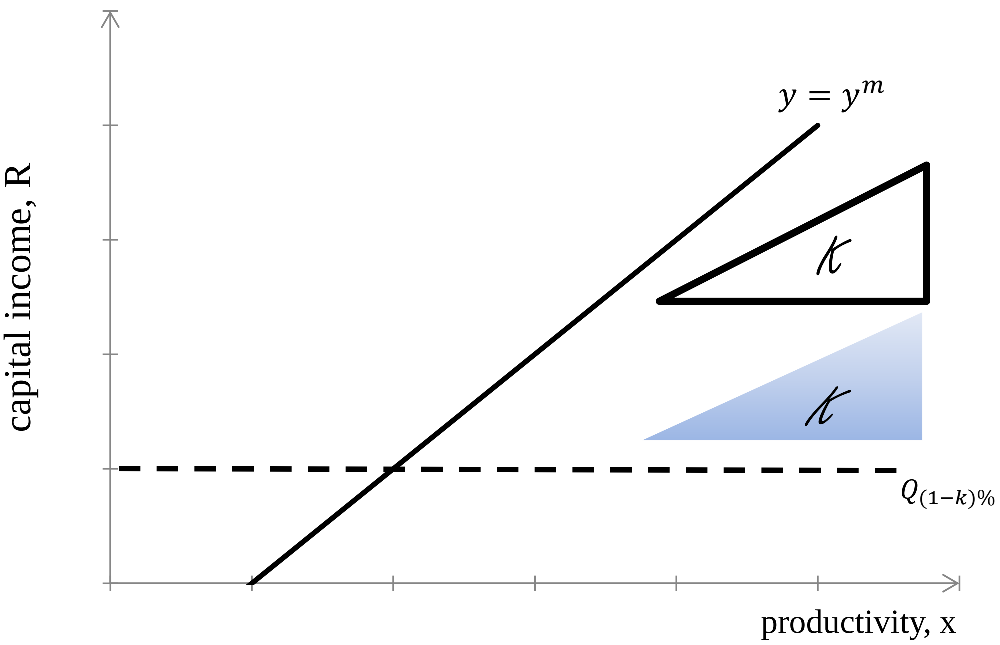


Fig. 2.3 Increase in Capital Income Inequality

On the other hand, the reduction in $\bar{y}$ implies that the individuals not in the set $\mathscr{K}$ will receive fewer transfers and, therefore, work more. From Assumption 2.2, the individual earning the median labor income is not in the set $\mathscr{K}$ and, thus, y^m will increase. The upshot is that $m = \bar{y}/y^m$ is decreased. Hence, the effect of the increase in capital income going to the top capital-income recipients is to reduce the gap between *taxable* mean and median labor income. Hence an increase in overall income inequality can coexist with a reduction in labor income inequality. Since $\frac{dt}{dm} > 0$, it follows that an increase in capital income inequality unambiguously lowers the tax rate chosen.

Proposition 2.1. *Suppose the top capital-income recipients are sufficiently productive that they also earn labor income above the median labor income (Assumption 2.2), and consider an increase in capital-income inequality represented by an increase in the capital income earned by the top capital-income recipients. Then the labor income tax rate t falls as capital income inequality rises.*

The proof of Proposition 2.1 is in Appendix A.2. In direct contrast to Meltzer and Richard (1981) government size diminishes with increased capital income inequality. If inequality increases such that the share of capital income going to the top income recipients increases, then the preferred tax rate falls because the (capital) rich are supplying less taxable labor income and hence the capacity of the median voter to redistribute is reduced.

The intuition here is that if capital income is highly concentrated within a small group of top (capital) rich, then these top rich individuals who are also with high productivity work less (and avoid being taxed their capital income through various ways). The aggregate tax revenue generated from income from labor (which is hard to escape) will be smaller, and therefore the aggregate level of redistribution is reduced.

The key issue is the extent to which the median voter can effectively redistribute through the tax system. As discussed above there are good reasons to believe that taxation of relatively mobile capital is considerably more difficult than taxation of labor income. If the rich are rich primarily due to capital income, perhaps because of the rising capital share, and perhaps due to successful reclassification of their income streams, then the capacity of the median voter to redistribute is curtailed. Moreover if rising inequality translates into further reductions in the supply of taxable labor then it follows that the demand for redistribution will fall.

2.3 Evidence

2.3.1 Data and Descriptive Statistics

The empirical analysis examines a panel of fifteen OECD countries over the period 1960-2007.[16] Following Pickering and Rockey (2011) and Facchini et al. (2017), the dependent variable is total government outlays as a percentage share of GDP, extracted from the OECD

[16]Specifically the countries included are Australia, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, Norway, Spain, Sweden, the United Kingdom, and the United States. Current data availability for the top income share precludes using other countries. The sample ends in 2007 due to the substantial toll on government outlays in many countries following the global financial crisis.

Economic Outlook database. Figure 2.4 depicts these data, showing all countries experienced an upward trend in the earlier years followed by a period of stasis or even slight decline since around 1990. Table 2.1 contains descriptive statistics of all the variables used in the analysis.

Figure 2.5 depicts the top income share data for all 15 countries. Note that the increases in the top income share to some extent coincides with the reversal of the growth of government noted above. Clearly there are interesting differences across the countries, for instance stronger recent increases in the English-speaking countries as discussed by Piketty and Saez (2006). The argument advanced in this chapter is the following: as the top income share increases, the supply of taxable labor of the rich falls, and hence support for taxation of labor income falls.

As noted above previous empirical literature has generally been unsupportive of the original Meltzer and Richard (1981) hypothesis. If the mechanism put forward in the present chapter is important, and capital-income inequality and productivity differences are correlated with each other, then arguably previous analyses have suffered from an omitted variable bias. A measure of productivity heterogeneity is thus also included in the empirical analysis. This measure is taken from the University of Texas Inequality Project's Estimated Household Income Inequality data.[17] These data (denoted by $UTIP$) use Theil's T statistic - measured across sectors within each country - to estimate wage inequality. Assuming competitive labor markets, then wage inequality should be capturing underlying heterogeneity in productivity.[18] Figure 2.6 depicts these data, which also exhibit increases in recent years, varying across countries. This measure is thus close to Meltzer and Richard (1981) original conception of the driver of the demand for redistribution - productivity-based inequality.

[17] See Galbraith and Kum (2005).

[18] To utilize Theil's T statistic - measured across sectors within each country - shows the evolution of economic inequality. We can do this with many different data sets, including at the regional or provincial level. With the $UTIP$ data, we can review changes in global inequality both across countries and through time. Nothing comparable can be done with previous data set (i.e. Deininger and Squire), for the measurements are too sparse and too inconsistent.

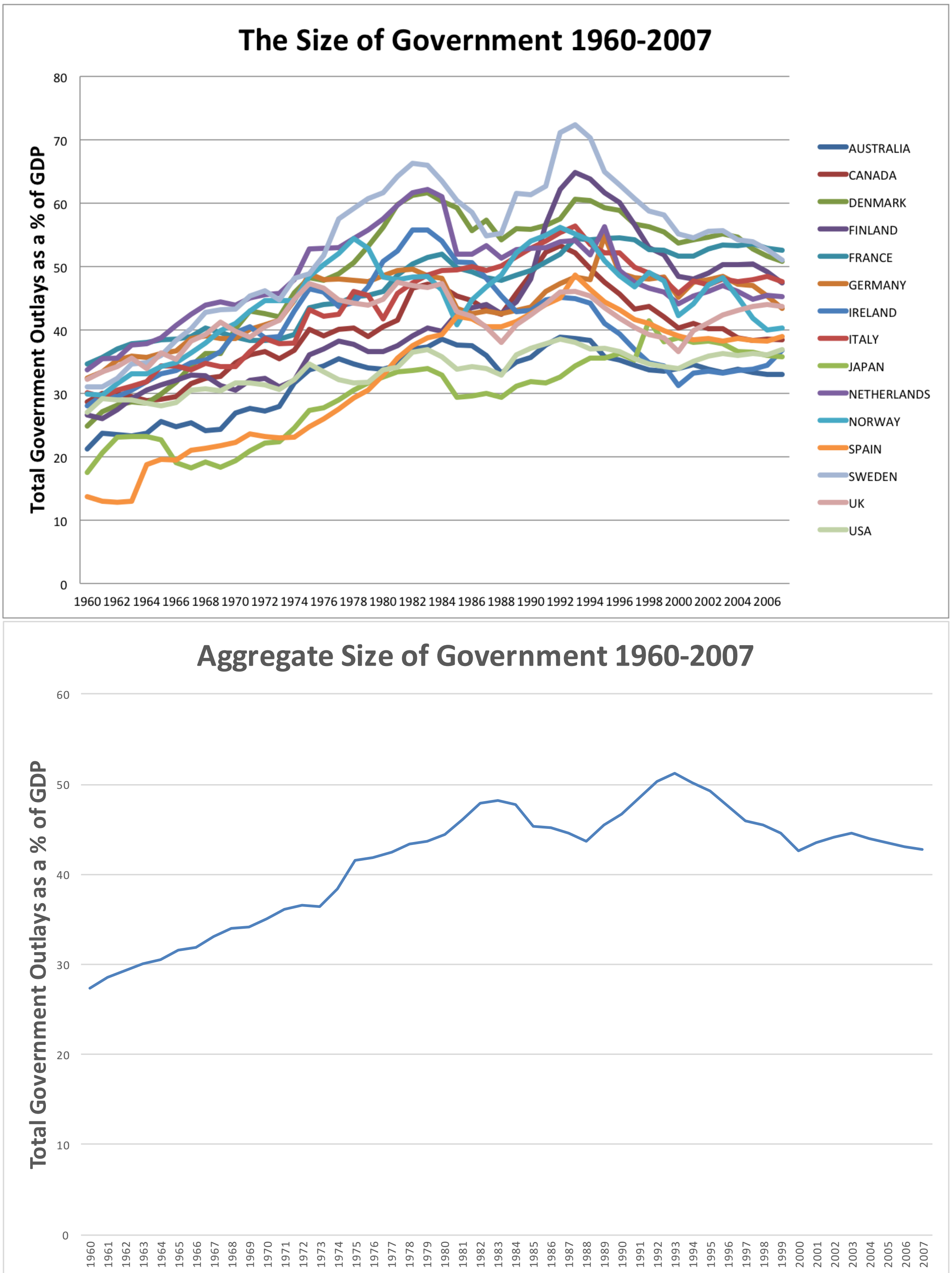
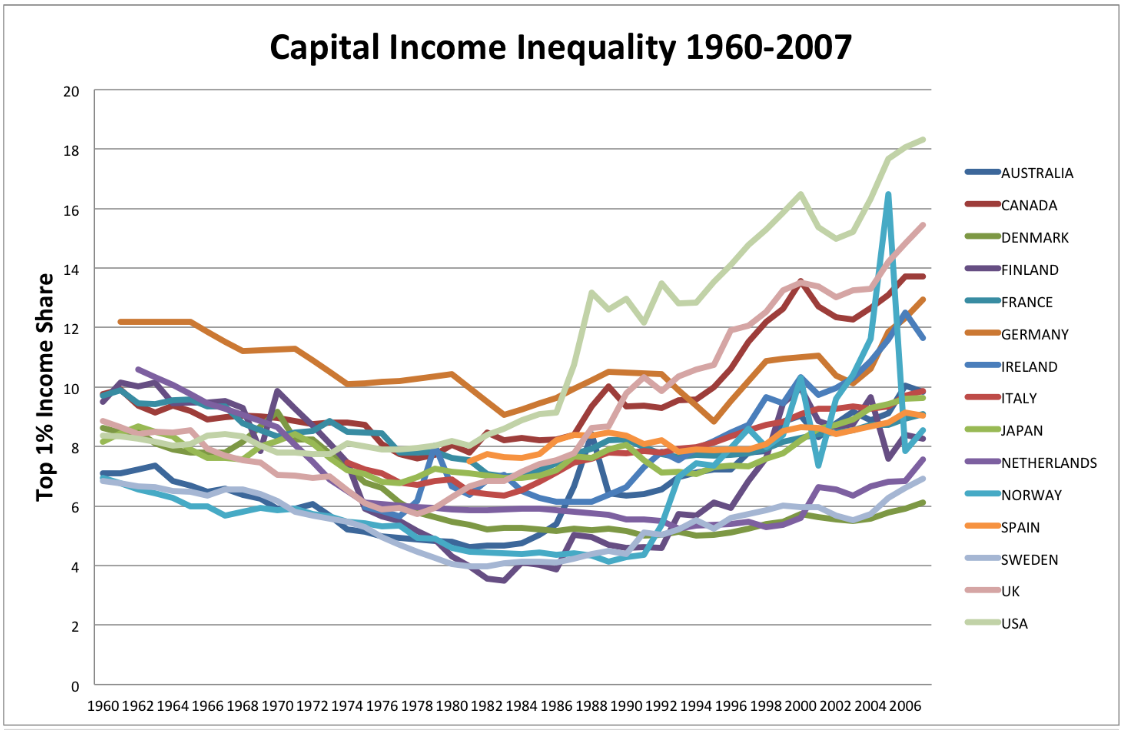


Fig. 2.4 The Size of Government, 1960-2007

Table 2.1 Descriptive Statistics

	obs	mean	std. dev.	min	max
OUTLAYS	720	41.50	10.14	12.8	72.4
TOPINC	617	7.89	2.44	3.49	18.33
UTIP	625	34.61	3.25	27.42	43.16
SHARE	548	67.97	6.03	44.74	82.10
FP	622	57.73	12.94	27.96	82.46
ln(y)	720	2.94	0.439	1.57	3.93
IDEO	683	0.043	0.117	-0.266	0.337
*PROP*1564	720	65.20	2.62	57.63	69.89
*PROP*65	720	12.79	2.84	5.73	21.02
TRADE	710	53.58	28.40	8.93	178.25
YGAP	720	0.026	1.34	-4.75	5.83
OIL_EX	720	4.60	12.05	0	72.36
OIL_IM	720	15.06	15.86	0	72.36
INTERNET	720	10.91	22.50	0	87.76
KAOPEN	564	1.41	1.22	-1.89	2.39

Notes: *OUTLAYS* denotes total government outlays as a percentage of GDP - taken from the OECD Economic Outlook database. *TOPINC* is the top 1% income share - taken from the WID. *UTIP* is the University of Texas Inequality Project's Estimated Household Income Inequality. *SHARE* is the business sector labor share - taken from the OECD database. *FP* is the female labor force as a percentage of the female population between 15 and 64 - also taken from the OECD database. y is real GDP per capita in \$000s of 2005 prices - taken from the Penn World Tables. *IDEO* is ideology used in Pickering and Rockey (2011). *PROP*1564 and *PROP*65 are respectively the proportion of the population aged between 15 and 64, and 65 and above - taken from WDI database. *TRADE* is the sum of exports and imports as a percentage of GDP. *YGAP* is the difference between the actual output and its trend value in percentage - also taken from WDI database. *OIL_EX* and *OIL_IM* are respectively the oil price times a dummy variable equal to 1 if net exports of oil are positive; and the oil price times a dummy variable equal to 1 if net exports of oil are negative - taken from US Energy Information Administration. *INTERNET* is the number of internet users per 100 people - also taken from WDI database. *KAOPEN* is the Chinn and Ito (2006) index for financial openness.



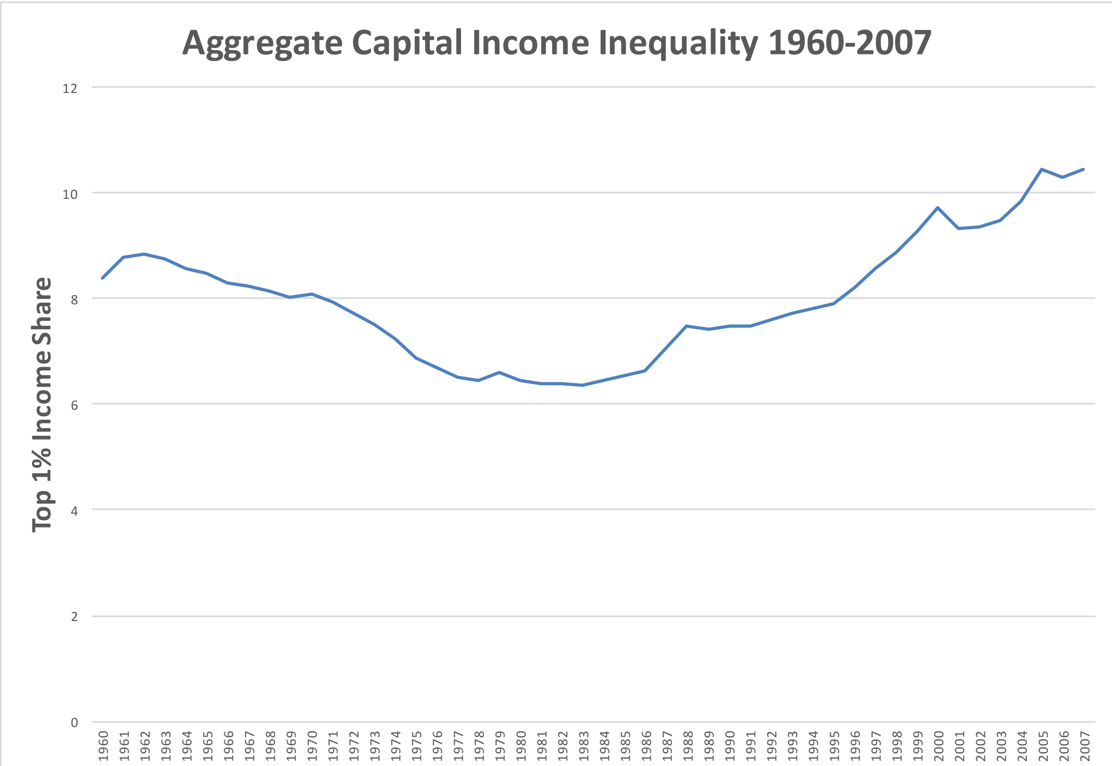


Fig. 2.5 Capital Income Inequality, 1960-2007

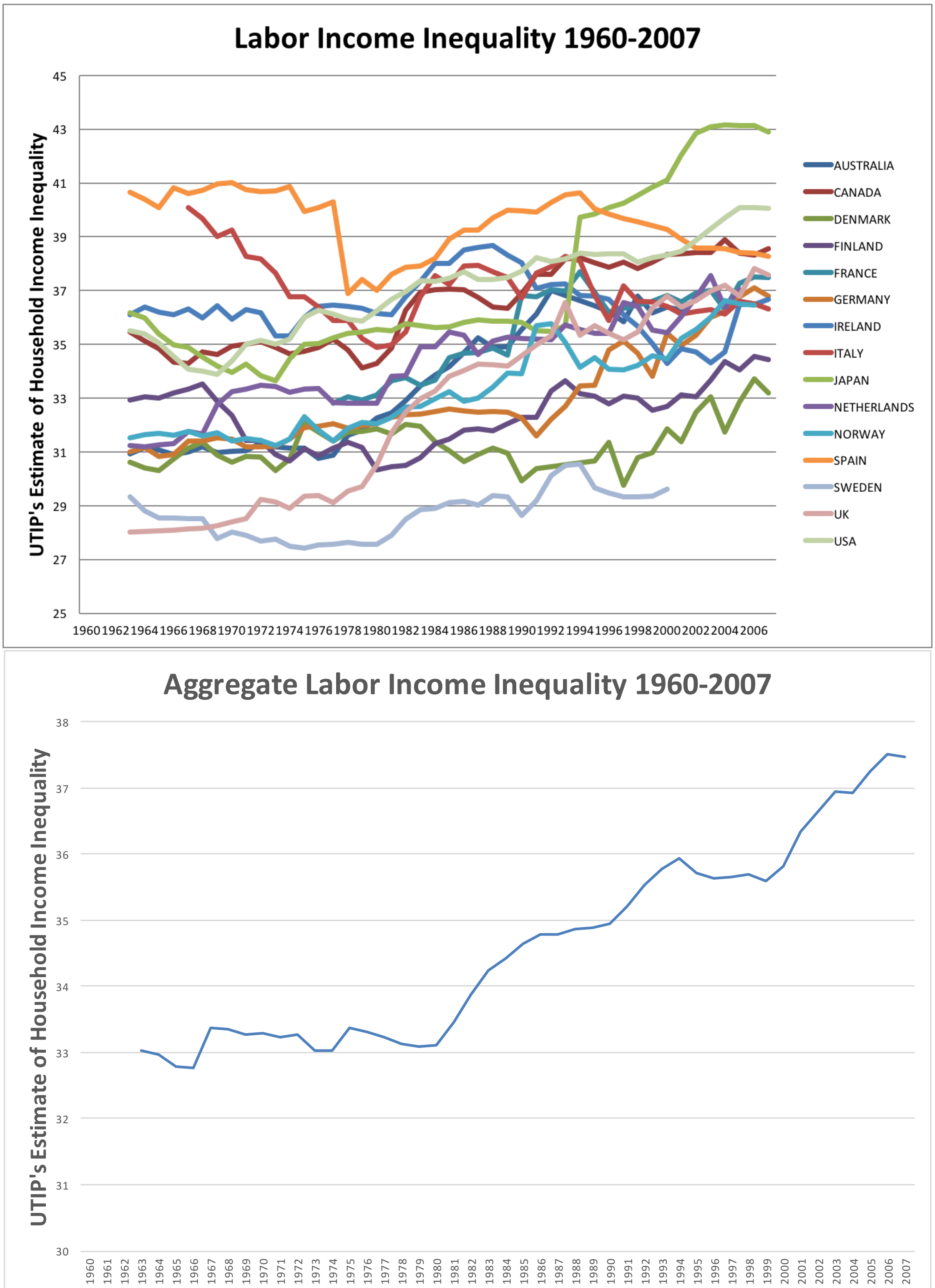


Fig. 2.6 Labor Income Inequality, 1960-2007

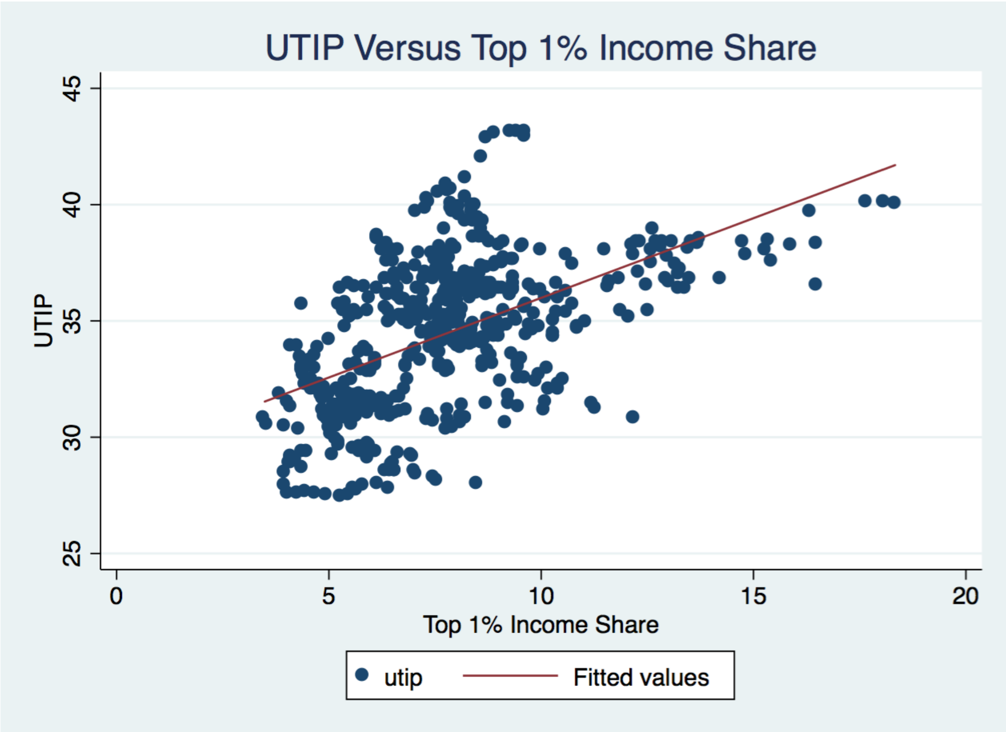


Fig. 2.7 Labor Income Inequality and Capital Income Inequality, 1960-2007

A natural objection here is that the top income share will also be picking up productivity-induced inequality. Inevitably there is a correlation between productivity inequality as measured by $UTIP$ and the income share of the top 1%, but this is somewhat weaker than might be expected. Figure 2.7 depicts a scatter plot of the two series, exhibiting a correlation coefficient of around 0.53. Hence there is meaningful separate information in the two series. Our argument is that the top income share is especially informative about capital income inequality rather than productivity-induced labor income inequality. The small sample of countries depicted in Figure 2.1 discussed in the introduction lends some credence to this argument.

The analysis includes control variables following Facchini et al. (2017). Controls include the natural logarithm of GDP per capita in constant chained PPP US$ ($\ln(y)$), taken from the Penn World Tables (e.g. see Ram, 1987). Ideology (denoted $IDEO$) and its interaction with income (denoted $INTERACT$) as used in Pickering and Rockey (2011), are also included as standard. Following Facchini et al. (2017) the labor share of income (denoted $SHARE$)

from the OECD database is also included to capture (falling) cost-push effects. Following Kau and Rubin (2002) and Winer et al. (2008) female participation (FP) in the labor force is also included. Further controls follow Persson and Tabellini (2003). Demographic effects are encapsulated in the percentage of the population between 15 and 64 years of age and the percentage over the age of 65 (denoted $PROP1564$ and $PROP65$), taken from the World Development Indicators (WDI) database. Following Rodrik (1998) the trade share (the sum of exports and imports as a percentage of GDP, denoted $TRADE$) is also employed in the regression analysis.[19]

Total government outlays in OECD countries vary counter-cyclically. There may also be cyclical movements in inequality. To address this potential problem the regression analysis employs the Persson and Tabellini (2003) cyclical control variables - the output gap (denoted $YGAP$) and oil price effects (depending on whether or not the country is a net oil-exporter or importer, denoted OIL_EX and OIL_IM) are also included in the analysis when annual data are used. To summarize, the frist approach to estimate the effect of inequality on total government outlays is to consider the following econometric model:

$$OUTLAYS_{i,t} = \beta_1 TOPINC_{i,t} + \beta_2 UTIP_{i,t} + \mathbf{x}'_{i,t}\Gamma + \alpha_i + \eta_t + u_{i,t} \tag{2.16}$$

where i represents each country and t represents each time period, all control variables analyzed above are included in the vector $\mathbf{x}_{i,t}$, α_i are country dummies, η_t are period dummies, and $u_{i,t}$ is the error term.

[19]The results would all still stand if we ideally incorporate polity variables as one further control variable. The reason why we omit here is to easily compare with the template work by Facchini et al. (2017), and that the sample countries we have are all with high democracy scores.

2.3.2 Panel Estimation

Table 2.2 contains estimation results from fixed-effects panel regressions with total outlays as a percentage of GDP as the dependent variable. Column 1a represents the current consensus, augmenting the benchmark specification in Facchini et al. (2017) with productivity-induced inequality ($UTIP$), and finding it to be highly insignificant. This insignificance coheres with the findings in Perotti (1996), Persson and Tabellini (2003), Mello and Tiongson (2006), and Shelton (2007). Column 1b further augments this specification with capital income inequality. The estimated coefficient for capital income inequality is negative, with a p-value of 1.7% and the estimated relationship is sizable: A one standard deviation increase in capital income inequality is statistically associated with government size which is smaller by 2.63% of GDP, consistent with the theoretical reasoning given here. It is also noteworthy that the coefficient estimate for productivity-induced labor income inequality increases substantially, though is still not statistically significant. Following Facchini et al. (2017) results are also presented (in columns 2a and 2b) using five-year averages of the data, and the results essentially duplicate those in column 1, establishing that the observed correlation is not caused by the cyclical features in the data.

Column 3 of Table 2.2 contains Arellano-Bond dynamic panel estimation results extending the specification used in column 2 to include the lagged dependent variable ($L.OUTLAYS$). Here the negative relationship between government size and capital income inequality holds up, and indeed the coefficient estimate pertaining to labor income inequality is now positive, consistent with the Meltzer and Richard (1981) hypothesis, and significantly different from zero at the 5% level. This evidence suggests that previous tests of the Meltzer and Richard (1981) hypothesis were hampered by the conflation of capital and labor income inequality.

Table 2.2 Panel Estimation Results with Fixed Effects

	(1a)	(1b)	(2a)	(2b)	(3)
L.OUTLAYS					0.423***
					(0.096)
TOPINC		-1.079**		-1.134***	-0.632*
		(0.401)		(0.367)	(0.361)
UTIP	0.139	0.730	0.132	0.497	0.932**
	(0.422)	(0.460)	(0.496)	(0.408)	(0.375)
SHARE	0.473***	0.364**	0.695***	0.522***	0.696***
	(0.155)	(0.127)	(0.200)	(0.163)	(0.164)
FP	-0.064	-0.019	-0.055	-0.096	0.141
	(0.188)	(0.192)	(0.221)	(0.257)	(0.112)
$\ln(y)$	-7.484*	-1.139	-6.217	2.342	-0.641
	(3.771)	(3.786)	(4.501)	(4.222)	(4.419)
IDEO	-53.942**	-38.339	-58.428**	-34.991	-7.101
	(22.948)	(23.633)	(23.033)	(24.511)	(19.769)
INTERACT	1.434	0.516	1.448*	0.388	0.145
	(0.877)	(0.918)	(0.805)	(0.863)	(0.723)
*PROP*1564	0.593	0.239	0.885	0.549	0.168
	(0.507)	(0.503)	(0.579)	(0.591)	(0.358)
*PROP*65	1.967***	1.102*	1.926***	1.318**	0.160
	(0.620)	(0.609)	(0.604)	(0.572)	(0.388)
TRADE	-0.026	-0.046	-0.031	-0.087	-0.057
	(0.047)	(0.046)	(0.054)	(0.053)	(0.065)
YGAP	-0.682***	-0.562***			
	(0.158)	(0.180)			
OIL_EX	0.031	0.003			
	(0.049)	(0.036)			
OIL_IM	0.052	0.045*			
	(0.030)	(0.024)			
Obs	506	462	113	113	98
No. Countries	15	15	15	15	15
Data	Annual	Annual	5-year averages	5-year averages	5-year averages
R^2 (within)	0.42	0.48	0.41	0.47	

Notes: Panel regressions of government outlays as a percentage share of GDP including fixed effects, *SHARE*, *FP*, $\ln(y)$, *IDEO*, *INTERACT*, *PROP*1564, *PROP*65, *TRADE*, *YGAP*, *OIL_EX*, *OIL_IM* as control variables. Column (3) contains Arellano-Bond estimation with lagged values of both the predetermined and endogenous variables as instruments. Robust standard errors are shown in parentheses. Standard errors are clustered by country. *, **, and *** respectively denote significance levels at 10%, 5% and 1%.

2.3.3 Instrumental Variables Estimation

The empirical analysis presented above establishes a robust negative statistical association between government size and capital income inequality in the presence of a substantial set of controls. However, these results do not establish causality, insofar that the movements in capital income inequality may be endogenous to the size of government, or alternatively both variables may co-move in response to an unobserved driver not accounted for in the controls. What is required for identification is a source of exogenous variation in capital income inequality. In this section we describe and deploy two potential instruments. An advantage of using two independent instruments is that it enables an overidentification test of the exclusion restriction that the instruments are not correlated with the error term in the second stage regression.

The first instrument is the number of internet users in percentage of the total population ($INTERNET$), encapsulating technological change.[20] Skill-biased technological change has been advanced as a (if not the) principle driver of rising inequality in general terms (for example in Goldin and Katz, 2009). Conceivably this process has especially underpinned increasing capital income inequality.[21] Atkinson et al. (2011) indeed document that a large part of the top income share derives from capital income.[22]

There are a number of channels through which advancing information technology could increase capital income inequality. One, as noted above is simply the mechanism advanced in Piketty (2014): if capital income rises with fixed ownership concentration, then capital inequality rises. Another stems from the observation that information technology is ‘weightless’ and in such circumstances the distinction between labor and capital income becomes somewhat arbitrary. Thus one can equally describe Mark Zuckerberg as being an

[20] Taken from the WDI database.

[21] Note that any effect of technological change through labor income inequality, or the labor share, is closed off due to these variables separately being included as controls in the analysis. It is still nonetheless possible that technology is correlated with the error term in the second-stage regression (i.e. violating the exclusion restriction), though the mechanism is not easy to see given the extensive set of controls. Moreover the exclusion restriction is tested below using the Hausman over-identification test.

[22] For instance in their figure 3 capital gains, capital income and business income represent well over half of the income of the top 0.1% in the US.

extremely productive worker, or as having created a company with enormous capital value. Relatedly, information technology plausibly has allowed many diverse activities to upscale their operations, resulting in significant increases in profitability which has in no small part been manifest in increased capital income for share owners or business partners. What is relevant for the theory above is *liability* for labor, as distinct from capital, income taxation. In particular in the case of new information technology, the new high earners face an interesting problem of how to classify their income.

Plausibly, and indeed empirically as observed above in related situations, they (or their accountants) will classify and organize their income so as to minimize taxation obligations. Given that it is almost universally the case that top marginal labor income taxes are higher than the (effective) top marginal capital income taxes, then income will likely be declared as capital income. To summarize, new technology has resulted in enormous rewards for a small number of people who have substantially registered these rewards in the form of capital income.

Our second instrument encapsulates exogenous variation in what we term as financial inclusiveness. By definition capital income requires capital ownership, and historically such ownership has not been widespread, even in the OECD. A necessary condition for mass ownership of capital assets and equity in particular is an established level of financial inclusion. A well developed financial system is one where it is easy, for all members of the population, to acquire (and sell) different types of capital assets. When financial inclusion is low, then conceivably at least some forms of asset ownership are not feasible for much of the population, and likely those with low income. Following this line of reasoning we conjecture that capital income inequality falls, conditionally, with financial inclusion.

The standard measure of financial inclusion is the ratio of stock market capitalization to GDP. However there are two problems with using this measure as an instrument in the context of our research objective. Firstly stock market capitalization is unlikely to be exogenous: a large public sector by construction implies a small private sector, hence lower stock market capitalization all else equal. Secondly, and more prosaically, the standard source for these data (the World Bank Global Financial Development Database) provides data only from 1989. To uncover exogenous variation in financial inclusion we use the Chinn-Ito index for financial openness (*KAOPEN*), an institutional measure that Chinn and Ito (2006) establish leads to changes in financial development, and therefore financial inclusion once legal systems and institutions are sufficiently developed (conditions which apply in the OECD). Notably these authors rule out reverse causality from financial inclusion to financial openness hence the Chinn-Ito index more plausibly satisfies the exogeneity requirement. To summarize the argument: The Chinn and Ito (2006) index exogenously drives financial inclusion. Exogenous increases in financial inclusion permit wider asset ownership thereby causing capital income inequality to fall. Hence we posit that capital income inequality exogenously falls with increases in the Chinn-Ito index.[23]

Table 2.3 contains the results of the IV estimation. Column 1 contains results using only the $INTERNET$ instrument, and column 2 contains results using only the *KAOPEN* instrument. The first-stage coefficient estimates for both instruments exhibit signs as hypothesized. Capital income inequality is estimated to (conditionally) increase with internet coverage, and the hypothesis that this particular instrument is weak can be rejected given that the F-statistic of the first stage regression exceeds 14. On the other hand capital income inequality is estimated to conditionally fall with capital market openness. The F-stat in this instance does not quite reach the threshold value of 10, but is not far off. Column 3 employs both instruments, with the advantage that this enables application of the overidentification test. The null hypothesis here is that the exclusion restriction is violated, and clearly the test statistic does not indicate rejection of this hypothesis. This test result thus supports the exclusion restriction that the instruments are not correlated with the second-stage error term.

Using the results from column 3, the coefficient estimate for $TOPINC$ in the second stage indicates that a one standard deviation increase in this variable *all else equal* causes a fall in

[23]Dabla-Norris et al. (2015) find that overall inequality actually increases with financial openness. The mechanism discussed therein is skills-bias – financial openness productively adds especially to the highly-skilled, thus increasing wage-inequality. It should be clear that this is a distinct hypothesis from ours, which emphasizes access to capital markets. Note again that labor income inequality is controlled for in both the first and second stages of the IV estimation. Hence the estimated effect of the Chinn-Ito index on capital income inequality is already conditional on any effect it has on labor income inequality.

the size of government of about 6% (i.e. using the data in Table 2.1 around 60% of a standard deviation). Importantly the assumption that all else is equal here is strong: we have already documented the positive correlation between $TOPINC$ and labor income inequality ($UTIP$), and indeed the coefficient estimate for the latter variable suggests an offsetting effect if both types of inequality simulataneously increase. What is clear from these results is that the effects of inequality in general terms are more complex than implied in the original Meltzer and Richard (1981) model. Labor income inequality now positively affects government size - consistent with Meltzer and Richard (1981). The top income share - which we interpret as a proxy especially for capital income inequality - negatively affects the size of government. This is consistent with the theoretical reasoning in this chapter. When it is difficult to tax capital income, then those who rely on labor income become averse to labor income taxation.

Columns 4 and 5 contain estimation results using 5-year averages of the data. For these regressions the lag of the top income share is used as an instrument, because $INTERNET$ and $KAOPEN$ are not sufficiently strong in this setting, where much of the time variation is averaged out. In column 4 $TOPINC$ is again estimated to have a significantly negative impact on government size, whilst labor income inequality ($UTIP$) remains positive and statistically significant. The negative impact of $TOPINC$ survives the addition of the lagged dependent variable in column 5, though the impact of productivity-induced labor income inequality is here reduced.

The concerns motivating IV estimation for capital income inequality ($TOPINC$) should also apply to labor income inequality ($UTIP$). To address this, in Table 2.4 we instrument for both $TOPINC$ and $UTIP$ by employing $INTERNET$, $KAOPEN$, the fifth lag of the top income share, and the fifth lage of labor income inequality. In this specification the estimated coefficient for $TOPINC$ is again negative. The hypothesis that these instruments are weak can be rejected given that the F-statistic of the first stage regression exceeds 23, and the result of the overidentification test supports the exclusion restriction that the instruments are not correlated with the second-stage error term.

Table 2.3 Instrumental Variable Estimation Results

	(1)	(2)	(3)	(4)	(5)
L.OUTLAYS					0.523^{***} (0.067)
TOPINC	-4.105^{***} (1.186)	-2.462^{**} (1.213)	-3.404^{***} (0.903)	-1.754^{***} (0.392)	-0.653^{**} (0.326)
UTIP	1.932^{***} (0.485)	1.420^{***} (0.460)	1.753^{***} (0.370)	0.836^{**} (0.350)	0.269 (0.280)
Obs	462	457	457	112	112
No. Countries	15	15	15	15	15
Method	IV	IV	IV	IV	IV
Data	Annual	Annual	Annual	5-year averages	5-year averages
Instruments	*INTERNET* 0.017^{***} (0.004)	*KAOPEN* -0.246^{***} (0.082)	*INTERNET* 0.015^{***} (0.004) *KAOPEN* -0.208^{**} (0.082)	*L.TOPINC* 0.801^{***} (0.056)	*L.TOPINC* 0.825^{***} (0.062)
F	14.78	9.038	10.64	205.3	177.9
p_{χ^2}			0.359		

Notes: IV is estimated by two-stage-least squares. First stage coefficients are reported below the named instruments in the Instruments row. F is an F-statistic for the statistical significance of the instruments in the first stage regression. p_{χ^2} is the p-value for the Chi-squared test of overidentifying restrictions. See also notes for Table 2.2 for other details.

Table 2.4 Instrumental Variable Estimation Results

	(1)	
TOPINC	-1.787***	
	(0.307)	
UTIP	0.628	
	(0.409)	
SHARE	0.292***	
	(0.077)	
Obs	406	
No. Countries	15	
Method	IV	
Data	Annual	
Instrument for	*TOPINC*	*UTIP*
L5.UTIP	0.348***	0.573***
	(0.040)	(0.043)
L5.TOPINC	0.586***	-0.035
	(0.042)	(0.045)
INTERNET	0.002	0.018***
	(0.004)	(0.004)
KAOPEN	-0.065	-0.083
	(0.067)	(0.073)
F	23.74	
p_{χ^2}	0.475	

Notes: IV is estimated by two-stage-least squares. First stage coefficients are reported below the named instruments in the Instruments row. F is an F-statistic for the statistical significance of the instruments in the first stage regression. p_{χ^2} is the p-value for the Chi-squared test of overidentifying restrictions. In this method we instrument for both $TOPINC$ and $UTIP$ using $L5.TOPINC$, $L5.UTIP$, $INTERNET$, and $KAOPEN$. See also notes for Table 2.2 for other details.

2.4 Conclusion

This chapter analyzes how inequality in the capital income distribution affects the size of government. Capital income is quite distinct from labor income. We define it as rental income, and also model it as untaxed, hence redistribution is financed solely by taxation applied to labor income, and voters have preferences over the tax rate based on their position in the capital income distribution. Despite the fact that there are two underlying sources of heterogeneity in the populations, the median voter is still the unique Condorcet winner because tax preferences are monotonic in labor income.

The result relating taxation levels to capital income inequality is novel. In contrast to Meltzer and Richard (1981) increased capital-income inequality now leads to smaller government. Agents who are endowed with capital income are less averse to labor-income taxation. The choice of labor income tax depends on the distribution of capital income: if the share of capital income of the rich increases, then their taxable labor supply falls and the preferred tax rate falls because the median voter has a reduced capacity to redistribute through taxation.

The relationship between the size of government and inequality is tested in a panel of OECD countries, augmenting the analysis of Pickering and Rockey (2011) and Facchini et al. (2017) to include capital income inequality as an additional explanatory variable. The measure of capital income inequality in the analysis is the top 1% income share. Consistent with the theory, government size is found to be negatively associated with capital income inequality. Moreover controlling for the top income share renders a consistently positive estimate for the impact of labor income inequality on government size, in line with the original Meltzer and Richard (1981) hypothesis. The negative impact of capital income inequality on government size survives a variety of econometric specifications, including when capital income inequality is instrumented with variables encapsulating technology and access to the capital market

Chapter 3

Inequality and Growth in the Twenty-First Century

3.1 Introduction

Is inequality necessarily harmful for growth? One important benchmark is an endogenous growth model by Persson and Tabellini (1994),[1] who embed the Meltzer and Richard (1981) argument and argue that if in a society the political decisions regarding redistribution generate economic policies that tax investment and constrain growth-promoting activities, then inequality should harm growth because it increases redistributive tax pressures.

A substantial amount of evidence has attempted to test the impact of inequality on growth, but the literature has not provided a satisfactory conclusion so far. For example, earlier cross-country OLS studies (e.g. see Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Perotti, 1996; and Deininger and Squire, 1998) all find negative consequence of higher inequality for economic performance. However, estimation using panel data, for example assembled by Deininger and Squire (1996), generally challenges the negative effect of inequality on growth found in cross-country regressions. Barro (2000) finds little overall link between income inequality and economic growth in a panel of countries, reporting a negative effect in

[1] Alesina and Rodrik (1994) and Bertola (1993) also provide similar anecdotes.

poor countries and a positive effect in rich countries. Perhaps the most surprising result is Forbes (2000). By controlling for country-specific effects and period effects, she finds that in the short and medium-term, an increase in the level of income inequality in a country has a positive and significant relationship with subsequent growth rates.[2]

In response to this puzzle, new theoretical literature has proposed mechanisms through which greater levels of income inequality can promote economic growth. For instance, Galor and Moav (2004) study the effect of inequality on growth along the process of development. In the early stages of development, when physical capital accumulation is the prime engine of growth, inequality stimulates growth as it channels resources towards individuals with more incentive to save. The positive effect of inequality on growth is reversed when human capital accumulation instead of physical capital is the primary engine for growth, where equality alleviates constraints on human capital accumulation and therefore stimulates growth.[3]

The mechanism analyzed in this chapter instead revisits Persson and Tabellini (1994) more closely. Their argument applies to income inequality stemming from differences in taxable labor income. However, traditional supplied labor is not the only source of income, especially for the rich, and moreover, the labor share of income has declined in recent years whilst capital income has recently become more unequal as well as more important (see Piketty, 2014; Saez and Zucman, 2016). Chapter two, building upon Meltzer and Richard (1981), links rising capital income inequality to declining redistribution: if inequality increases such that the share of capital income going to the top capital-income recipients increases, then the preferred tax rate falls because the (capital) rich are supplying less taxable labor income and hence the capacity of the median voter to redistribute is reduced.

[2] Li and Zou (1998) also find the positive link by using an improved data set on income inequality again compiled in Deininger and Squire (1996). More recent empirical work is that of Frank (2009), who, estimating a dynamic panel data model but using regional data from different U.S. states, provides evidence that the long-run relationship between inequality and growth in the United States is positive and in principle driven by the upper end of the income distribution.

[3] Moreover, Foellmi and Zweimuller (2006) study an innovation-based growth model and identify that an increasing unequal distribution of income affects the incentive to innovate through a price effect, where greater inequality allows innovators to charge higher prices, and a market-size effect, with an opposite direction. It turns out that the price effect always dominates the market-size effect, and thus increased inequality simulates growth.

Hence this chapter instead asks how inequality stemming from capital income affects economic growth in an endogenous growth model. Following chapter two, the key issue is that labor income is taxable, whilst income from capital is harder to tax. Evidence abounds of tax evasion or avoidance in the case of the latter.[4] It is harder to escape from PAYE. Like labor income, the capital income distribution is again right-skewed with the majority of individuals endowed with limited (or zero) assets or wealth, compelled to supply labor for their income, which is taxed. On the other hand, those paid in capital income are to a meaningful extent able to avoid the same tax obligation, then the capital-rich are relatively less exposed to taxation. Persson and Tabellini (1994) argue that productivity-induced income inequality leads to lower growth as distortionary taxes increase and harm capital accumulation. When income differences are generated by capital income, the ability of the median voter to redistribute through taxation is constrained, while such redistributive policies are financed by distortionary taxes, in principle, affecting capital accumulation and growth-promoting activities. If capital income inequality increases (and it is the rich who enjoy capital income) such that labor tax rates fall, then the subsequent rate of economic growth increases because distortionary taxes fall and investment is facilitated. In direct contrast to Persson and Tabellini (1994), increased inequality in capital income leads to higher growth.

The relationship between inequality and growth is investigated empirically using a panel of OECD countries, augmenting the analysis of Forbes (2000) to include a measure of capital income inequality as an additional explanatory variable. As direct measures of capital income inequality are not widely available, this is proxied by the top 1% total income share, taken from the World Wealth and Income Database (WID).[5] Though previous literature implies that countries undergoing faster income concentration should be suffering more from economic problems, I show that since the early 1990s, the period commonly viewed as the beginning of the adverse effects of inequality in much of the advanced world, there is no negative association between inequality and GDP per capita. Figure 3.1 depicts a plot of the change in GDP per capita and the top 1% total income share in OECD countries over 1960-2010.

[4]Some may argue that dividends are often taxed at source. Note that only cash dividends can be taxed while the rich can reclassify it through different ways, or simply transform it to stock dividends to avoid being taxed.

[5]See also arguments as well as Figure 2.1 discussed in chapter two.

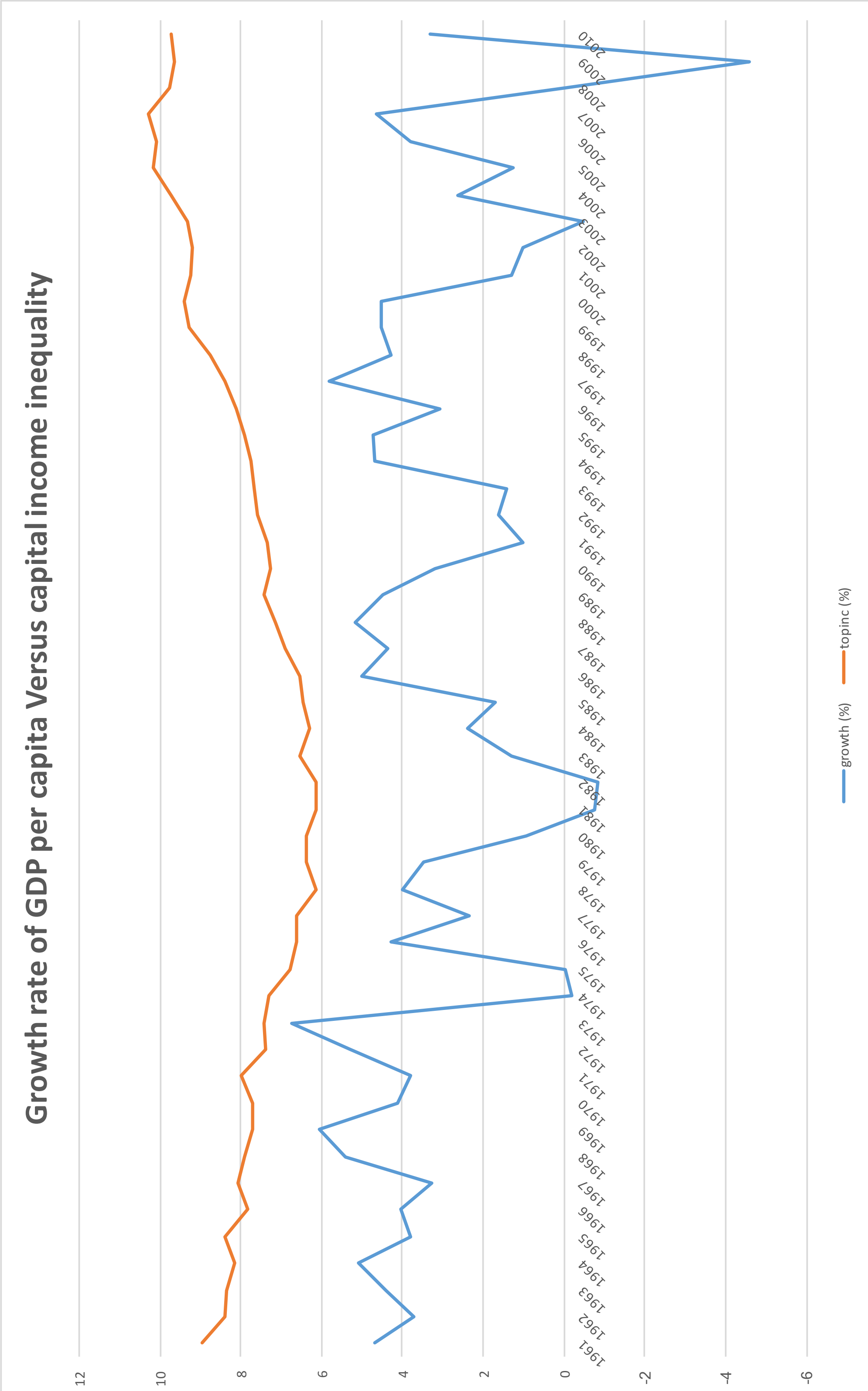


Fig. 3.1 Growth and Top 1% Income Share

The empirical analysis below also separately includes specific measures of productivity-induced labor income inequality as distinct from capital income inequality. As discussed in chapter two, the two measures are empirically as well as conceptually distinct from one another. Consistent with the theory proposed, an increase in capital income inequality has a positive and significant relationship with subsequent economic growth. A one standard deviation increase in capital income inequality is statistically correlated with a 0.9% increase in average annual growth over the next five years. The positive relationship holds up when different sample sets or omitted variables are considered, and also when difference and system generalized method of moments techniques are used to deal with potential endogeneity problems. I also find that once capital income inequality is controlled for, then the impact of labor income inequality becomes negative, consistent with Persson and Tabellini (1994) and in contrast to the empirical work using aggregative measures of inequality.

This chapter is part of a small literature that attempts to get a better grasp of the empirical picture with respect to the growth-inequality relationship. Earlier empirical contributions include Voitchovsky (2005), Castello-Climent (2010), and Halter et al. (2014). The first-mentioned paper questions previous empirical literature that uses aggregate indicators of inequality (e.g. Gini coefficient) which may mask different impacts of the upper and bottom part of the income distribution on growth. Castello-Climent (2010), consistent with Barro (2000), states that the results of inequality are different for rich and poor countries, finding a positive effect in the group of rich countries but a negative effect in the poor one.[6] Finally, Halter et al. (2014), by contrast, examine this relationship in the time dimension rather than the regional dimension (among rich and poor countries), and indeed find a positive effect in the short term but a negative effect further in the future. None of these papers, however, links inequality in the distribution of capital income to economic growth.

The next section theoretically analyzes how the rate of growth changes with capital income inequality. Section 3.3 contains the empirical work, and section 3.4 concludes.

[6] Banerjee and Duflo (2003) revisit both Perotti (1996) and Barro (2000) specifications, and argue that the growth rate is an inverted U-shaped function of changes in inequality. They believe that this non-linearity can account for different estimates of the relationship between inequality and growth within previous research.

3.2 The Model

This model revisits Persson and Tabellini (1994) to include labor income taxation instead of wealth taxation. As in chapter two individuals differ in terms of their capital income endowment as well as in terms of productivity. I study an overlapping generations model with constant population, where individuals live for two periods. Individuals born in period t, indexed by i, have preferences defined over consumption when young c^i, leisure when young l^i, and consumption when old d^i, represented by a strictly concave, continuous, twice-differentiable utility function $v_t^i = U(c_t^i, l_t^i, d_{t+1}^i)$. Consumption and leisure are both normal goods. Following the original, I first analyze the equilibrium behavior conditional on a given tax policy and then address the tax policy choice itself.

3.2.1 Economic Environment

Income may be derived from both labor and capital, and the stock of asset, k, accumulated on average by the previous generation has a positive externality on the income of the newborn generation as in Persson and Tabellini (1994). All individuals possess a unit of time to allocate to labor n^i, or leisure $l^i = 1 - n^i$. Individual labor income $y_t^i = n^i e^i k_t$ depends on productivity, e^i, as well as hours worked, and is taxed at a linear rate τ. Capital income varies exogenously across individuals and is denoted by $R^i k_t$, taxed at a linear rate ϑ. Following Meltzer and Richard (1981), consumption is also financed by lump-sum redistribution, r, common to all individuals, hence the budget constraints are:

$$c_t^i + k_{t+1}^i = (1 - \tau_t) n^i e^i k_t + r_t + (1 - \vartheta_t) R^i k_t \tag{3.1}$$

$$d_{t+1}^i = \gamma k_{t+1}^i \tag{3.2}$$

where k^i is the individual accumulation of asset, and γ is the exogenous rate of return on asset.[7] Individuals make decision between consumption and investment when young,

[7] Throughout the chapter I use superscripts to denote individual-specific variables and no superscripts to denote average variables.

financed by labor and capital income as well as lump-sum transfers, and benefit from the return on that investment when old. Note that the stock of aggregate capital is accumulated as average productivity of all individuals increases. With homothetic preferences, the ratio of consumption in the two periods is independent of wealth and labor income taxation, $\frac{d_{t+1}^i}{c_t^i} = D$. Equivalently, every individual has the same "saving rate".

To clarify the argument, capital income taxation, $\vartheta_t R^i k_t$, is assumed to be fixed in this model.[8] The key issue is that labor income is taxable, whilst income from capital is harder to tax (or at very low average rates). Evidence abounds of tax evasion or avoidance in the case of the latter (perhaps also due to capital mobility and international tax competition). Those paid in capital income are to a meaningful extent able to avoid the same tax obligation, then the capital-rich are relatively less exposed to taxation. Indeed, Alstadsæter et al. (2017) find that tax evasion is extremely concentrated in Scandinavian countries, and specifically, in 2006 the richest 0.01% of households, with net assets of at least \$40m, underpaid by over 30%. Evidence from OECD data is also in support of that capital income taxation is fixed (and unresponsive to inequality).[9] Thus this chapter focuses on the choice of the labor income tax.

Each individual chooses labor supply so as to maximize:

$$\begin{aligned} v_t^i = U\Big[&\frac{\gamma}{\gamma+D}\Big((1-\tau_t)n^i e^i k_t + r_t + (1-\vartheta_t)R^i k_t\Big), 1-n^i, \\ &\frac{\gamma D}{\gamma+D}\Big((1-\tau_t)n^i e^i k_t + r_t + (1-\vartheta_t)R^i k_t\Big)\Big]. \end{aligned} \tag{3.3}$$

The first-order condition is:

$$\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t U_c - U_l + \frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t U_d = 0 \tag{3.4}$$

[8] Note that in Persson and Tabellini (1994) labor income is assumed to be untaxed.

[9] For example, over recent decades, the ratios of capital income taxation to GDP in OECD countries have been relatively constant around 2-3%.

which determines the labor supply, $n[(1-\tau_t)e^i, r_t, R^i]$, for those who wish to work.[10] Since leisure is a normal good, I have that

$$\frac{\partial n^i}{\partial R^i} = -\frac{\frac{\partial^2 v_t^i}{\partial n^i \partial R^i}}{\frac{\partial}{\partial n^i}(\frac{\partial v_t^i}{\partial n^i})} < 0 \tag{3.5}$$

given the assumption that v is strictly concave.[11] Similarly, since consumption is a normal good I have that:

$$\begin{aligned}\frac{\partial c_t^i}{\partial R^i} &= \frac{\gamma k_t}{\gamma+D}\Big[1+\frac{\partial n^i}{\partial R^i}(1-\tau_t)e^i\Big], \\ &= \frac{\gamma k_t}{\gamma+D}\frac{\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t U_{cl} + \frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t U_{dl} - U_{ll}}{-\Delta} > 0,\end{aligned} \tag{3.6}$$

a condition which imposes additional restrictions on U_{cl} and U_{dl}. Hence, all else equal, people who are relatively capital-rich supply less labor and enjoy higher consumption.

There are two sources of heterogeneity that determine differences in before-tax labor income. Firstly productivity, as analyzed by Meltzer and Richard (1981), and secondly capital income endowments. At the individual level increases in productivity will all else equal increase

[10] Note again that k_t is given due to accumulation by the previous generation. Further, for simplicity (but without loss of generality) I henceforth assume that the joint distribution of e^i and R^i is such that $n^i > 0$ for all i, so that everyone supplies a strictly positive amount of market work.

[11] In detail, using (3.4), I have that

$$\begin{aligned}\frac{\partial n^i}{\partial R^i} &= \frac{\frac{\partial^2 v_t^i}{\partial n^i \partial R^i}}{-\frac{\partial}{\partial n^i}(\frac{\partial v_t^i}{\partial n^i})} \\ &= k_t \frac{(\frac{\gamma}{\gamma+D})^2(1-\tau_t)e^i k_t U_{cc} + (\frac{\gamma D}{\gamma+D})^2(1-\tau_t)e^i k_t U_{dd} - \frac{\gamma}{\gamma+D}U_{cl} + 2\frac{\gamma^2 D}{(\gamma+D)^2}(1-\tau_t)e^i k_t U_{cd} - \frac{\gamma D}{\gamma+D}U_{dl}}{-\Delta} < 0,\end{aligned}$$

with $\frac{\partial}{\partial n^i}(\frac{\partial v_t^i}{\partial n^i}) \equiv \Delta = \Big[\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t\Big]^2 U_{cc} + U_{ll} + \Big[\frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t\Big]^2 U_{dd} - 2\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t U_{cl} + 2\Big[\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t\Big]^2 D U_{cd} - 2\frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t U_{dl} < 0$.

labor income.[12] On the other hand increases in capital income will all else equal reduce the labor supply and, therefore, labor income. This underpins their proclivity towards taxation of labor income.

Average labor income can thus be written by integrating:

$$\bar{y}_t = k_t \int_0^{\infty} \int_0^{\infty} e^i n\Big[(1-\tau_t)e^i, r_t, R^i\Big] f(e^i, R^i) de^i dR^i \tag{3.8}$$

where $f(e^i, R^i)$ is joint distribution function of e^i and R^i. Individual productivity and capital endowments conceivably are correlated with each other to some extent: if, for example, high productivity individuals simultaneously enjoy high capital income. Finally, the government's balanced budget requirement (in per capita terms) is given by:

$$\vartheta_t R^i k_t + \tau_t \bar{y}_t = r_t. \tag{3.9}$$

For the average individual, by use of (3.2) and (3.8) I can thus solve for the growth rate of k

$$\begin{aligned} g_t &= \frac{k_{t+1} - k_t}{k_t} \\ &= \frac{D\Big[\int_0^{\infty}\int_0^{\infty} e^i n[(1-\tau_t)e^i, r_t, R^i] f(e^i, R^i) de^i dR^i + R\Big]}{\gamma + D} - 1 \end{aligned} \tag{3.10}$$

where R is average capital income. Note that analogous to (3.5), I have:

$$\frac{\partial n^i}{\partial r_t} = -\frac{\frac{\partial^2 v_t^i}{\partial n^i \partial r_t}}{\frac{\partial}{\partial n^i}\left(\frac{\partial v_t^i}{\partial n^i}\right)} < 0 \tag{3.11}$$

[12] Note that, as in Meltzer and Richard (1981), the sign of $\frac{\partial n^i}{\partial e^i}$ is indeterminate, but for any individual with positive labor income I have

$$\begin{aligned} \frac{\partial y_t^i}{\partial e^i} &= k_t\Big(n^i + e^i \frac{\partial n^i}{\partial e^i}\Big) \\ &= k_t \frac{e^i\Big[\frac{\gamma}{\gamma+D}(1-\tau_t)k_t U_c + \frac{\gamma D}{\gamma+D}(1-\tau_t)k_t U_d\Big] + n^i\Big[\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t U_{cl} + \frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t U_{dl} - U_{ll}\Big]}{-\Delta} > 0, \end{aligned} \tag{3.7}$$

must be positive given condition (3.6).

again given the assumption that v is strictly concave.[13] Hence for given productivity and capital income endowment, individual labor supply falls with increased redistribution. Therefore:

$$\frac{\partial \bar{y}_t}{\partial r_t} = k_t \int_0^\infty \int_0^\infty e^i \frac{\partial n^i}{\partial r_t} f(e^i, R^i) de^i dR^i < 0. \tag{3.12}$$

This establishes that the left-hand side of (3.9) is strictly decreasing with r.[14] Moreover, $\tau\bar{y}$ is non-negative and bounded above by τe, where e is average productivity. In turn, the right-hand side of (3.9) is strictly increasing with r. Thus, there is a unique value of r to satisfy (3.9) for any τ.

3.2.2 The Median Voter's Choice of Tax Policy

I now turn to the policy-setting decision. Crucially, the median voter is still a Condorcet winner even though the electorate is heterogeneous on two dimensions. The logic of this is that the preferred tax rate remains a monotonic function of the labor income alone, regardless of the underlying determinants of that labor income. Hence high labor income (whether induced by either high productivity or low capital income) will engender aversion to taxes, whilst low labor income (whether induced by low productivity or a generous capital income inheritance) will engender support for tax-financed redistribution. Formally, the median labor income-earner, m, is the median voter. She sets taxes to maximize utility subject to the budget constraints (3.1) and (3.2), the government budget constraint (3.9), and a rational anticipation of how taxation will affect the incentives to supply labor in the economy. The

[13] In detail, using (3.4), I have that

$$\frac{\partial n^i}{\partial r_t} = \frac{\frac{\partial^2 v_t^i}{\partial n^i \partial r_t}}{-\frac{\partial}{\partial n^i}\left(\frac{\partial v_t^i}{\partial n^i}\right)}$$
$$= \frac{(\frac{\gamma}{\gamma+D})^2(1-\tau_t)e^i k_t U_{cc} + (\frac{\gamma D}{\gamma+D})^2(1-\tau_t)e^i k_t U_{dd} - \frac{\gamma}{\gamma+D}U_{cl} + 2\frac{\gamma^2 D}{(\gamma+D)^2}(1-\tau_t)e^i k_t U_{cd} - \frac{\gamma D}{\gamma+D}U_{dl}}{-\Delta} < 0.$$

[14] As discussed above capital income taxation, $\vartheta_t R^i k_t$, is fixed.

first-order condition for the median voter with respect to the labor income tax rate is:

$$\bar{y}_t - y_t^m + \tau_t \frac{d\bar{y}_t}{d\tau_t} = 0 \tag{3.13}$$

where y^m is the labor income of the median voter. For a given ratio of mean to median labor income, the political equilibrium τ is constant over time, so that the time subscript t is suppressed henceforth. Let $\theta = 1 - \tau$ be the fraction of earned income retained. Condition (3.13) yields the following solution for the tax rate chosen by the median voter

$$\tau = \frac{m - 1 + \eta_r}{m - 1 + \eta_r + m\eta_\theta}, \tag{3.14}$$

with $\eta_r < 0$ and $\eta_\theta > 0$ the partial elasticities of average income (assumed constant, as in Meltzer and Richard, 1981), and labor income inequality $m = \bar{y}/y^m$.[15]

The key insight of Meltzer and Richard (1981) is that an increase in labor income inequality raises taxation, since an increase in income inequality raises m and from (3.14) I have that

$$\frac{d\tau}{dm} > 0. \tag{3.15}$$

I am interested in the consequences of higher capital income inequality as in chapter two. To study this issue I consider an increase in the capital income earned by the individuals in the set $\mathscr{K}$ of all individuals with capital income above $\mathscr{Q}_{99\%}$.[16] The effect of the increase in capital income going to the top capital-income recipients is to reduce the gap between *taxable* mean and median labor income. Hence an increase in overall income inequality can coexist with a reduction in labor income inequality. Since $\frac{d\tau}{dm} > 0$, it follows that an increase in capital income inequality unambiguously lowers the tax rate chosen.

Lemma 3.1. *Suppose the top capital-income recipients are sufficiently productive that they also earn labor income above the median labor income, and consider an increase in*

[15]Details are available in the Appendix B.1.

[16]I focus on the 99% percentile because in the empirical section that follows I use the income share of the top 1% as our measure of capital income inequality.

capital-income inequality represented by an increase in the capital income earned by the top capital-income recipients. Then the labor income tax rate τ falls as capital income inequality rises.

The proof of Lemma 3.1 is in Appendix B.2. This indicates that government size diminishes with increased capital income inequality, identical to Proposition 2.1 in chapter two. If inequality increases such that the share of capital income going to the top income recipients increases, then the preferred tax rate falls because the (capital) rich are supplying less taxable labor income and hence the capacity of the median voter to redistribute is reduced.

The key issue is the extent to which the median voter can effectively redistribute through the tax system. As discussed above there are good reasons to believe that taxation of relatively mobile capital is considerably more difficult than taxation of labor income. If the rich are rich primarily due to capital income, perhaps because of the rising capital share, and perhaps due to successful reclassification of their income streams, then the capacity of the median voter to redistribute is curtailed. Moreover if rising inequality translates into further reductions in the supply of taxable labor then it follows that the demand for redistribution will fall.

3.2.3 Capital Income Inequality and Growth

I now turn to the effect of capital income inequality on economic growth via the channel of redistribution. Combining (3.10) and the total derivative of $\bar{y}$, I have Lemma 3.2.

Lemma 3.2. *The growth rate falls as the labor income tax rate τ rises, e.g.,*

$$\frac{dg}{d\tau} = \frac{D}{\gamma + D} \frac{d\left[\int_0^\infty \int_0^\infty e^i n[(1-\tau)e^i, r, R^i] f(e^i, R^i) de^i dR^i + R\right]}{d\tau} < 0. \tag{3.16}$$

Thus all else equal, the higher is the labor income taxation, the lower is the growth rate. Appendix B.3 contains more mathematical details.

From the properties of the g and τ functions derived above, I can obtain Lemma 3.3.

Lemma 3.3. *A more unequal distribution of labor income decreases growth, e.g.,*

$$\frac{dg}{dm} = \frac{dg}{d\tau}\frac{d\tau}{dm} < 0. \tag{3.17}$$

This indicates that productivity-induced income inequality leads to lower growth as distortionary taxes increase and harm capital accumulation, which is identical in spirit to Persson and Tabellini (1994).[17] Now consider the consequences of higher capital income inequality and the mechanism analyzed above.

Proposition 3.1. *Greater inequality in capital income leads to reduced demand for redistribution and thus higher growth, if the top capital-income recipients are sufficiently productive (Lemma 3.1).*

In direct contrast to Persson and Tabellini (1994) economic growth increases with increased capital income inequality. When income differences are driven by capital income, the capacity of the median voter to redistribute through taxation is reduced since the capital-rich supply less (taxable) labor.[18] Such redistributive policies, financed by distortionary taxes, in principle, affect capital accumulation and growth-promoting activities which in turn is actually detrimental to growth. If capital income inequality increases (and it is the rich who enjoy capital income) such that the labor income tax rate chosen by the median voter falls as the (capital-poor) median voter cannot effectuate redistribution, then it generates smaller size of redistributive policies that are financed by less distortionary taxes. If declining distortionary taxes translate into further less restriction on aggregate capital accumulation (and less redistributive tax pressures), then the original Persson and Tabellini (1994) hypothesis gets reversed and it follows that subsequent economic growth will increase.

[17] Appendix B.4 analyzes how growth changes with labor income inequality without imposing capital income endowments.

[18] Analogous to Assumption 2.2, the condition that the top capital-income recipients are sufficiently productive indicates that the individuals that are in the top of the capital income distribution are never the decisive voter.

3.3 Evidence

3.3.1 Data and Econometric Specification

The empirical analysis examines a panel of OECD countries over the period 1965-2010.[19] Following Perotti (1996), the dependent variable is the average rate of growth of income per capita per five-year period as yearly growth rates incorporate short-run cyclical disturbances. For example, this means that growth rate in period 2 is averaged over 1971-1975 and is regressed on explanatory variables measured during period 1 (1966-1970).[20] This reduces yearly serial correlation from business cycles. The change from previous model is to include capital income inequality and labor income inequality instead of aggregate inequality. The final data set, with means, and standard deviations is contained in Table 3.1.

Figure 3.1 depicts the top income share data for all nineteen countries, showing all countries averagely experienced a downward trend in the earlier years followed by a period of stasis or even slight increase since around 1990. The argument proposed in this chapter is the following: as the top income share increases, distortionary taxes fall and investment is facilitated, which is likely to result in more accumulation and higher growth. Figure 3.2 depicts the raw correlation between the change in GDP per capita between 1960 and 2010 and the lagged top income share. Below I show that when I control for initial GDP per capita, human capital and market distortions, there is no evidence of a negative relationship between top income share and GDP per capita growth; on the contrary, the relationship is significantly positive in many specifications.

As noted above previous empirical literature including both country dummies and period dummies has generally been unsupportive of the original Persson and Tabellini (1994)

[19] Specifically the countries included are Australia, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. Current data availability for the top income share precludes using other countries.

[20] In practice, each explanatory variable is measured in 1970, except capital income inequality and labor income inequality, which are sometimes not available in a specific year and is taken from the year closest to 1970.

Table 3.1 Descriptive Statistics

	obs	mean	std. dev.	min	max
TOPINC	738	7.88	2.35	3.49	18.33
UTIP	746	35.18	3.54	27.42	44.12
y	969	20.71	9.37	1.07	53.77
*PROP*65	720	12.79	2.84	5.73	21.02
MEDU	190	3.02	1.38	0.32	6.59
FEDU	190	2.64	1.41	0.14	5.84
PPPI	864	85.88	25.20	34.58	179.06
*PROP*65	720	12.79	2.84	5.73	21.02
SHARE	548	67.97	6.03	44.74	82.10
OUTLAYS	861	39.98	10.65	12.8	72.4

Notes: *TOPINC* is the top 1% income share - taken from the WID. *UTIP* is the University of Texas Inequality Project's Estimated Household Income Inequality. Income *y* is real GDP per capita in $000s of 2005 prices - taken from the Penn World Tables. *MEDU* and *FEDU* are respectively the average years of secondary schooling in the male and female population aged over 25 - taken from Barro and Lee (1996). *PPPI* is the price level of investment measured as the PPP of investment over exchange rate relative to the United States - taken from the Penn World Tables. *PROP*65 is the proportion of the population aged 65 and above - taken from WDI database. *SHARE* is the business sector labor share - taken from the OECD database. *OUTLAYS* denotes total government outlays as a percentage of GDP - taken from the OECD Economic Outlook database.

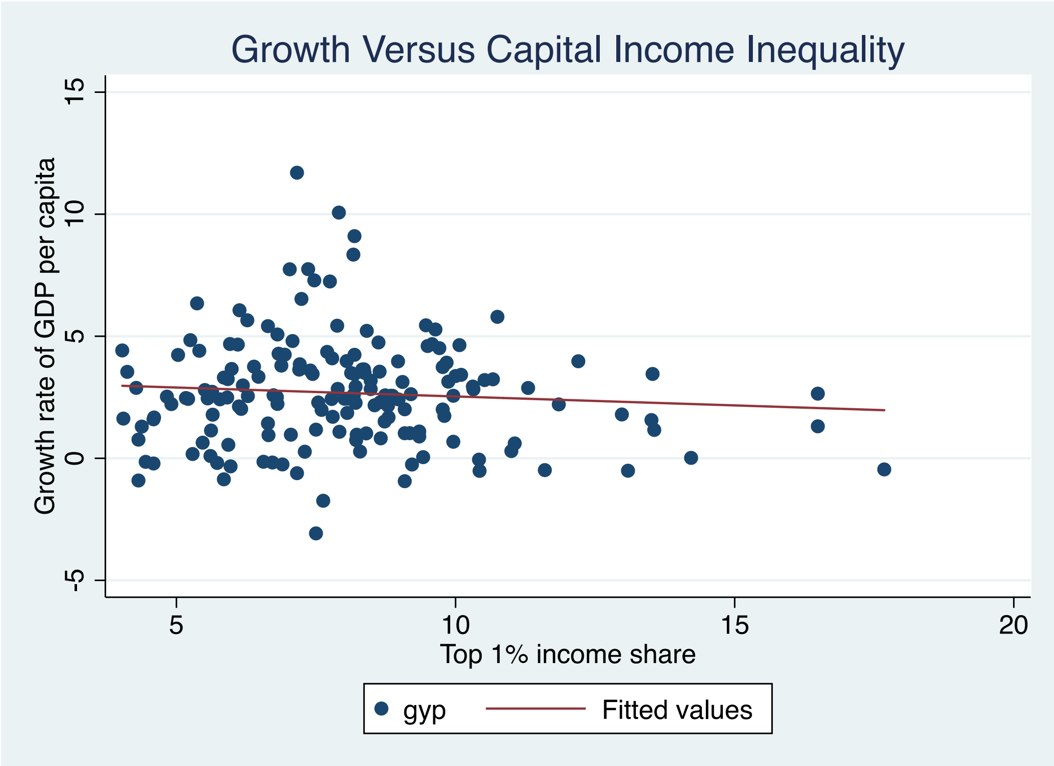


Fig. 3.2 Growth and Capital Income Inequality, 1960-2010

hypothesis. If the mechanism put forward in this chapter is important, and productivity differences and capital-income inequality are correlated with each other, then arguably previous analyses have suffered from an omitted variable bias. A measure of productivity heterogeneity, $UTIP$ (taken from the University of Texas Inequality Project's Estimated Household Income Inequality data), is therefore employed in the empirical analysis as in chapter two. As the top income share will also be picking up productivity-induced inequality, the correlation between productivity inequality as measured by $UTIP$ and the income share of the top 1% is somewhat weaker than might be expected, depicted by Figure 2.7 with smaller sample of OECD countries.[21] Hence there is meaningful separate information in the two series.[22] Figure 3.3 depicts the raw correlation between the change in GDP per capita between 1960 and 2010 and the lagged productivity-induced labor income inequality,

[21]Figure 6 in Luo (2017) depicts a scatter plot of the two series with full sample, exhibiting a correlation coefficient of around 0.49.

[22]The argument proposed in this chapter is that the top income share is especially informative about capital income inequality rather than productivity-induced labor income inequality.

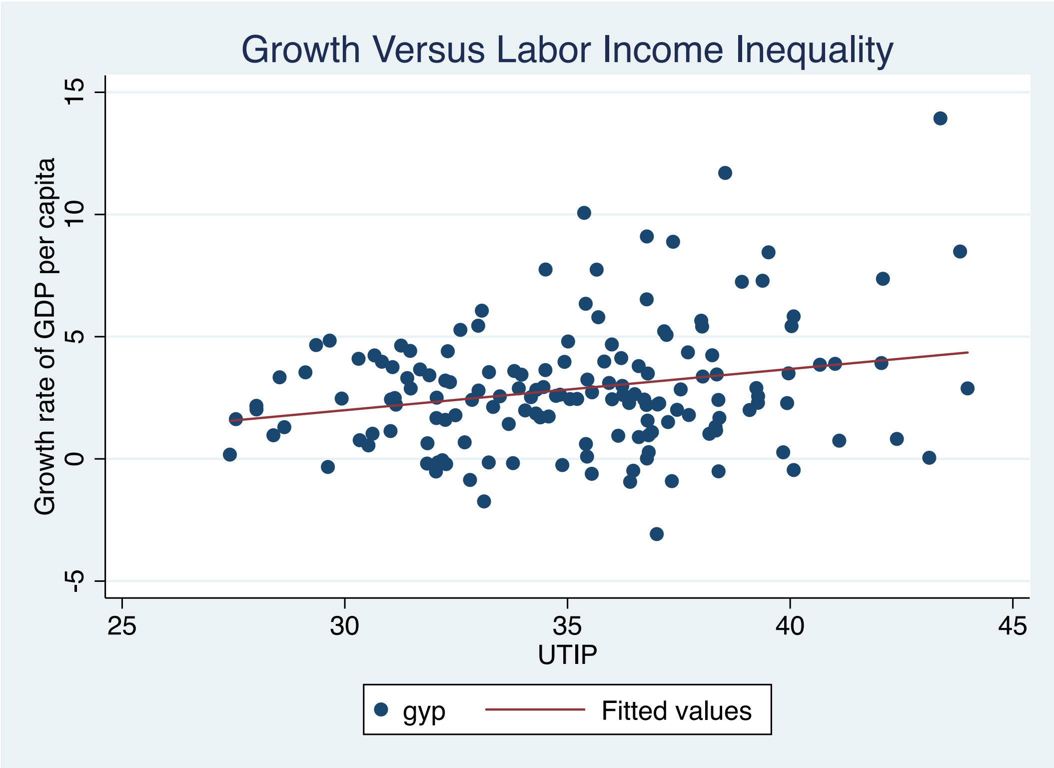


Fig. 3.3 Growth and Labor Income Inequality, 1960-2010

showing that there is no evidence of a negative relationship between labor income inequality and GDP per capita, in support of Forbes (2000).

The analysis includes control variables following Forbes (2000). Controls include per capita GDP in constant chained PPP US$ (denoted y). Per capita GDP y and the resultant growth rates are taken from the Penn World Tables. Following most empirical studies of income distribution and growth (e.g. Alesina and Rodrik, 1994; Persson and Tabellini, 1994) human capital effects are also included, and are represented by average years of secondary schooling in the male and female population aged over 25 (denoted $MEDU$ and $FEDU$), drawn from the data set compiled in Barro and Lee (1996). These two schooling variables proxy for the stock of human capital at the beginning of each of the estimation periods. The price level of investment (the PPP of investment over exchange rate relative to the United States, denoted $PPPI$) as used in Perotti (1996) is also employed in the regression analysis to capture market distortions that affect the cost of investment, also taken from the Penn World Tables. Finally, the country dummies are employed to control for time-invariant omitted-variable bias, and the period dummies are employed to control for global shocks that may affect aggregate growth in any periods but are not captured by other explanatory variables.

It is clearly possible to include a set of additional variables. However, as in Perotti (1996) this chapter mainly focuses on this simple specification for three considerations. First, in order to estimate the impact of inequality on growth it is important to make as few discrepancy as possible relative to typical growth model. Second, as the number of observations is limited by the availability of inequality data, this simplified specification will help maximize the number of degrees of freedom. Third, since some control variables used in standard-growth model (e.g. government expenditure) may be endogenous, focusing on stock variables measured at the start of each periods instead of flow variables measured throughout each periods can reduce the potential endogeneity problem. To summarize, the growth model central to be estimated is

$$\begin{aligned} GROWTH_{i,t} = {} & \beta_1 TOPINC_{i,t-1} + \beta_2 UTIP_{i,t-1} + \beta_3 y_{i,t-1} + \beta_4 MEDU_{i,t-1} \\ & + \beta_5 FEDU_{i,t-1} + \beta_6 PPPI_{i,t-1} + \alpha_i + \eta_t + u_{i,t} \end{aligned} \tag{3.18}$$

where i represents each country and t represents each time period, $GROWTH$ is average annual growth, α_i are country dummies, η_t are period dummies, and $u_{i,t}$ is the error term.

3.3.2 Panel Estimation

Table 3.2 contains estimation results from fixed-effects panel regressions with average annual growth rate as the dependent variable. Column 1 examines the original Persson and Tabellini (1994) hypothesis using five-year periods, 1965-2005, applying the benchmark specification in Forbes (2000) with productivity-induced inequality ($UTIP$), and finding its coefficient to be positive, though insignificant. This positive sign coheres with the results in Forbes (2000). Column 2 further augments this specification with capital income inequality. The estimated coefficient for capital income inequality is positive, with a p-value of 2.0% and the estimated relationship is sizable: A one standard deviation increase in capital income inequality is statistically correlated with a 0.9% increase in average annual growth over the next five years,[23] consistent with the theoretical reasoning given here. It is also noteworthy that the coefficient estimate for productivity-induced labor income inequality is now negative, though is still not statistically significant. Following Forbes (2000) results are also presented (in columns 4 and 5) using ten-year panels, and the results essentially duplicate those in columns 1 and 2, establishing that this observed short-term, positive relationship is not dampened over time.

Column 6 of Table 3.2 contains Arellano-Bond dynamic panel estimation results extending the specification used in columns 4 and 5 to include the lagged dependent variable ($GROWTH$). Here the positive relationship between capital income inequality and growth holds up, and indeed the coefficient estimate pertaining to labor income inequality is negative and significantly different from zero at the 10% level, consistent with the Persson and Tabellini (1994) hypothesis. This evidence suggests that previous tests of the Persson and Tabellini (1994) hypothesis were potentially hampered by the conflation of capital and labor income inequality. Columns 7-9 again test 1-3 using extended sample of 1965-2010 and duplicate their results.

Most of the coefficient estimates of control variables agree with those traditionally reported in typical literature. As indicated by models considering conditional convergence, the coefficient on initial income level is negative and statistically significant. Note also that the opposite signs on the coefficients of $MEDU$ and $FEDU$ are in line with the findings in Barro and Sala-I-Martin (2003) and Perotti (1996), who obtain the results based on a larger sample.[24] For a given male attainment, an increase in initial female attainment leads to less backwardness and thus slower subsequent growth since the economy converges toward steady state (see Barro and Sala-I-Martin, 2003).

[23] Note, however, that it is unlikely that any country's top income share could rise by this magnitude in a short period of time.

[24] The insignificance of the coefficients of $MEDU$ and $FEDU$ indicates that human capital accumulation in OECD countries is not the only one crucial driving force of economy, which in turn underpins the inequality-growth argument proposed by this chapter.

Table 3.2 Panel Estimation Results with Fixed Effects

	1965-2005						1965-2010		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.GROWTH			-0.0324			-0.0442			-0.00593
			(0.0911)			(0.203)			(0.0795)
L.TOPINC		0.412**	0.491***		0.529***	0.539***		0.424**	0.563***
		(0.161)	(0.164)		(0.140)	(0.187)		(0.179)	(0.124)
L.UTIP	0.119	-0.108	-0.189	0.198*	-0.0668	-0.265*	0.116	-0.114	-0.145
	(0.113)	(0.154)	(0.165)	(0.0974)	(0.111)	(0.143)	(0.120)	(0.154)	(0.142)
L.y	-0.459***	-0.612***	-0.828***	-0.557***	-0.610***	-0.792***	-0.336**	-0.544***	-0.825***
	(0.155)	(0.130)	(0.127)	(0.100)	(0.0964)	(0.138)	(0.123)	(0.113)	(0.0961)
L.MEDU	0.141	0.433	0.412	0.0947	0.915	1.038	-0.172	-0.137	-0.137
	(0.612)	(0.504)	(1.029)	(0.546)	(0.532)	(1.012)	(0.696)	(0.514)	(0.843)
L.FEDU	-0.369	-0.201	-0.439	0.0459	-0.462	-1.051	0.277	0.551	0.483
	(0.774)	(0.615)	(1.181)	(0.731)	(0.781)	(1.257)	(0.827)	(0.595)	(0.905)
L.PPPI	-0.00908	-0.00212	-0.00397	0.00215	0.0121	0.0263**	-0.00737	0.00171	0.00220
	(0.0127)	(0.00690)	(0.0124)	(0.0109)	(0.00954)	(0.0107)	(0.0124)	(0.00746)	(0.0110)
Obs	138	125	92	70	63	32	154	141	118
Countries	19	19	18	19	19	17	19	19	18
Periods	5-year	5-year	5-year	10-year	10-year	10-year	5-year	5-year	5-year
R^2 (within)	0.523	0.631		0.595	0.736		0.524	0.626	

Notes: Panel regressions of average annual per capita growth rate including fixed effects, *L.TOPINC*, *L.UTIP*, *L.y*, *L.MEDU*, *L.FEDU*, *L.PPPI*, and robust standard errors clustered by country in parentheses. Year dummies are included in all regressions. Columns (3) and (6) contain Arellano-Bond estimation with lagged values of both the predetermined and endogenous variables as instruments. Columns (7)-(9) again test (1)-(3) using extended sample 1965-2010. *, **, and *** respectively denote significance levels at 10%, 5%, and 1%.

3.3.3 Further Estimation

Previous work on the effect of income inequality on economic growth (Forbes, 2000) discusses the necessity to deal with potential endogeneity. Following the specification by Forbes (2000), column 1 of Table 3.3 applies difference GMM by Arellano and Bond (1991) to a panel covering 18 OECD countries during 1965-2010 in five-year periods. The basic difference GMM regression, eliminating the fixed effects and using lags of the endogenous variables as instruments, produces similar results presented in Table 3.2, in particular, significant and positive coefficient on lagged capital income inequality. While heightening the concern is the problem of weak instruments in difference GMM, which led to the development of system GMM by Arellano and Bover (1995) and Blundell and Bond (1998), and could reinforce endogeneity bias. The perfect p-value of 1.00 for the Hansen test is a classic sign of instrumental proliferation.[25]

The remaining columns 2-5 of Table 3.3 examine the sensitivity of the results to reducing the number of instruments. Column 2 firstly collapses the instruments. Columns 3 and 4 use two different lags from the instrument set, and column 5 combines the two modification. It should also be noted that the AR(2) test and the Hansen J test show that there is no further serial correlation, and the overidentifying restrictions are not rejected. As difference GMM can suffer from the problem of weak instruments, the rest columns of Table 3.3 utilise system GMM, which augments the equation estimated by difference GMM, simultaneously estimating an equation in levels with suitable lagged differences of endogenous variables as instruments. Therefore, columns 6-10 mimic columns 1-5 whilst instead using system GMM and produce similar results, which reinforce the proposed theory. Throughout Table 3.3 the positive coefficients on capital income inequality lose significance as the number of instruments falls.[26]

[25] More recent literature on weak instruments (in system GMM estimation in particular) has indicated that if instruments are weak, then inferences based on conventional Wald statistics can be misleading. However, this is still an open question, and we should not simply conclude that the system GMM estimator is not a useful tool for conducting cross-country growth empirics (see Bazzi and Clemens, 2013; Kraay, 2015).

[26] In the model, I theorize that greater capital income inequality leads to smaller tax burden on labor and thus higher subsequent economic growth rates. Appendix B.5 provides evidence showing a negative relationship between government size and growth.

Table 3.3 Difference and System GMM Regressions

	1965-2010									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
L.TOPINC	0.386**	0.441***	0.387**	0.480**	0.359	0.115*	0.211	0.0803	0.115*	0.459
	(0.160)	(0.167)	(0.188)	(0.203)	(0.341)	(0.0626)	(0.136)	(0.0622)	(0.0626)	(0.292)
L.UTIP	-0.147	-0.285	-0.123	-0.160	-0.0492	-0.0351	-0.204*	-0.0210	-0.0351	-0.469
	(0.169)	(0.224)	(0.240)	(0.169)	(0.437)	(0.0449)	(0.110)	(0.0454)	(0.0449)	(0.358)
L.y	-0.507***	-0.786***	-0.862***	-0.651***	-0.366	-0.170**	-0.211*	-0.134*	-0.170**	-0.330**
	(0.0958)	(0.0912)	(0.121)	(0.104)	(0.947)	(0.0788)	(0.110)	(0.0770)	(0.0788)	(0.164)
L.MEDU	-0.457	-0.726	-0.766	-0.608	-3.286	0.390	-0.437	-0.134	0.390	-2.115
	(0.460)	(1.020)	(0.972)	(0.533)	(4.913)	(0.492)	(0.745)	(0.499)	(0.492)	(1.322)
L.FEDU	0.832*	0.882	0.727	0.971	3.536	-0.162	0.900	0.195	-0.162	2.475**
	(0.490)	(1.295)	(1.121)	(0.662)	(5.658)	(0.452)	(0.827)	(0.387)	(0.452)	(1.258)
L.PPPI	0.000338	0.0166	0.00672	0.000378	-0.000266	-0.0177**	-0.0166	-0.0220***	-0.0177**	-0.0139
	(0.00628)	(0.0129)	(0.0117)	(0.00728)	(0.0242)	(0.00724)	(0.0134)	(0.00750)	(0.00724)	(0.0147)
Obs	118	118	118	118	118	141	141	141	141	141
Countries	18	18	18	18	18	19	19	19	19	19
Periods	5-year	5-year	5-year	5-year	5-year	5-year	5-year	5-year	5-year	5-year
Hansen test	4.44	9.57	7.16	5.55	exactly identified	3.31	4.52	11.93	5.54	2.34
AR(2) p-value	0.366	0.322	0.264	0.300	0.997	0.749	0.710	0.863	0.749	0.720
Estimator	difference GMM	difference GMM	difference GMM	difference GMM	difference GMM	system GMM	system GMM	system GMM	system GMM	system GMM
Method to reduce count		collapse	lags 1-1	lags 1-2	collapse & lags 1-1		collapse	lags 1-1	lags 1-2	collapse & lags 1-1

Notes: In columns (1)-(5) estimations use the difference GMM of Arellano and Bond (1991), with robust standard errors. In columns (6)-(10) estimations use the system GMM of Arellano and Bover (1995) and Blundell and Bond (1998), with robust standard errors. “collapse” stands for collapsed instruments; “lags” stands for restricting the number of lags used in generating instuments from the endogenous variables. Year dummies are included in all regressions. Endogenous variables used as instruments: *L.TOPINC*, *L.UTIP*, *L.y*, *L.MEDU*, *L.FEDU*, *L.PPPI*. *, **, and *** respectively denote significance levels at 10%, 5%, and 1%.

Table 3.4 tests the robustness and contains estimation results from fixed-effects panel regressions using five-year periods. Column 1 uses the same specification as column 2 of Table 3.2 but excluding Asian countries (i.e. Japan and Korea) to examine whether the regional coverage of the sample affects the results. Apart from the regional coverage, not surprisingly, the representative of very poor countries is extremely limited due to the unavailability of the top income share statistics. Alternatively, the relationship between capital income inequality and growth may depend on the stage of development of a country. I split the sample into wealthy and poor countries based on initial income level in 1965, and then reestimate equation (3.18) for two groups (reported in columns 2 and 3). Note that no matter which sample selection is utilized, the relationship between capital income inequality and growth remains positive and statistically significant.

Column 4 of Table 3.4 includes the percentage of population over the age of 65 (denoted $PROP65$) as an additional control variable following the argument of Perotti (1996). This demographic variable may be correlated with income inequality as among retirees both average income and inequality are lower. In turn, if the population in a country is older, then the demand for social security is potentially higher and hence, more taxation distortions and slower subsequent growth. The coefficient on this demographic variable is negative and statistically significant at the 5% level, supporting the mechanism proposed. Further, inequality stemming from capital income is likely to be correlated with the labor share of income (denoted $SHARE$). As in Facchini et al. (2017) a recent declining labor share has played a part in explaining the slowdown in the growth of government size and therefore, less distortions and higher growth. In fact, no matter whether I control for $PROP65$ or the labor share, as in columns 4 and 5, the coefficient on capital income inequality is positive and statistically significant at the 5% level. Note also that throughout columns 1-5 of Table 3.4 the coefficient estimates for labor income inequality are consistently negative (though not significant). If interpreted causally, the estimated effect of capital income inequality on growth remains sizable: An increase in $TOPINC$ by one standard deviation is associated with an increase in average rate of growth of GDP per capita by around 0.7%.

Table 3.4 Sensitivity Analysis

	1965-2005					1965-2010				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
L.TOPINC	0.304**	0.350**	0.466*	0.325**	0.440**	0.265**	0.295**	0.474**	0.328**	0.440**
	(0.114)	(0.146)	(0.209)	(0.132)	(0.171)	(0.110)	(0.117)	(0.189)	(0.132)	(0.175)
L.UTIP	-0.0493	-0.0460	-0.179	-0.0611	-0.126	0.000345	0.00976	-0.184	-0.0585	-0.110
	(0.121)	(0.129)	(0.195)	(0.108)	(0.134)	(0.113)	(0.143)	(0.200)	(0.0933)	(0.152)
L.y	-0.439***	-0.455***	-0.676*	-0.616***	-0.738***	-0.383***	-0.404***	-0.611**	-0.573***	-0.655***
	(0.119)	(0.0712)	(0.302)	(0.107)	(0.0999)	(0.0634)	(0.0986)	(0.221)	(0.0763)	(0.101)
L.MEDU	-0.131	0.786	0.451	-0.162	0.316	-0.806	0.0346	0.130	-0.716	-0.259
	(0.460)	(0.964)	(1.358)	(0.706)	(0.617)	(0.480)	(0.915)	(0.704)	(0.604)	(0.662)
L.FEDU	0.339	0.0446	-0.855	0.327	0.0651	1.186**	0.806	-0.0412	1.202*	0.783
	(0.509)	(1.018)	(2.694)	(0.711)	(0.661)	(0.505)	(0.785)	(1.126)	(0.606)	(0.701)
L.PPPI	-0.000104	0.00300	0.00235	0.00495	0.00433	0.00234	0.00891	0.00305	0.00657	0.00629
	(0.00806)	(0.00851)	(0.0279)	(0.00863)	(0.00769)	(0.00736)	(0.00883)	(0.0234)	(0.00754)	(0.00907)
L.PROP65				-0.463**					-0.461***	
				(0.216)					(0.130)	
L.SHARE					-0.0550					-0.0658
					(0.0561)					(0.0579)
Obs	113	63	62	125	108	127	69	72	141	124
Countries	17	9	10	19	18	17	9	10	19	18
Data	Excluding Asia	Higher income	Lower income	Full	Full	Excluding Asia	Higher income	Lower income	Full	Full
Periods	5-year	5-year	5-year	5-year	5-year	5-year	5-year	5-year	5-year	5-year
R^2 (within)	0.634	0.608	0.769	0.659	0.676	0.637	0.582	0.773	0.670	0.667

Notes: Regression specification is the same as column (2) of Table 3.2, and robust standard errors clustered by country in parentheses. Year dummies are included in all regressions. Column (1) excludes Asian countries. Columns (2) and (3) respectively correspond to higher and lower levels of initial income in 1965. Column (4) includes *L.PROP65* as a further control, and column (5) includes *L.SHARE* as a further control. Columns (6)-(10) again test (1)-(5) using extended sample 1965-2010.

The contribution of this empirical work is that it does not follow previous empirical studies using aggregate income inequality measure (i.e. Gini coefficient), but instead splitting the aggregate income inequality measure into capital income inequality and labor income inequality. If only labor income inequality is incorporated, then evidence shows its coefficient to be positive, in direct contrast with earlier cross-country OLS studies (e.g. see Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Perotti, 1996; and Deininger and Squire, 1998). When capital income inequality is controlled (as the main innovation of this chapter), the relationship between labor income inequality and growth is found to be negative, consistent with those paper conjectured.[27]

3.4 Conclusion

This chapter analyzes how inequality in the capital income distribution affects growth. Capital income is quite distinct from labor income. I define it as rental income, and also model it as difficult to tax, hence redistribution is financed mainly by taxation applied to labor income, and voters have preferences over the labor income tax rate based on their position in the capital income distribution. Despite the fact that there are two underlying sources of heterogeneity in the populations, the median voter is still the unique Condorcet winner because tax preferences are monotonic in labor income.

The result relating growth to capital income inequality is novel. In contrast to Persson and Tabellini (1994) increased capital-income inequality now leads to higher growth. Agents who are endowed with capital income are less averse to labor-income taxation. If the share of capital income of the rich increases such that their taxable labor supply falls and the preferred tax rate falls, then the subsequent rate of economic growth increases because distortionary taxes fall and capital accumulation is less constrained.

[27] The empirical model presented above incorporates panel estimation as well as difference and system GMM estimation (and some sensitivity check), which plausibly indicates that it is adequate for distinguishing the impact of a second order effort, and it could engage in the discussion of inequality and growth in current literature.

The relationship between inequality and growth is tested in a panel of 19 OECD countries, augmenting Forbes (2000) to include capital income inequality as an additional explanatory variable. The measure of capital income inequality in the analysis is the top 1% income share. Consistent with the theory, subsequent growth rate is found to be positively associated with capital income inequality. Moreover controlling for the top income share renders a consistently negative estimate for the impact of labor income inequality on growth, in line with the original Persson and Tabellini (1994) hypothesis. The positive impact of capital income inequality on growth survives in a variety of econometric specifications, including when difference and system GMM techniques are employed.

Chapter 4

Demography and the Composition of Taxes

4.1 Introduction

How does population aging affect fiscal policy? Razin et al. (2002), building on Meltzer and Richard (1981) and Saint-Paul (1994), theorize that increases in the dependency ratio lead to lower labor tax rates and a reduction in the generosity of social transfers in democracies. The logic is that the median voter is a worker, who increasingly dislikes redistributing as the retired population increases. This chapter develops the Razin et al. (2002) hypothesis to consider the composition of taxes, in particular the setting of income versus expenditure taxes. The main theoretical prediction is that the extent of taxes on income relative to expenditure falls with population aging. The logic is similar to the original paper. Income taxes are paid solely by workers, whilst expenditure taxes are paid by both generations. If the median voter is a worker, then increasing the size of the retired population compels a shift in tax composition towards expenditure taxes. International panel evidence supports this hypothesis.

Empirical evidence generally has not supported the Razin et al. (2002) hypothesis, which focuses only on income taxes. For example, Figure 4.1 depicts the raw correlation between

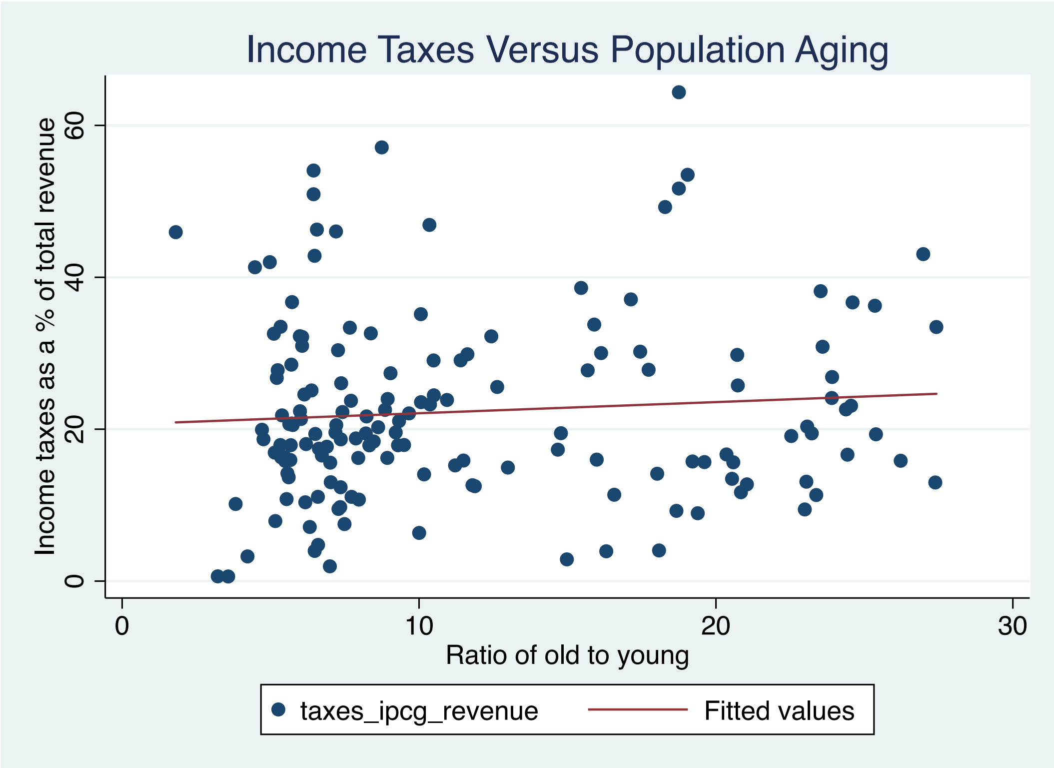


Fig. 4.1 Correlation between Aging and Income Taxes, 1990-2014

income taxes as a portion of total revenue over 1990-2014 and the ratio of the population above 65 to the population between 15 and 64, indicating no negative association between aging and the level of income taxes. On the other hand, Disney (2007), Sanz and Velazquez (2007), and Shelton (2008) all find a positive relationship between population aging and the size of the welfare state in regression analyses. In response to this finding, a new theoretical literature has appeared through which a larger proportion of retirees can lead to higher income taxes and more generous social transfers. Galasso and Profeta (2007) propose that population aging has a political effect: the median voter becomes older and hence more willing to support a larger system. Further theoretical analyses studying welfare states and aging, as also argued by Simonovits (2007), extend beyond the mechanism analyzed in Razin et al. (2002), trying to link increasing income taxes to aging.

The theory proposed in this chapter is motivated by recent rises in taxes on goods and services and concurrent population aging. Revenue sources outside of income taxes (i.e.

expenditure taxes) are thus empirically becoming increasingly important components of total revenue (e.g. approximately 30% as a share of total tax revenue in the United Kingdom). This is potentially a paradox as it might have been expected that countries with a relatively larger retired population would have higher labor income taxes and lower expenditure taxes, reflecting the increased political clout of the retired population who presumably would prefer labor income taxes.

Given the increased empirical importance of expenditure taxes this chapter examines how population aging affects the composition of taxes. Despite the fact that there is an enormous literature focusing on optimal taxation, starting with Diamond and Mirrlees (1971), relatively little of this literature provides a positive analysis of the political-economic determination of tax composition. One related work by Pickering and Rajput (2018) links the composition of taxes to inequality. Other related work analyzes the adoption of tax instruments with the development process (see Aidt and Jensen, 2009; Keen and Lockwood, 2010), but this literature neglects the effect of demography on the composition of taxes.

In an overlapping generations model, taxes are levied on both income and expenditure, which finance redistribution to both the working and retired population, with a balanced budget period by period.[1] As in Razin et al. (2002) the choice of the median voter is the unique Condorcet winner. In the case of a positive population growth rate, the median voter is a young individual (because the younger generation is more numerous than the older generation), hence determining political equilibrium tax rates.

The theoretical findings in this chapter cohere with Razin et al. (2002) in that an increased fraction of retirees in the population always leads to lower labor income taxes as long as the decisive voter is young. The results relating to expenditure taxes are novel - it depends on the initial level of expenditure taxes. At lower initial levels of expenditure taxes, an increase in the share of retirees in the population also leads to lower expenditure tax rates. However, if the initial levels of expenditure taxes pass some threshold level, then the increased size of the retired population leads to higher expenditure taxes, even if this comes at the price of greater

[1] I have been aware of high debt-to-GDP ratio especially in OECD countries. Note that this chapter mainly focuses on the composition of taxes rather than debt decision.

deadweight expenditure tax losses, because the decisive voter wants to shift the tax burden onto the retired population.[2] Nonetheless, the main proposition is that the composition of taxes, defined as the extent to which taxes are levied on income relative to expenditure, unambiguously falls with the share of retirees.

The relationship between population aging and the extent of taxes on income relative to taxes on expenditure is investigated empirically using a panel of over 100 countries over the period 1990-2014. Following Pickering and Rajput (2018) the dependent variable is constructed by the ratio of taxes on income, profits and capital gains (as a share of total tax revenue) to taxes on goods and services (as a share of total tax revenue), and the key demography measure is the percentage of the population over the age of 65. These data are all taken from the World Development Indicators database. The empirical work analyzes both panel data and cross-country regressions using within-country averages. Consistent with the theory, the extent of taxes on income relative to expenditure is found to be negatively associated with the fraction of the retired population. This relationship is robust across different econometric specifications employed. In the panel estimation with fixed effects, a one standard deviation increase in the fraction of the retired population is statistically associated with a fall of 0.63 in the ratio of income to expenditure taxes, holding all else equal. The magnitude of this estimated correlation is sizable - implying more than a half of the raw standard deviation in the policy variables. This negative relationship holds up significantly in countries with higher degrees of democracy, in support of the mechanism proposed in this chapter. Moreover I also separately examine how the different tax instruments co-move with the proportion of retirees. The data indicate a negative correlation between income taxes and the fraction of the retired population, and a positive correlation between expenditure taxes and the fraction of the retired population.

The next section theoretically analyzes how the composition of taxes changes with population aging. Section 4.3 contains the empirical work, and section 4.4 concludes.

[2] In practice, Japan, as a country with problem of aging, provides some interesting anecdotes. In April 2014 the government of Japan increased its expenditure tax rate from 5% to 8%, and after that, they are currently planning to again raise the rate to 10% in October 2019.

4.2 The Model

4.2.1 Economic Environment

This section analyzes the impact of population aging on the composition of taxes. The model extends Razin et al. (2002) to include expenditure taxes as well as income taxes, consisting of an overlapping generations model with population growth rate (n), where individuals live for two periods: a working (young) period and a retirement (old) period. The utility u of an individual born in period t depends on his consumption in the two periods ($c_{1,t}$ and $c_{2,t+1}$):

$$u_t = u(c_{1,t}, c_{2,t+1}) \tag{4.1}$$

where u is a strictly concave and twice-differentiable utility function. Due to the presence of expenditure (or consumption) taxes $\tau_{c,t}$ and $\tau_{c,t+1}$, consumption is less than expenditure in period t ($x_{1,t}$) and in period $t+1$ ($x_{2,t+1}$):

$$c_{1,t} = (1 - \tau_{c,t})x_{1,t} \tag{4.2}$$

$$c_{2,t+1} = (1 - \tau_{c,t+1})x_{2,t+1}. \tag{4.3}$$

First-period expenditure, $x_{1,t}$, and second-period expenditure, $x_{2,t+1}$, are determined by the budget constraints

$$x_{1,t} = (1 - \tau_{y,t})y_t + r_t \tag{4.4}$$

$$x_{2,t+1} = r_{t+1} \tag{4.5}$$

where labor income y_t is taxed at a linear income tax rate $\tau_{y,t}$, and r_t and r_{t+1} are lump-sum redistribution in period t and in period $t+1$.

The budget of the government is assumed to be balanced period by period. Redistribution (r_t) is financed by consumption and income tax revenue, and is paid to both working and retired people as in Razin et al. (2002) (r_t is assumed to be equal for young and old, as in Razin et al. (2002)), hence

$$r_t N_0\Big[(1+n)^{t-1}+(1+n)^t\Big] = \tau_{c,t}\bar{x}_t N_0\Big[(1+n)^{t-1}+(1+n)^t\Big]+\tau_{y,t}\bar{y}_t N_0(1+n)^t \qquad (4.6)$$

where N_0 is the initial size of young individuals, and $\bar{x}_t$ and $\bar{y}_t$ are the average levels of expenditure and income in period t. Furthermore expenditure equals income at the aggregate level in period t,

$$\bar{x}_t N_0\Big[(1+n)^{t-1}+(1+n)^t\Big] = \bar{y}_t N_0(1+n)^t. \qquad (4.7)$$

Combining (4.6) and (4.7), lump-sum redistribution equals

$$r_t = (\tau_{c,t}+\tau_{y,t})\frac{1+n}{2+n}\bar{y}_t. \qquad (4.8)$$

Since the budget of government is balanced period by period, it follows that redistribution in period $t+1$, r_{t+1}, is independent of tax rates, $\tau_{y,t}$ and $\tau_{c,t}$, in period t, analogous to Razin et al. (2002). In voting on tax rates $\tau_{y,t}$ and $\tau_{c,t}$, individuals living in period t thus take r_{t+1} as exogenous because there is no serial dependence between r_t and r_{t+1}. The political-economic equilibrium for tax rates, $\tau_{y,t}$ and $\tau_{c,t}$, is then determined by majority voting of individuals alive in period t, without being affected by preceding or future generations.

A final ingredient is that mean income is modeled to be declining in taxes in order to capture the spirit of Meltzer and Richard (1981) as in Pickering and Rockey (2011),

$$\bar{y}_t = y_t^*(1-\frac{\delta_y\tau_{y,t}\bar{y}_t}{y_t^*}-\frac{\delta_c\tau_{c,t}\bar{y}_t}{y_t^*}) \qquad (4.9)$$

where y_t^* is potential income, and $0<\delta_y<1$ and $0<\delta_c<1$ capture the sensitivity of income respectively to income and expenditure taxes. The parameters δ_y and δ_c reflect deadweight losses, either as the result of the costs of tax collection and/or their influences on economic activity. High values of δ_y and δ_c indicate high costs for taxes collection or alternatively low levels of tax base. The properties of (4.9) indicate that $\frac{d\bar{y}_t}{d\tau_{y,t}}=-\delta_y\bar{y}_t$ and $\frac{d\bar{y}_t}{d\tau_{c,t}}=-\delta_c\bar{y}_t$, and hence that the proportionate deadweight losses, $\frac{d\bar{y}_t/d\tau_{y,t}}{\bar{y}_t}$ and $\frac{d\bar{y}_t/d\tau_{c,t}}{\bar{y}_t}$, are constant, which therefore rules out scale effects.

4.2.2 Income Taxes

This section now addresses policy-setting. The pivotal voter alive in period t chooses the vector of policies $\mathbf{q} = \{\tau_{y,t}, \tau_{c,t}, r_t\}$ in order to maximize their utility. Substituting (4.4) and (4.5) into (4.2) and (4.3) and then substituting into (4.1) and using (4.8) gives

$$u_t = u\Big[(1-\tau_{c,t})\Big((1-\tau_{y,t})y_t + (\tau_{c,t}+\tau_{y,t})\frac{1+n}{2+n}\bar{y}_t\Big), (1-\tau_{c,t+1})r_{t+1}\Big] \tag{4.10}$$

which indicates that the policy problem is multidimensional (τ_c and τ_y).[3] Given that $y_t = \bar{y}_t$, all young people in period t have the same level of income, and hence, the same utility function as above.[4] As long as the rate of population growth is positive, $n > 0$, the condition that there are more young individuals than old individuals (or more working individuals than retired individuals) always holds. This implies that the median voter in determining equilibrium tax rates is still among the working-age population. Therefore, the preferred policy of the median voter is the unique Condorcet winner even though the policy problem is two-dimensional.

For a given n, the political equilibrium (τ_y or τ_c) is constant over time, so that the time subscript t is suppressed henceforth.

Maximization of (4.10) with respect to τ_y, given (4.9), yields:

$$\begin{aligned}\frac{\partial u}{\partial \tau_y} &= V(\tau_y, n) \\ &= -y_d + \frac{1+n}{2+n}\bar{y} - (\tau_c+\tau_y)\frac{1+n}{2+n}\delta_y\bar{y} = 0\end{aligned} \tag{4.11}$$

[3] In voting on tax rates in period t individuals take $\tau_{c,t+1}$ as exogenous.

[4] Ideally I would incorporate hetergeneity in income, but this leads to no existence of a Condorcet winner in the multidimensional policy problem according to the indirect utility function above. Note, however, that the results still hold if income hetergeneity is incorporated as this assumption is only for simplicity (but without loss of generality).

where y_d is the income of the decisive voter. The mathematical derivations are contained in the Appendix C.1. Equation (4.11) is analogous to the well-known condition (13) in Meltzer and Richard (1981) and equation (11) in Razin et al. (2002) which determines the preferred income tax rate of the decisive voter. Equation (4.11) then delivers the result in Razin et al. (2002) that an increase in n raises the political equilibrium income tax rate as discussed below.

According to equation (4.11), τ_y is defined by the first-order condition. Given the assumption that u is strictly concave, I have the second-order condition

$$\frac{\partial^2 u}{\partial \tau_y^2} = \frac{\partial V(\tau_y, n)}{\partial \tau_y} \leq 0. \tag{4.12}$$

Here I examine the effect of changes in the population growth rate and thus population aging on the equilibrium income tax rate. Total differentiation of (4.11) with respect to n implies

$$\frac{d\tau_y}{dn} = -\frac{{}^{\partial V(\tau_y,n)}\big/_{\partial n}}{{}^{\partial V(\tau_y,n)}\big/_{\partial \tau_y}}. \tag{4.13}$$

Since $\frac{\partial V(\tau_y,n)}{\partial \tau_y} \leq 0$ (see eq. [4.12]), it follows that the direction of the effect of changes in n on the equilibrium income tax rate, τ_y is determined by the sign of $\frac{\partial V(\tau_y,n)}{\partial n}$. By differentiating equation (4.11) with respect to n, I conclude that

$$\frac{\partial V(\tau_y, n)}{\partial n} = \frac{1}{(2+n)^2}\bar{y} - (\tau_c + \tau_y)\frac{1}{(2+n)^2}\delta_y \bar{y}. \tag{4.14}$$

If the sign of $\partial V(\tau_y, n)/\partial n$ is positive, then an increase in the rate of population growth, n, increases the political economy equilibrium income tax rate, τ_y. On inspection of the right-hand side of (4.14), it contains one positive term, $\frac{\bar{y}}{(2+n)^2}$, whereas the other term is negative. Consider for concreteness the case in which the decisive voter is young and the population growth rate rises (the fraction of retirees falls). In this case, there is a decline in the amount of tax revenue collected from the median voter that “leaks” to the retirees, who become a smaller share of the population with the higher n. This is a pro-tax factor. However, the per capita marginal efficiency cost of distortionary taxation, $-(\tau_c+\tau_y)\frac{1}{(2+n)^2}\delta_y\bar{y}$, rises, which is an anti-tax factor. In this instance, there is no ambiguity: given that $\tau_c+\tau_y<1$ and $\delta_y<1$, the pro-tax factor always dominates the anti-tax factor and hence $\partial V(\tau_y,n)/\partial n$ is positive.

Lemma 4.1. *The equilibrium income tax rate rises as the rate of population growth rises, i.e.,*

$$\frac{d\tau_y}{dn}>0.$$

In this case, an increase in n (smaller fraction of retirees) increases the political economy equilibrium income tax rate, with exactly the same underpinning as that provided in Razin et al. (2002) (who only consider income taxes). The aging of population affects the equilibrium income tax rate in two directions: the increased number of retirees raises the demand for benefits while at the same time restrains the willingness of the working-age population to accede to higher income taxes, as current workers are net losers from the welfare state. If the decisive voter is not among the retirees, as is still the case in all western countries, then the increased size of the retired population leads to lower income taxes, since the decisive voter is adversely affected because she is a net contributor to the welfare system.

4.2.3 Expenditure Taxes

Now I turn to the political equilibrium choice of the expenditure tax rate. Maximization of (4.10) with respect to τ_c, given (4.9), yields

$$\begin{aligned}\frac{\partial u}{\partial \tau_c}&=W(\tau_c,n)\\&=(1-\tau_c)\left[\frac{1+n}{2+n}\bar{y}-(\tau_c+\tau_y)\frac{1+n}{2+n}\delta_c\bar{y}\right]-\left[(1-\tau_y)y_d+(\tau_c+\tau_y)\frac{1+n}{2+n}\bar{y}\right]=0.\end{aligned} \tag{4.15}$$

The mathematical derivations are again contained in the Appendix C.1. According to (4.15), τ_c is defined by the first-order condition. Given the assumption that u is strictly concave, I have the second-order condition

$$\frac{\partial^2 u}{\partial \tau_c^2} = \frac{\partial W(\tau_c, n)}{\partial \tau_c} \leq 0. \tag{4.16}$$

I next examine the effect of changes in the population growth rate and thus population aging on the equilibrium expenditure tax rate. Total differentiation of (4.15) with respect to n implies

$$\frac{d\tau_c}{dn} = -\frac{\partial W(\tau_c,n)/\partial n}{\partial W(\tau_c,n)/\partial \tau_c}. \tag{4.17}$$

Since $\frac{\partial W(\tau_c,n)}{\partial \tau_c} \leq 0$ (see eq. [16]), it follows that the direction of the effect of changes in n on the equilibrium expenditure tax rate, τ_c, is determined by the sign of $\frac{\partial W(\tau_c,n)}{\partial n}$.

By differentiating equation (4.15) with respect to n, I conclude that

$$\frac{\partial W(\tau_c, n)}{\partial n} = (1-\tau_c)\left[\frac{1}{(2+n)^2}\bar{y} - (\tau_c+\tau_y)\frac{1}{(2+n)^2}\delta_c \bar{y}\right] - (\tau_c+\tau_y)\frac{1}{(2+n)^2}\bar{y}. \tag{4.18}$$

If the sign of $\partial W(\tau_c, n)/\partial n$ is negative, then an increase in the rate of population growth, n, decreases the political economy equilibrium expenditure tax rate, τ_c. The right-hand side of (4.18) contains one positive term, $(1-\tau_c)\frac{\bar{y}}{(2+n)^2}$, whereas the other terms are negative. Again consider for concreteness the case in which the decisive voter is young and the population growth rate rises (the fraction of retirees falls). As with the case of income taxes, within the bracket the protax factor always dominates the distortionary element (which is an antitax factor), $-(\tau_c+\tau_y)\frac{1}{(2+n)^2}\delta_c\bar{y}$, given that $\tau_c + \tau_y < 1$ and $\delta_c < 1$, so the whole term

$$(1-\tau_c)\left[\frac{1}{(2+n)^2}\bar{y} - (\tau_c+\tau_y)\frac{1}{(2+n)^2}\delta_c \bar{y}\right]$$

is positive. Note that the last term on the right-hand side of (4.18) is also negative, which indicates that the sign of $d\tau_c/dn$ depends on the value of τ_c.

Lemma 4.2. *The effect of changes in the population growth rate on the equilibrium expenditure tax rate is given by:*

$$\frac{d\tau_c}{dn}\begin{cases} > 0 & \text{if } \tau_c < \tau_c^* \\ < 0 & \text{if } \tau_c > \tau_c^* \end{cases}$$

This lemma describes a threshold relation. At lower initial levels of τ_c, an increase in the population growth rate also leads to higher expenditure tax rates as with the income tax rates because at low tax levels the deadweight losses are relatively small. However, if τ_c is greater than some threshold level, greater deadweight expenditure tax losses drive the decisive voter to opt for lower tax rates.

At higher initial levels of τ_c, an increase in n (smaller fraction of retirees) decreases the political economy equilibrium expenditure tax rate. The aging of population affects the equilibrium expenditure tax rate in two directions: the greater number of retirees raises the demand for benefits, while at the same time the willingness of the working-age and retired population to progress to higher expenditure taxes reduces, since both working and retired people are required to contribute to the welfare state. If the decisive voter is among the working population, then the increased size of the retired population leads to higher expenditure taxes, even if this comes at the price of greater deadweight expenditure tax losses, since the decisive voter wants to raise the tax burden on the retired population, and larger fraction of retirees indicates larger size of taxation base from expenditure from the old.

4.2.4 The Composition of Taxes

Now I consider the composition of taxes. The ratio of income to expenditure taxes is given by

$$T \equiv \frac{\tau_y}{\tau_c}. \tag{4.19}$$

Combining equations (4.11) and (4.15) yields proposition 4.1. The proof of this is in the Appendix C.2.

Proposition 4.1. *The ratio of income to expenditure taxes rises as the rate of population growth rises, and the ratio falls as n falls, i.e.,*

$$\frac{dT}{dn} > 0.$$

Proposition 4.1 is straightforward but novel. Increases in the fraction of the population that is retired (or falls in population growth rate) lead to increases in consumption taxes relative to income taxes. If the fraction of the retired population increases, then income taxes fall due to smaller proportion of working-age population who are net contributing to income taxes given all else equal. In order to guarantee the degree of redistribution, there will be a tradeoff between income and expenditure taxes preferred by the decisive voter. Although the political clout of the retired population who presumably would prefer labor income taxes is increasing, the working-age population (including the decisive voter) wants to shift the tax-burden onto the retired population, rather than being taxed labor income only from the young who are acting as a net contributor to the welfare system. In short, the extent to which taxes are levied on income relative to expenditure unambiguously falls with the share of retirees.

For completeness, this chapter shall also consider briefly the case in which the median voter is among the retired population.[5] The political economy equilibrium income tax rate in this case maximizes the transfer r. In contrast, when the median voter is a member of the working-age population, the political economy equilibrium income tax rate maximizes r plus another term, $(1-\tau_y)y$ that is decreasing in τ_y.[6] Thus the political economy equilibrium income tax rate "jumps" upward when the old become a majority, that is, as the young/old balance switches from being more young to being more old. While the political economy equilibrium expenditure tax rate is independent of the young/old balance if the median voter is among the retired population. In this case expenditure tax rate depends on the value

[5]This may occur if population growth happens through longevity rather than fertility, or the old are more likely to vote than the young.

[6]The median voter is not among the retirees - as is probably still the case in all developed countries as well as developing countries.

of income tax rate (this means that expenditure tax rate could be increasing in τ_y). These opposing effects on the ratio of income to expenditure tax rate are next examined empirically.

4.3 Evidence

4.3.1 Data and Descriptive Statistics

The main agenda in this section is to test the hypothesis proposed above - whether and how the ratio of income to expenditure taxes across countries and time systematically changes with the fraction of the population that is retired. This empirical analysis focuses on a panel of over 100 countries. Cross-country annual data on income and expenditure tax revenue are available over the period 1990-2014 from the World Development Indicators (hereafter WDI). This chapter also reports results from cross-country regressions with data measured by within-country averages. Moreover I separately examine how different categories of tax measures respectively co-move with the fraction of retirees.

Following Pickering and Rajput (2018), the main dependent variable is the ratio of income taxes to expenditure taxes, $T = \frac{\tau_y}{\tau_c}$, constructed by the ratio of taxes on income, profits and capital gains (as a percentage share of total tax revenue) to taxes on goods and services (as a percentage share of total tax revenue). Both are extracted from the WDI database. In practice (and also within countries) rates of tax vary with different types of income and goods, but the measure of ratio proposed is a way to capture the extent to which taxes are levied on income relative to expenditure. Due to the relatively small value in the data of taxes on goods and services in the case of some countries, following Pickering and Rajput (2018) I use the natural logarithm of T, $\ln(T)$, in the below regression analysis. The argument proposed predicts that the extent of taxes on income relative to expenditure declines with an increased fraction of the retired population.

The measure of the proportion of retirees in the population used is the percentage of the population over the age of 65 (denoted $PROP65$), which is also taken from the WDI database.

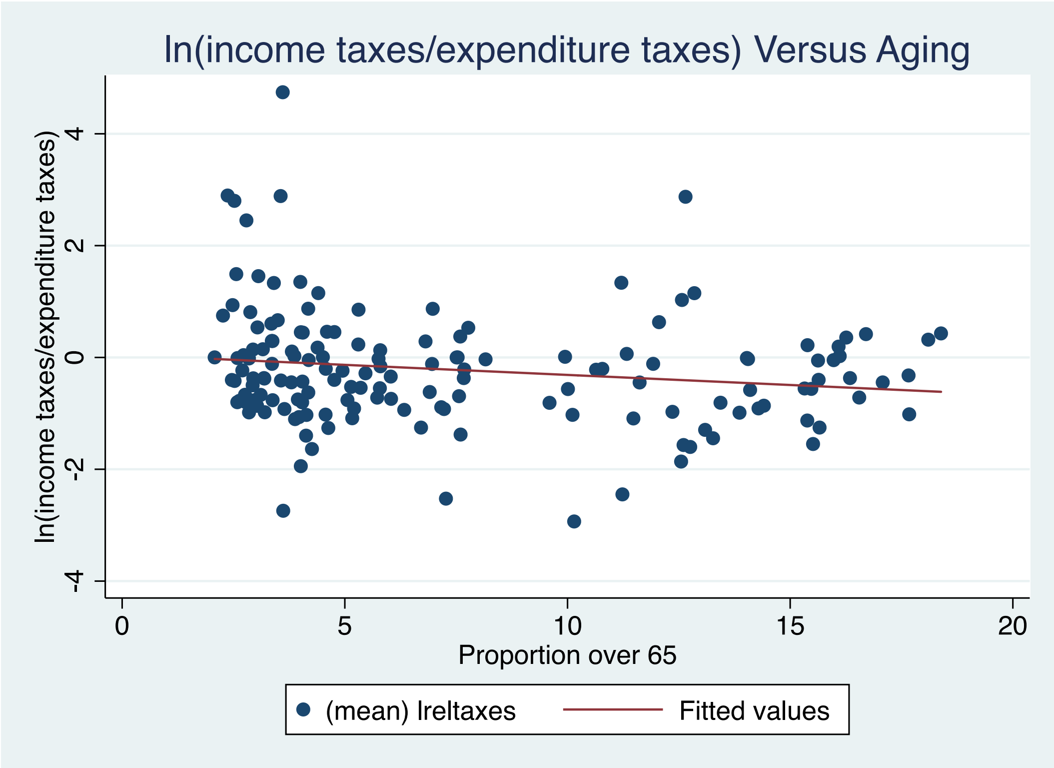


Fig. 4.2 Correlation between Aging and the Composition of Taxes, 1990-2014

This measure of the retired fraction is preferable to the dependency ratio used by Razin et al. (2002). The dependency ratio includes children as well as retirees which would have different impacts on taxes as shown by Shelton (2008). Figure 4.2 depicts a scatter plot of the logarithm ratio of income to expenditure taxes and $PROP65$, showing a negative relationship, in support of the proposed theory.

Following Pickering and Rajput (2018), one important determinant of the composition of taxes is the level of development, so I include the natural logarithm of GDP per capita in constant chained PPP US$ ($\ln(y)$), taken from the Penn World Tables, as a first control in the regression analysis. As another measure of the development level and institutional capacity, hence OECD membership (denoted $OECD$) is also employed as a further control variable. To fully capture demographic effects, the econometric analysis includes the percentage of the population between 15 and 64 years of age (denoted $PROP1564$, also from the WDI database) as an additional control. Another potential determinant of the composition of taxes is inequality (Pickering and Rajput, 2018), taken from the University of Texas Inequality Project's Estimated Household Income Inequality data of Galbraith and Kum (2005), so the inequality measure (denoted $UTIP$) is also included as a control.

Governments collect tax revenue through means beyond taxation on income and consumption. One important source is the revenue from import duties and tariffs due to openness (Rodrik, 1998). Thus the trade share (the sum of exports and imports as a percentage of GDP - denoted $TRADE$) is also employed in the regression analysis. Apart from these control variables the natural logarithm of the total population size (denoted $\ln(POP)$) is included as well, to some extent capturing any scale (dis-)economies related to particular forms of tax collection.[7] These data are also taken from the WDI database.

There may also be cyclical movements in policy variables. To address this potential problem the regression analysis includes the output gap (the difference between aggregate output and its trend value in percentage - denoted $YGAP$) as a further control.

The policy variables may also be affected by the degree of democracy through various channels, so the democracy score (with -10 denoting the highest level of autocracy, and 10 denoting the highest level of democracy) provided by the Polity IV project is included as a final control (denoted $POLITY2$).

Table 4.1 contains descriptive statistics of the variables used in the regression analysis. Note that there is considerable dispersion in how countries raise their tax revenue. Over the sample period taxes on income on average represent a smaller fraction of total revenue than taxes on goods and services (Figure 4.3 depicts a scatter plot of the two series).[8] This indicates that the capacity to raise revenue through income taxes is normally limited in those countries with low income. For instance in the OECD members, taxes on income are on average 32% of total revenue, whilst outside the OECD income taxes account for just 20% of total revenue on average (Figure 4.4 depicts a line plot of the two series).

[7]For example, larger size of population means more sources of tax collection, which may lead to scale economies or scale diseconomies depending on its size.

[8]For example, Macao has a very large size of expenditure taxes relative to income taxes.

Table 4.1 Descriptive Statistics

	obs	mean	std. dev.	min	max
$\ln(T)$	2091	-0.261	1.00	-4.00	5.09
τ_y	2142	22.45	12.75	0.349	66.72
τ_c	2141	29.06	13.76	0.024	89.22
$PROP65$	4633	7.01	4.72	0.335	25.08
$PROP1564$	4633	61.20	7.01	45.29	85.81
$UTIP$	1556	42.79	6.71	22.75	59.96
$\ln(y)$	3652	8.58	1.29	5.03	11.73
$OECD$	5325	0.138	0.345	0	1
$TRADE$	4263	86.94	51.96	0.309	531.7
$\ln(POP)$	5302	15.07	2.35	9.11	21.04
$POLITY2$	3790	3.04	6.69	-10	10
$YGAP$	4668	0	0.034	-0.609	0.505

Notes: τ_y denotes taxes on income, profits and capital gains as a percentage of revenue - taken from the World Development Indicators (WDI). τ_c denotes taxes on goods and services as a percentage of revenue - also taken from the WDI. $T = \frac{\tau_y}{\tau_c}$. $PROP1564$ and $PROP65$ are respectively the proportion of the population aged between 15 and 64, and 65 and above. $UTIP$ is the University of Texas Inequality Project's Estimated Household Income Inequality. y is real GDP at chained PPPs in millions of 2005 US dollars per capita - taken from the Penn World Tables. $OECD$ is a dummy variable denoting OECD membership. $TRADE$ is the sum of exports and imports as a percentage of GDP. POP is the size of country population. $POLITY2$ is a measure of democracy provided by the Polity IV project, with -10 denoting the highest level of autocracy, and 10 denoting the highest level of democracy. $YGAP$ is the difference between the actual output and its trend value in percentage.

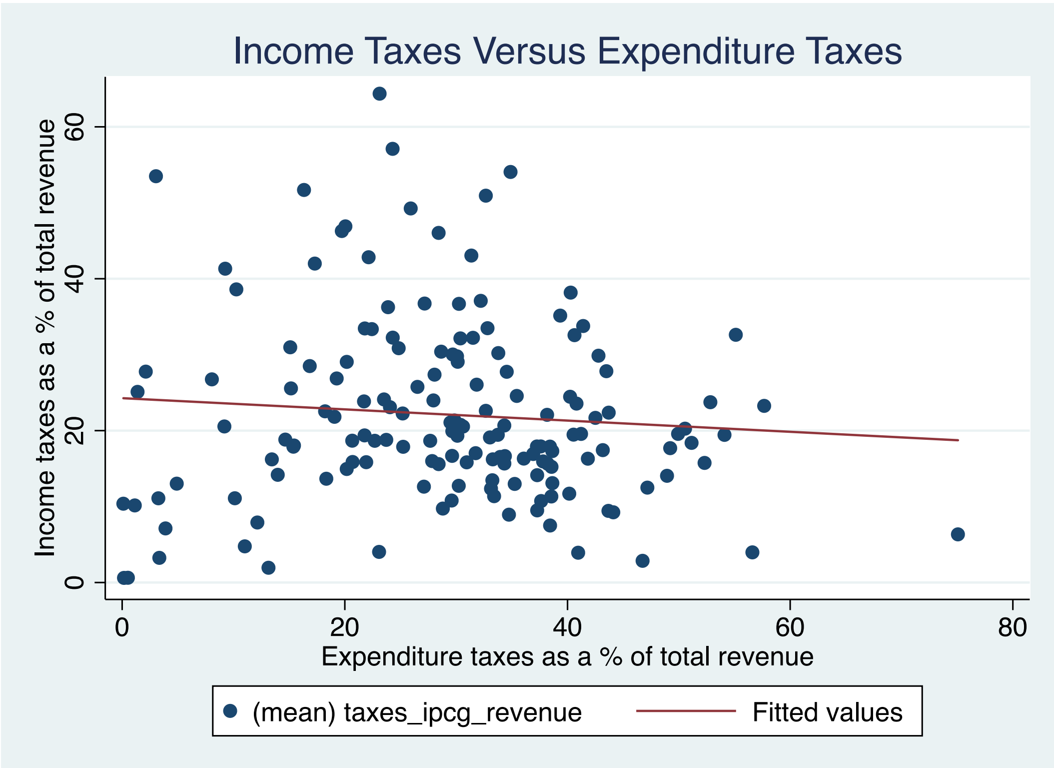


Fig. 4.3 Correlation between Income Taxes and Expenditure Taxes, 1990-2014

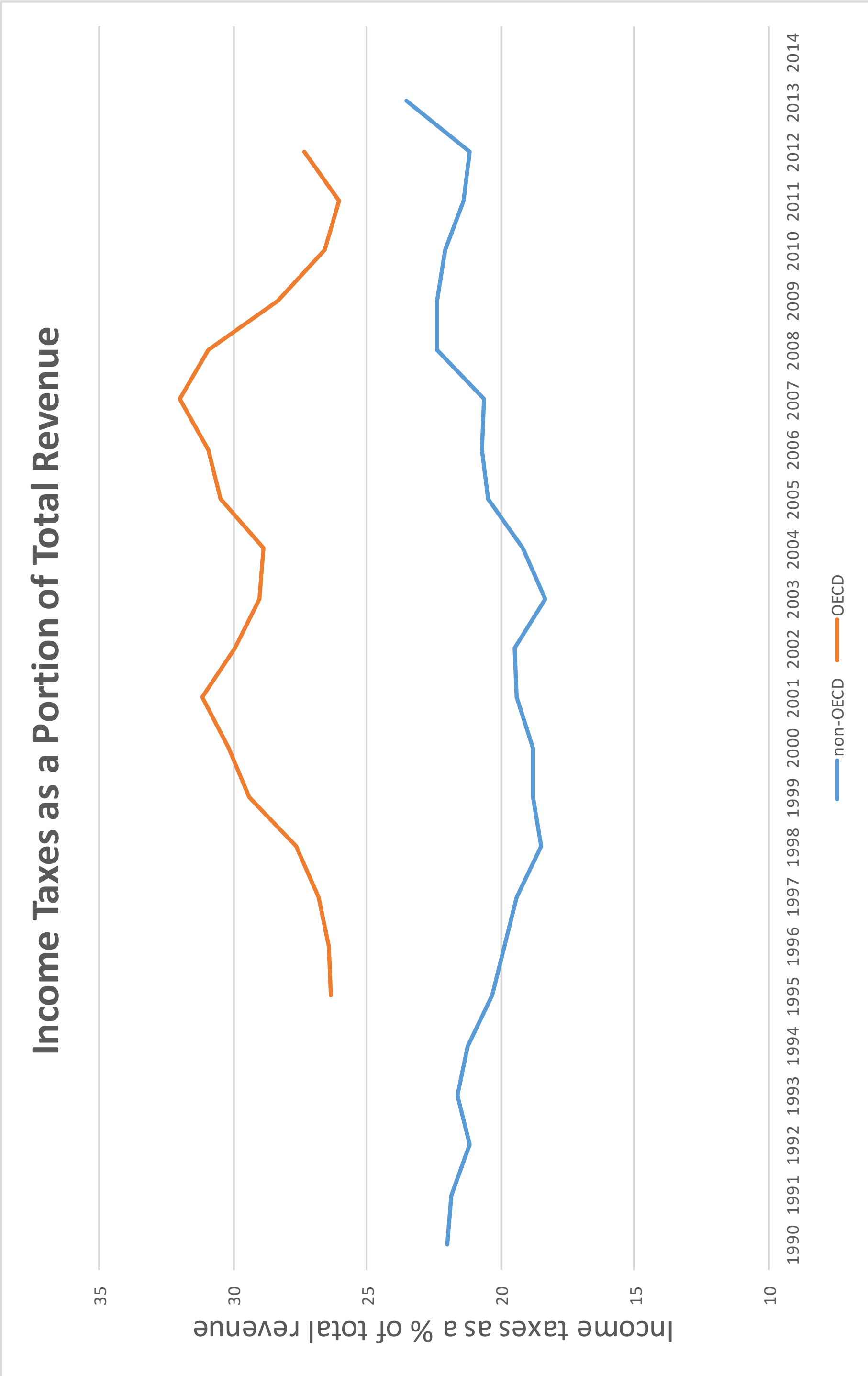


Fig. 4.4 Income Taxes in OECD and non-OECD countries, 1990-2014

The *PROP*65 data cover 155 countries, and numerically range from 1.37 (Qatar) to 18.39 (Italy) on average, with higher numbers meaning greater proportion of the retired population. Notably, these data are positively correlated with GDP per capita across the sample, with a correlation coefficient of around 0.72. Richer countries have greater retired fraction of the population than poorer countries. This emphasizes the need to include controls for the level of economic development, or the retired fraction may become a proxy for other drivers of the policy variables.

4.3.2 Panel Estimation

Columns 1 and 2 of Table 4.2 contain estimation results examining the impact of population aging on the ratio of income to expenditure taxes using OLS. Column 1 is a simple specification with the fraction of the retired population (*PROP*65) and a number of control variables using annual data, with robust standard errors clustered by country. Column 2 extends the regression of column 1 to include time effects. The use of time effects will substantially control for the potential problem of a secular trend. In these specifications the sign of the coefficient estimate relating to the fraction of the retired population is negative in all cases, and all are statistically significant at the 1% level. This is consistent with the theory - an increase in the retired fraction increases expenditure taxes relative to income taxes. Columns 3 and 4 repeat the analysis of columns 1 and 2 using country fixed effects panel estimation instead. The results using panel estimation support those already found. The estimated statistical significance of the fraction of the retired population is unaffected and even remains at the 5% level in column 3. Using the estimate from column 3 of Table 4.2, a one standard deviation increase in the fraction of the retired population is statistically associated with a fall of 0.63 in the ratio of income to expenditure taxes, holding all else equal. The magnitude of this estimated correlation is sizable - implying more than a half of the raw standard deviation in the policy variables.

Table 4.2 Basic Estimation Results - the Composition of Taxes (annual data)

Dep Var: ln(T)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
*PROP*65	-0.0954***	-0.104***	-0.133**	-0.125*	-0.124**	-0.0722		
	(0.0231)	(0.0243)	(0.0519)	(0.0643)	(0.0576)	(0.0625)		
*PROP*1564	-0.0716***	-0.0755***	-0.0525**	-0.0452**				
	(0.0190)	(0.0193)	(0.0246)	(0.0224)				
RATIO							-0.0647*	-0.0231
							(0.0338)	(0.0354)
UTIP	-0.000347	-0.00770	-0.0126	-0.00986	-0.0145	-0.00911	-0.0162	-0.00933
	(0.0143)	(0.0153)	(0.0184)	(0.0175)	(0.0208)	(0.0191)	(0.0222)	(0.0192)
ln(y)	0.715***	0.713***	1.072***	1.024***	0.842***	0.853***	0.755***	0.832***
	(0.156)	(0.151)	(0.231)	(0.226)	(0.202)	(0.196)	(0.189)	(0.191)
OECD	0.311	0.338	-0.0399	0.0703	-0.0406	0.0825	-0.0661	0.0674
	(0.226)	(0.221)	(0.0833)	(0.117)	(0.0704)	(0.114)	(0.0712)	(0.121)
TRADE	0.00112	0.000618	0.000637	0.000294	0.00121	0.000895	0.000832	0.000792
	(0.00133)	(0.00131)	(0.00212)	(0.00232)	(0.00207)	(0.00235)	(0.00202)	(0.00243)
ln(*POP*)	0.150***	0.137**	0.00172	0.430	-0.508	0.535	-0.513	0.782
	(0.0562)	(0.0575)	(0.610)	(0.948)	(0.598)	(0.956)	(0.610)	(0.980)
*POLITY*2	-0.0166	-0.0189	-0.0134	-0.0110	-0.0183	-0.0139	-0.0201*	-0.0144
	(0.0214)	(0.0217)	(0.0124)	(0.0121)	(0.0112)	(0.0113)	(0.0109)	(0.0114)
YGAP	0.253	-0.132	-0.937	-0.541	-0.964	-0.512	-1.010	-0.515
	(0.989)	(1.028)	(0.734)	(0.707)	(0.792)	(0.736)	(0.817)	(0.737)
Observations	796	796	796	796	796	796	796	796
Countries	87	87	87	87	87	87	87	87
Data	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel
Fixed Effects?	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects?	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.422	0.435	0.172	0.195	0.143	0.178	0.130	0.175

Notes: Columns (1) and (2) contain results using OLS regressions of the composition of taxes, ln(T), including *PROP*1564, *UTIP*, ln(y), *OECD*, *TRADE*, ln(*POP*), *POLITY*2, and *YGAP* as control variables. Columns (3) and (4) contain results using Panel regressions with country fixed effects. Columns (5) and (6) again test columns (3) and (4) without *PROP*1564 as a control. Columns (7) and (8) instead use the ratio of the population above 65 to those between the ages of 15 and 64, *RATIO*, to measure population aging, and mimic columns (3) and (4). Robust standard errors are shown in parentheses. Standard errors are clustered by country. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

One may argue that the existence of $PROP1564$ counteracts the efficiency of $PROP65$ as a measure of population aging. To address the potential counteractive effect of $PROP1564$, columns 5 and 6 of Table 4.2 again test columns 3 and 4 with full control variables except $PROP1564$. Further, columns 7 and 8 instead use the ratio of the population above 65 to those between the ages of 15 and 64, $RATIO$, as in Shelton (2008), to measure population aging, and mimic columns 3 and 4. The results similarly demonstrate an increased tendency to use expenditure taxes rather than income taxes as population aging increases.

It is natural to investigate whether or not the results reported change with the degree of democracy, given that the theory proposed is based on the median voter framework. Table 4.3 thus extends the regression results by splitting the sample by levels of democracy. Column 1 contains results for countries with stronger democratic credentials (i.e. with the democracy score $POLITY2$ of 7 or above over the sample period). Column 2 contains results for countries with weaker democratic credentials (i.e. with $POLITY2$ of less than 7). The democracy criterion is strengthened further in columns 3 and 4, and the sample is split according to $POLITY2 \gtrless 8$. When the sample is separated it becomes clear that the negative relationship between the fraction of the retired population and the ratio of income to expenditure taxes holds only in the subsample of democratic regimes. This is in line with the theory, which assumes a complete franchise. In column 3 the p-value for the estimated coefficient for the fraction of the retired population is 1.1%, and the estimated effect is sizable: A one standard deviation increase in the fraction of the retired population is statistically associated with the policy variable $\ln(T)$ which is smaller by 0.78, holding all else equal.

Table 4.3 Estimation Results - the Composition of Taxes (annual data)

Dep Var: ln(T)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
*PROP*65	-0.161**	-0.0852	-0.165**	-0.116	-0.111*	-0.124***	-0.129**	-0.128
	(0.0623)	(0.108)	(0.0629)	(0.0955)	(0.0580)	(0.0406)	(0.0623)	(0.248)
*PROP*1564	-0.138***	-0.0152	-0.119**	-0.0115	-0.106**	-0.0774*	-0.0253	-0.100**
	(0.0442)	(0.0181)	(0.0507)	(0.0167)	(0.0496)	(0.0433)	(0.0289)	(0.0432)
UTIP	-0.0252	0.00239	-0.0232	0.000338	-0.0231	-0.00980	-0.0437*	0.0199
	(0.0198)	(0.0225)	(0.0206)	(0.0211)	(0.0169)	(0.0237)	(0.0251)	(0.0189)
ln(y)	1.647***	0.882***	1.628***	0.875***	0.892*	1.595***	1.171***	0.967***
	(0.455)	(0.308)	(0.465)	(0.299)	(0.447)	(0.296)	(0.412)	(0.328)
OECD	0.0458	-0.0559	0.0532	-0.0501		-0.0244	-0.000149	
	(0.102)	(0.127)	(0.0877)	(0.118)		(0.0655)	(0.0939)	
TRADE	0.00210	0.000153	0.00126	0.000869	0.00486	-0.000143	0.000949	0.00140
	(0.00265)	(0.00285)	(0.00259)	(0.00280)	(0.00449)	(0.00262)	(0.00219)	(0.00354)
ln(*POP*)	1.419	-0.991*	1.833*	-0.981*	0.715	3.173**	-0.955	0.747
	(1.122)	(0.531)	(1.081)	(0.515)	(1.340)	(1.352)	(1.234)	(0.811)
*POLITY*2	-0.0429	-0.0100	0.0149	-0.0106	0.397	-0.154**	-0.0249*	-0.00594
	(0.145)	(0.0219)	(0.156)	(0.0213)	(0.365)	(0.0716)	(0.0142)	(0.0131)
YGAP	-2.197	-0.133	-1.749	-0.146	-0.451	-1.664	-2.434*	-0.873
	(1.351)	(1.020)	(1.405)	(1.015)	(1.129)	(1.408)	(1.239)	(0.850)
Observations	520	276	496	300	273	223	559	237
Countries	61	37	56	41	38	32	67	31
Data	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel
Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	POLITY2 ≥ 7	POLITY2 < 7	POLITY2 ≥ 8	POLITY2 < 8	POLITY2 ≥ 8 & High τ_y	POLITY2 ≥ 8 & Low τ_y	High y	Low y
R^2	0.255	0.218	0.278	0.201	0.248	0.347	0.214	0.215

Notes: Regression specification is the same as column 3 of Table 4.2. Columns (1) and (2) respectively correspond to higher and lower democracy levels (according to $POLITY2 \gtrless 7$). Columns (3) and (4) instead respectively correspond to higher and lower democracy levels (according to $POLITY2 \gtrless 8$). Columns (5) and (6) respectively correspond to higher and lower levels of τ_y under the sample with higher democracy level ($POLITY2 \geq 8$). Columns (7) and (8) respectively correspond to higher and lower levels of income.

Note that it is also of interest to ask whether there are other stories, related to the Laffer curve, explaining the recent tendency of increasing expenditure taxes. The Laffer curve suggests that when income tax rates increase from low levels, the tax revenue collected by government also increases. If tax rates keep increasing after a certain point, then it would cause people not to work as hard as before, thereby reducing tax revenue. One common explanation for the appearance of expenditure taxes is that income taxes were close to their Laffer curve peaks. Columns 5 and 6 of Table 4.3 split the sample with stronger democratic credentials ($POLITY2 \geq 8$) by τ_y (determined by the median value of τ_y). If the story of the Laffer curve explains the results, when τ_y increases from lower to higher levels, then the tax revenue generated by taxes on income relative to expenditure firstly increases and then declines. This indicates that the sign of the coefficient on $PROP65$ should be reversed at lower and higher levels of τ_y, holding all else equal. Statistical significance in columns 5 and 6 implies that the estimates are stable across these subsamples, which in turn supports the theory proposed.

It is also natural to see whether the results vary with level of development. Columns 7 and 8 of Table 4.3 split the sample by levels of GDP per capita (determined by the median value of GDP per capita). As can be seen in all cases, the ratio of income to expenditure taxes is negatively correlated with the fraction of the retired population. However, this negative relationship holds significantly only in the group of countries with higher income level. Rich countries commonly have larger fractions of the retired population, and therefore tax revenue collected by taxes on income is reduced relative to expenditure. In column 7 the p-value for the estimated coefficient for the fraction of the retired population is 4.2%, and the estimated effect is sizable: a one standard deviation increase in the fraction of the retired population is statistically associated with a reduction of 0.61 in the policy variable $\ln(T)$.

Table 4.4 Estimation Results - the Composition of Taxes (annual data)

Dep Var: ln(T)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PROP65	-0.117*	-0.0968	-0.117*	-0.124	-0.123*	-0.278				
	(0.0665)	(0.113)	(0.0631)	(0.100)	(0.0622)	(0.229)				
RATIO							-0.0317**	-0.138**	-0.0701**	-0.0831
							(0.0132)	(0.0599)	(0.0343)	(0.214)
UTIP	-0.0290	0.000814	-0.0308	-0.000954	-0.0456*	0.0216	0.0254	-0.0227	-0.0487*	0.0259
	(0.0250)	(0.0225)	(0.0226)	(0.0208)	(0.0272)	(0.0215)	(0.0184)	(0.0227)	(0.0292)	(0.0234)
ln(y)	1.044***	0.826***	1.080***	0.834***	1.066***	0.679*	0.999***	0.860***	0.954***	0.560
	(0.310)	(0.285)	(0.289)	(0.276)	(0.390)	(0.350)	(0.192)	(0.239)	(0.355)	(0.367)
OECD	-0.127*	-0.0541	-0.0887	-0.0492	-0.0112		-0.117	0.0289	-0.0310	
	(0.0739)	(0.122)	(0.0700)	(0.114)	(0.0805)		(0.164)	(0.0949)	(0.0786)	
TRADE	0.00123	0.000656	0.000462	0.00121	0.00123	0.00194	-0.00386	0.00302	0.00111	0.00129
	(0.00271)	(0.00294)	(0.00264)	(0.00284)	(0.00230)	(0.00443)	(0.00260)	(0.00286)	(0.00224)	(0.00432)
ln(*POP*)	0.195	-1.121*	0.832	-1.083*	-1.166	-0.121	-0.882	-0.619	-1.124	-0.294
	(1.420)	(0.584)	(1.335)	(0.561)	(1.288)	(0.692)	(1.215)	(0.633)	(1.308)	(0.641)
POLITY2	-0.186	-0.0125	-0.0664	-0.0125	-0.0250*	-0.0125		-0.0125	-0.0247*	-0.0175
	(0.158)	(0.0203)	(0.168)	(0.0199)	(0.0139)	(0.0128)		(0.0124)	(0.0139)	(0.0117)
YGAP	-1.990	-0.149	-1.603	-0.149	-2.397*	-1.512	1.918	-1.027	-2.327*	-1.511
	(1.341)	(1.039)	(1.382)	(1.028)	(1.239)	(1.093)	(1.147)	(0.881)	(1.241)	(1.066)
Observations	520	276	496	300	559	237	281	515	559	237
Countries	61	37	56	41	67	31	31	62	67	31
Data	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel
Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	POLITY2 ≥ 7	POLITY2 < 7	POLITY2 ≥ 8	POLITY2 < 8	High y	Low y	POLITY2 $= 10$	POLITY2 < 10	High y	Low y
R^2	0.148	0.214	0.196	0.200	0.206	0.151	0.334	0.171	0.198	0.115

Notes: Columns (1)-(6) use Panel regressions with country fixed effects without *PROP1564* as a control. Columns (1) and (2) respectively correspond to higher and lower democracy levels (according to $POLITY2 \geqslant 7$). Columns (3) and (4) instead respectively correspond to higher and lower democracy levels (according to $POLITY2 \geqslant 8$). Columns (5) and (6) respectively correspond to higher and lower levels of income. Columns (7)-(10) instead use the ratio of the population above 65 to those between the ages of 15 and 64, *RATIO*, to measure population aging. Columns (7) and (8) respectively correspond to $POLITY2 = 10$ and $POLITY2 < 10$. Columns (9) and (10) respectively correspond to higher and lower levels of income. Robust standard errors are shown in parentheses. Standard errors are clustered by country. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

As a robustness check, columns 1-6 of Table 4.4 again use panel regressions with country fixed effects including full controls except $PROP1564$, with robust standard errors clustered by country. Columns 1 and 2 respectively correspond to higher and lower democracy levels (according to $POLITY2 \gtrless 7$), and columns 3 and 4 explore further stronger democratic requirement (i.e. $POLITY2 \gtrless 8$). Columns 5 and 6 instead split the sample of countries according to higher and lower levels of income. The estimation results support the findings in Table 4.3. Columns 7-10 then present regressions of the taxes composition on the ratio of old to young, $RATIO$. The results presented in columns 1-4 of Table 4.3 and columns 1-4 of Table 4.4 clearly establish that the estimated effect is predominantly driven by countries with high $POLITY2$ scoring scale. Note that these data are not classically normally-distributed, since there is a cluster of countries scoring 10. While many countries are with 'intermediate' $POLITY2$ scores, indicating a substantial political volatility (i.e. democratic reversals). This is likely to create further volatility in fiscal policy decision. Columns 7 and 8 therefore split the sample according whether or not $POLITY2 = 10$, showing a 'perfect' democracy throughout the period. Both of the relevant coefficient estimates are negative and statistically different from zero. Notably column 9 again confirms that the estimated effect holds significantly in rich countries.

In Tables 4.5 and 4.6 results are presented respectively for income taxes, τ_y, and expenditure taxes, τ_c, the numerator and denominator in the main dependent variable, using Panel regressions as above. In Table 4.5 the findings for income taxes, τ_y, are quite similar to the results found for $\ln(T)$ though with lower significance levels. Increases in the share of retirees are generally found to be negatively correlated with the extent to which taxes are levied on income, but more so in the stronger democracies. In countries where full sample is included or $POLITY2 \geq 7$, the estimated effect remains negative, though is not statistically significant. When the stronger democratic criterion (i.e. $POLITY2 \geq 8$) is employed, the estimated effect increases and is statistically significant at the 10% level. When the sample is refined further to those countries with $POLITY2 = 10$ throughout the same period (in columns 6 and 7), and utilizing instead the ratio of old to young ($RATIO$) measure of aging, the negative coefficient estimate is sustained, although statistical significance is in this instance low. If all control variables are excluded except ln(y), then the p-value of the coefficient estimate pertaining to $RATIO$ in column 6 improves to $p = 0.023$. Using the estimate of column 3, a one standard deviation increase in the fraction of the retired population is statistically associated with a reduction of 6.73 in τ_y, holding all else equal. Given that this is nearly about half of a standard deviation in the policy variable, the magnitude of the estimated correlation is sizable.

This weak estimated relationship indicates that there is a very slight variation in income taxes within countries over the sample period. This in turn emphasizes the need and motivation to examine the relationship between population aging and the extent to which taxes are levied on income relative to expenditure, instead of income taxes only.

In Table 4.5 the results relating to the control variables are of some interest. One regularity is that consistent with Perotti (1996), Benabou (1996), and results in chapter two, who have generally challenged the Meltzer and Richard (1981) hypothesis, there is a clear negative relationship between income taxes and income inequality (though at weak significance levels). This indicates that a more unequal distribution of income implies divergence between mean and median income and so, under universal suffrage, reduces income taxes. In addition as shown by Besley and Persson (2014) there is a positive relationship with income per capita, which likely indicates greater potential to tax in richer countries. Further, trade is found to be positively associated with income taxes as in Rodrik (1998), which shows a greater potential to tax in countries with higher level of openness.

Table 4.5 Estimation Results - Income Taxes (annual data)

Dep Var: τ_y	(1)	(2)	(3)	(4)	(5)	(6)	(7)
*PROP*65	-0.340	-1.275	-1.425*	-0.844	-0.892		
	(0.687)	(0.765)	(0.792)	(0.727)	(0.699)		
*PROP*1564	-0.307	-1.377**	-1.364**				
	(0.307)	(0.573)	(0.658)				
RATIO						-0.213	-0.429
						(0.173)	(0.738)
UTIP	-0.147	-0.345*	-0.312	-0.378	-0.394	0.412**	-0.271
	(0.247)	(0.201)	(0.243)	(0.249)	(0.266)	(0.197)	(0.284)
ln(*y*)	7.378	16.57***	16.63**	10.57**	10.41**	11.00***	5.903
	(4.489)	(6.128)	(6.548)	(4.100)	(4.071)	(3.498)	(4.655)
OECD	-2.855	-0.0816	-0.141	-1.768	-1.737	-1.508	-3.185
	(2.394)	(0.680)	(0.753)	(1.347)	(1.372)	(2.706)	(2.685)
TRADE	0.0410*	0.0653**	0.0693**	0.0549**	0.0589**	-0.0111	0.0671**
	(0.0211)	(0.0289)	(0.0297)	(0.0268)	(0.0274)	(0.0284)	(0.0323)
ln(*POP*)	-2.303	26.68**	30.57**	14.85	19.48	-18.29	-5.042
	(9.922)	(12.16)	(12.24)	(14.89)	(14.90)	(15.34)	(9.814)
*POLITY*2	-0.382	0.147	0.414	-1.271	-0.517		-0.370
	(0.238)	(1.594)	(1.750)	(1.693)	(1.996)		(0.255)
YGAP	-13.17	-23.05	-16.67	-20.39	-14.30	22.06*	-16.25
	(12.89)	(18.41)	(20.04)	(17.85)	(19.53)	(12.83)	(14.38)
Observations	826	539	515	539	515	295	531
Countries	89	62	57	62	57	32	63
Data	Panel	Panel	Panel	Panel	Panel	Panel	Panel
Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	POLITY2 $\geq$ 7	POLITY2 $\geq$ 8	POLITY2 $\geq$ 7	POLITY2 $\geq$ 8	POLITY2 = 10	POLITY2 $<$ 10
R^2	0.090	0.242	0.263	0.181	0.208	0.261	0.100

Notes: Estimations contain results using Panel regressions with country fixed effects of income taxes, τ_y, including *PROP*1564, *UTIP*, ln(*y*), *OECD*, *TRADE*, ln(*POP*), *POLITY*2, and *YGAP* as control variables. Column (1) includes full data sample; Columns (2) and (3) respectively correspond to different democracy levels ($POLITY2 \geq 7$ and $POLITY2 \geq 8$). Columns (4) and (5) again mimic columns (2) and (3) without *PROP*1564 as a control. Columns (6) and (7) instead use the ratio of the population above 65 to those between the ages of 15 and 64, *RATIO*, to measure population aging, respectively correspond to $POLITY2 = 10$ and $POLITY2 < 10$. Robust standard errors are shown in parentheses. Standard errors are clustered by country. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

Table 4.6 Estimation Results - Expenditure Taxes (annual data)

Dep Var: τ_c	(1)	(2)	(3)	(4)	(5)	(6)	(7)
*PROP*65	2.166***	2.318***	2.459***	1.861**	1.872**		
	(0.739)	(0.713)	(0.766)	(0.801)	(0.806)		
*PROP*1564	0.869**	1.447***	1.476**				
	(0.351)	(0.515)	(0.590)				
RATIO						0.485*	2.088**
						(0.251)	(0.905)
UTIP	0.228	0.428**	0.403*	0.467*	0.498**	0.132	0.294
	(0.222)	(0.195)	(0.216)	(0.239)	(0.222)	(0.473)	(0.270)
ln(*y*)	-11.01***	-17.31***	-18.66***	-10.97***	-11.88***	-12.06**	-6.643**
	(2.808)	(4.735)	(4.965)	(3.446)	(3.439)	(5.128)	(3.143)
OECD	-1.702	-0.740	-1.030	1.079	0.726	1.820*	-3.411
	(2.343)	(3.168)	(3.089)	(2.270)	(2.175)	(0.964)	(4.018)
TRADE	0.0152	0.0131	0.0251	0.0222	0.0350	0.0710*	-0.00922
	(0.0277)	(0.0300)	(0.0281)	(0.0320)	(0.0304)	(0.0359)	(0.0335)
ln(*POP*)	-1.007	-12.34	-14.34	0.505	-1.944	14.43	9.005
	(11.57)	(13.53)	(14.14)	(16.22)	(16.70)	(21.48)	(12.52)
*POLITY*2	-0.0797	0.488	0.661	1.992	1.667		-0.0686
	(0.184)	(1.560)	(1.708)	(1.567)	(1.862)		(0.165)
YGAP	7.993	26.43**	26.49**	24.27**	24.68**	-8.405	8.095
	(7.769)	(11.69)	(11.72)	(11.90)	(11.37)	(22.93)	(9.098)
Observations	797	521	496	521	496	281	516
Countries	87	61	56	61	56	31	62
Data	Panel	Panel	Panel	Panel	Panel	Panel	Panel
Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	POLITY2 $\geq$ 7	POLITY2 $\geq$ 8	POLITY2 $\geq$ 7	POLITY2 $\geq$ 8	POLITY2 = 10	POLITY2 $<$ 10
R^2	0.139	0.249	0.266	0.165	0.183	0.210	0.109

Notes: Estimations contain results using Panel regressions with country fixed effects of expenditure taxes, τ_c, including *PROP*1564, *UTIP*, ln(y), *OECD*, *TRADE*, ln(*POP*), *POLITY*2, and *YGAP* as control variables. Column (1) includes full data sample; Columns (2) and (3) respectively correspond to different democracy levels ($POLITY2 \geq 7$ and $POLITY2 \geq 8$). Columns (4) and (5) again mimic columns (2) and (3) without *PROP*1564 as a control. Columns (6) and (7) instead use the ratio of the population above 65 to those between the ages of 15 and 64, *RATIO*, to measure population aging, respectively correspond to $POLITY2 = 10$ and $POLITY2 < 10$. Robust standard errors are shown in parentheses. Standard errors are clustered by country. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

Table 4.6 contains estimation results relating to τ_c, the extent to which taxes are raised through expenditure on goods and services. In contrast to income taxes, increases in the share of retirees are generally found to be positively related to the extent to which expenditure taxes are used, and again this result is particularly strong in the stronger democracies. In countries where $POLITY2 < 7$, the estimated relationship is positive, though it is not statistically significant, whilst in countries where $POLITY2 \geq 7$, the estimated effect is statistically significant at the 1% level. When the stronger democratic requirement (i.e. $POLITY2 \geq 8$) is applied, the estimated effect remains positive and statistically significant at the 1% level, whilst in countries where $POLITY2 < 8$, the estimated relationship is found to be positive, though at a weaker significance level. The coefficient estimate is also positive and statistically significant when the sample is refined further to those pure democracies ($POLITY2 = 10$), and utilizing instead the $RATIO$ measure of aging in column 6. Using the estimate of column 5, a one standard deviation increase in the fraction of the retired population is statistically associated with a increase of 8.84 in τ_c, holding all else equal. As with $\ln(T)$, this again represents more than a half of the raw standard deviation in τ_c, so this is still a sizable effect.

There are some differences between the results relating to the controls for income taxes and expenditure taxes. For instance there is a clear positive relationship between expenditure taxes and income inequality, which is opposite to the findings in income taxes and implies that if the median voter becomes relatively poor, then he is likely to tax more on expenditure instead of income. Further in contrast to τ_y there is a negative relationship between τ_c and income per capita, which reflects the ability to collect revenue through taxes on income in particular.

4.3.3 Cross-Country Estimation

Table 4.7 presents OLS estimation results using cross-country averages to examine the effect of population aging on the ratio of income to expenditure taxes, with robust standard errors. This econometric analysis at least has the advantage of addressing potential cyclicality in the data. The estimated effect using cross-country regression still remains negative and is statistically significant at the 1% level when all controls except ln(y) are dropped or full controls are incorporated (in columns 1 and 2). Using the estimate in column 2, a one standard deviation increase in the proportion of the retired population is statistically associated with a reduction of 0.46 in the policy variable ln(T). In Tables 4.8 and 4.9 results are again presented respectively for τ_y and τ_c whilst using cross-country estimation. As in Razin et al. (2002) population aging leads to smaller income taxes, and the estimated effects are generally statistically significant and are moreso in the stronger democracies. On the other hand, population aging is found to be positively related with expenditure taxes but at weaker significance levels, while this relationship holds in regimes with higher democratic scores (i.e. $POLITY2 \geq 9$).

In the case of cross-country estimation in income and expenditure taxes, some results relating to the control variables are interesting. Notably the extent of democracy is positively associated with both τ_y and τ_c. This means that revenue relied on τ_y and τ_c is increasingly related with the stronger level of democracy. Further in line with Baunsgaard and Keen (2010), trade is negatively related to τ_c (though at weak significance level), which implies that globalization might constrain the capacity to raise revenue through taxes on goods and services especially more pressure in countries without alternative sources to collect revenue.

Table 4.7 Estimation Results - the Composition of Taxes (cross-country data)

Dep Var: ln(T)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PROP65	-0.112***	-0.0976***	-0.0763**	-0.0970**	-0.0819**	-0.0933**		
	(0.0276)	(0.0273)	(0.0299)	(0.0365)	(0.0314)	(0.0357)		
PROP1564		-0.0589**	-0.0375	0.0229				
		(0.0229)	(0.0426)	(0.0467)				
RATIO							-0.112***	-0.0897***
							(0.0216)	(0.0231)
UTIP		0.0327*	0.0486*	0.0736***	0.0573**	0.0695***	0.0686***	0.0387*
		(0.0180)	(0.0253)	(0.0213)	(0.0222)	(0.0200)	(0.0205)	(0.0227)
ln(y)	0.422***	0.783***	0.909***	1.029***	0.863***	1.051***	1.662***	0.478***
	(0.120)	(0.163)	(0.295)	(0.293)	(0.304)	(0.299)	(0.314)	(0.143)
OECD		0.505	0.0401	0.187	0.111	0.141	0.0468	0.302
		(0.341)	(0.441)	(0.383)	(0.448)	(0.392)	(0.345)	(0.497)
TRADE		0.000658	0.0000832	0.000832	-0.000364	0.00107	-0.00361*	0.00150
		(0.00156)	(0.00208)	(0.00215)	(0.00225)	(0.00218)	(0.00208)	(0.00216)
ln(*POP*)		0.0969	0.161*	0.228**	0.147	0.233**	0.256**	0.0356
		(0.0612)	(0.0947)	(0.0982)	(0.0994)	(0.0991)	(0.101)	(0.0679)
POLITY2		-0.0332*	0.117	0.103	0.115	0.120		-0.0316
		(0.0200)	(0.120)	(0.179)	(0.123)	(0.169)		(0.0204)
Observations	146	111	56	45	56	45	27	84
Countries	146	111	56	45	56	45	27	84
Data	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country
Sample	Full	Full	POLITY2 ≥ 7	POLITY2 ≥ 8	POLITY2 ≥ 7	POLITY2 ≥ 8	POLITY2 $= 10$	POLITY2 < 10
R^2	0.150	0.436	0.550	0.631	0.540	0.628	0.869	0.397

Notes: Estimations contain results using cross-country OLS regressions of the composition of taxes, ln(T), including *PROP*1564, *UTIP*, ln(y), *OECD*, *TRADE*, ln(*POP*), and *POLITY*2 as control variables, with robust standard errors in parentheses. Column (1) includes full data sample, only with ln(y) as a control; Column (2) includes full data sample, with full control variables. Columns (3) and (4) respectively correspond to different democracy levels ($POLITY2 \geq 7$ and $POLITY2 \geq 8$). Columns (5) and (6) again mimic columns (3) and (4) without *PROP*1564 as a control. Columns (7) and (8) instead use the ratio of the population above 65 to those between the ages of 15 and 64, *RATIO*, to measure population aging, respectively correspond to $POLITY2 = 10$ and $POLITY2 < 10$. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

Table 4.8 Estimation Results - Income Taxes (cross-country data)

Dep Var: τ_y	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
*PROP*65	-0.585	-1.372***	-1.369**	-2.090***	-1.345**	-2.019***		
	(0.365)	(0.424)	(0.603)	(0.753)	(0.584)	(0.739)		
*PROP*1564		-0.407	0.160	0.432				
		(0.326)	(0.578)	(0.835)				
RATIO							-2.046***	-0.930***
							(0.528)	(0.311)
UTIP		0.194	0.600	1.022**	0.562	0.944**	1.029*	0.291
		(0.309)	(0.498)	(0.437)	(0.445)	(0.405)	(0.553)	(0.335)
ln(y)	4.349***	6.620***	8.101**	15.09***	8.297**	15.49***	24.19***	3.790**
	(1.417)	(1.846)	(3.183)	(3.905)	(3.352)	(4.132)	(5.176)	(1.832)
OECD		7.656	5.691	7.643	5.388	6.781	8.585	0.496
		(4.713)	(6.844)	(6.798)	(6.615)	(6.940)	(6.872)	(5.879)
TRADE		0.0155	-0.0362	-0.0525	-0.0342	-0.0481	-0.118**	0.0396
		(0.0248)	(0.0516)	(0.0530)	(0.0501)	(0.0521)	(0.0573)	(0.0296)
ln(*POP*)		1.870**	1.108	1.393	1.165	1.498	0.644	1.763**
		(0.823)	(1.517)	(1.552)	(1.462)	(1.577)	(1.954)	(0.828)
*POLITY*2		0.485*	3.319*	0.653	3.330*	0.973		0.523*
		(0.274)	(1.734)	(3.498)	(1.707)	(3.337)		(0.281)
Observations	147	112	56	45	56	45	27	85
Countries	147	112	56	45	56	45	27	85
Data	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country
Sample	Full	Full	POLITY2 ≥ 7	POLITY2 ≥ 8	POLITY2 ≥ 7	POLITY2 ≥ 8	POLITY2 $= 10$	POLITY2 < 10
R^2	0.110	0.307	0.426	0.509	0.425	0.505	0.707	0.239

Notes: Estimations contain results using cross-country OLS regressions of income taxes, τ_y, including $PROP1564$, $UTIP$, ln(y), $OECD$, $TRADE$, ln(POP), and $POLITY2$ as control variables, with robust standard errors in parentheses. Column (1) includes full data sample, only with ln(y) as a control; Column (2) includes full data sample, with full control variables. Columns (3) and (4) respectively correspond to different democracy levels ($POLITY2 \geq 7$ and $POLITY2 \geq 8$). Columns (5) and (6) again mimic columns (3) and (4) without $PROP1564$ as a control. Columns (7) and (8) instead use the ratio of the population above 65 to those between the ages of 15 and 64, $RATIO$, to measure population aging, respectively correspond to $POLITY2 = 10$ and $POLITY2 < 10$. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

Table 4.9 Estimation Results - Expenditure Taxes (cross-country data)

Dep Var: τ_c	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
*PROP*65	0.953**	0.295	0.307	0.868*	0.309	0.860*		
	(0.369)	(0.420)	(0.394)	(0.503)	(0.360)	(0.485)		
*PROP*1564		0.793**	0.0154	-0.438				
		(0.369)	(0.522)	(0.615)				
RATIO							0.543*	0.614*
							(0.314)	(0.353)
UTIP		-0.322	-0.791**	-0.913***	-0.793**	-0.839**	-0.703*	-0.281
		(0.305)	(0.368)	(0.283)	(0.344)	(0.314)	(0.363)	(0.372)
ln(y)	-2.566	-6.691***	-7.999**	-9.850***	-7.985**	-9.743***	-14.83***	-3.142*
	(1.653)	(2.139)	(3.375)	(3.076)	(3.339)	(3.266)	(3.894)	(1.849)
OECD		-2.458	-2.756	-3.487	-2.787	-3.364	1.783	-0.763
		(4.306)	(3.849)	(4.175)	(3.679)	(4.409)	(4.634)	(7.102)
TRADE		-0.0281	-0.0396	-0.00373	-0.0394	-0.00992	0.00990	-0.0328
		(0.0227)	(0.0338)	(0.0201)	(0.0314)	(0.0198)	(0.0197)	(0.0294)
ln(*POP*)		-1.044	-2.525**	-2.773**	-2.521**	-2.942**	-3.371***	-0.0294
		(0.847)	(1.000)	(1.335)	(0.954)	(1.305)	(1.100)	(1.058)
*POLITY*2		0.507*	0.162	3.354	0.173	2.672		0.454
		(0.294)	(1.852)	(3.612)	(1.775)	(3.486)		(0.298)
Observations	147	111	45	35	45	35	27	84
Countries	147	111	45	35	45	35	27	84
Data	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country	Cross-country
Sample	Full	Full	POLITY2 ≥ 8	POLITY2 ≥ 9	POLITY2 ≥ 8	POLITY2 ≥ 9	POLITY2 $= 10$	POLITY2 < 10
R^2	0.056	0.221	0.463	0.619	0.463	0.610	0.733	0.186

Notes: Estimations contain results using cross-country OLS regressions of expenditure taxes, τ_c, including *PROP*1564, *UTIP*, ln(y), *OECD*, *TRADE*, ln(*POP*), and *POLITY*2 as control variables, with robust standard errors in parentheses. Column (1) includes full data sample, only with ln(y) as a control; Column (2) includes full data sample, with full control variables. Columns (3) and (4) respectively correspond to different democracy levels ($POLITY2 \geq 8$ and $POLITY2 \geq 9$). Columns (5) and (6) again mimic columns (3) and (4) without *PROP*1564 as a control. Columns (7) and (8) instead use the ratio of the population above 65 to those between the ages of 15 and 64, *RATIO*, to measure population aging, respectively correspond to $POLITY2 = 10$ and $POLITY2 < 10$. *, **, and *** respectively denote significant levels at 10%, 5% and 1%.

4.4 Conclusion

This chapter analyzes how population aging affects the composition of taxes. In an overlapping generations model taxes are levied on both labor income and expenditure, financing redistribution to both generations.

An increased share of retirees in the population leads to lower income taxes, but it leads to higher expenditure taxes at initial high levels of expenditure taxes, whilst leads to lower expenditure taxes otherwise. The results relating to the composition of taxes, defined as the extent of taxes on income relative to expenditure, are novel. Increases in the fraction of the retired population in the model lead to increases in expenditure taxes relative to income taxes, because the working-age population (including the median voter) wants to shift the tax-burden onto the retired population.

The relationship between population aging and the composition of taxes is tested using international panel data, including the fraction of the population that is retired as an explanatory variable. Data for taxes composition and demography are all from the WDI database. Consistent with the theory, the extent of taxes on income relative to expenditure is found to be negatively associated with the fraction of the retired population.[9] Moreover, income taxes as a proportion of total revenue fall with aging in support of the Razin et al. (2002) hypothesis, whilst expenditure taxes as a proportion of total revenue increase with aging. The empirical results hold across various econometric specifications employed. In particular the fact that the results found hold significantly in countries with strong democratic credentials is supportive of the mechanism proposed in this chapter.

[9] This is in line with intuition: when the median voter is of working age, then population aging increases the demand for expenditure rather than income taxes in order to increase the tax burden on the retired population.

Chapter 5

Conclusion

5.1 Summary

This thesis contributes to three important topics on income inequality and fiscal policy. Chapter 2 derives a median voter model with the twist that income inequality is engendered from differences in capital income as well as differences in labor productivity. The chapter analyzes how inequality in the distribution of capital income affects the size of government. Capital income is quite distinct from labor income. I define it as rental income, and also model it as difficult to tax. Therefore, redistribution is financed only by taxation applied to labor income, and voters have preferences over the tax rate based on their position in the distribution of capital income. Despite that there are two underlying sources of heterogeneity in the populations, the median voter is still the unique Condorcet winner as tax preferences are monotonic in labor income.

The result in Chapter 2 relating taxation levels to capital income inequality is novel. In contrast to Meltzer and Richard (1981), increased capital income inequality leads to smaller government. Individuals who are endowed with capital income are less averse to labor-income taxation. If the share of capital income of the rich rises, then their taxable labor supply declines and the preferred labor tax rate declines as the median voter has a reduced capacity to redistribute through taxation. The relationship between the size of government and inequality is tested in a panel of OECD data, including capital income inequality as an additional explanatory variable. The measure of capital income inequality in the analysis is the top 1 percent income share. Consistent with the theory proposed, government size is found to be negatively associated with capital income inequality. Moreover, controlling for the top income share renders a consistently positive estimate for the impact of labor income inequality on government size, in line with the original Meltzer and Richard (1981) hypothesis. The negative impact of capital income inequality on government size also survives when capital income inequality is instrumented with variables encapsulating technology and access to the capital market.

The model constructed in Chapter 2 is extended in Chapter 3 to investigate how economic growth is affected by inequality in the capital income distribution in an endogenous growth model. The template in this chapter is Persson and Tabellini (1994), who argue that productivity-induced income inequality leads to lower growth as distortionary taxes increase and harm capital accumulation. However, the chapter theorizes that if inequality stems from differences in income from capital, then labor tax rate falls, implying reduced tax distortions, leading to higher growth. The relationship between inequality and growth is tested in a panel of OECD data, augmenting the work of Forbes (2000) to include capital income inequality as an additional explanatory variable. Consistent with the theory put forward in the chapter, growth is found to be positively associated with capital income inequality. Controlling for capital income inequality yields a negative relationship between labor income inequality and growth, as Persson and Tabellini (1994) originally conjectured, and in contrast to previous empirical work (in particular Forbes, 2000) challenging their hypothesis. The positive impact of capital income inequality on growth also holds in various econometric specifications, including when difference and system GMM techniques are used to deal with the potential endogeneity problem.

The analysis (distinction between capital income inequality and labor income inequality) presented in Chapter 2 and Chapter 3 can be applied in most inequality literature and it can be applied in international data as only OECD countries are currently taken into account.

In particular, it is very interesting to apply this analysis in the case of China. This is due to the fact that there are a great number of unobserved capital income in China, and this will conceivably lead to high capital income inequality level. Current observed inequality level may be underestimated. If capital income is in fact highly concentrated within the top (capital) rich in China, then its consequence will be somewhat attractive.

Chapter 4, building on Razin et al. (2002), instead investigates the effect of demography on the composition of taxes. In an overlapping generations model, taxes are levied on both labor income and expenditure, financing redistribution to both generations (working and retired population). The results in this chapter relating to the composition of taxes, defined as the extent of taxes on income relative to expenditure, are novel. When the median voter is of working age, then population aging increases the demand for expenditure rather than income taxes as the working-age population wants to shift the tax-burden onto the retired population. International panel data in the empirical section of the chapter confirm this prediction. Data for taxes composition and the fraction of the retired population (as an explanatory variable) are all from the WDI database. Consistent with the theory derived, the extent of taxes on income relative to expenditure is found to be negatively associated with the fraction of the retired population. Moreover, income taxes as a share of total revenue decline with aging in support of the Razin et al. (2002) hypothesis, whilst expenditure taxes as a share of total revenue rise with aging. In particular, the results found hold significantly in countries with strong democratic credentials is supportive of the mechanism proposed in this chapter.

The analysis of Chapter 4 can be applied in the design of tax system as well as the income taxation literature. The composition of taxes is always neglected by the government and researchers. Tax system is set ideally by economists, but politician cannot follow this ideal tax design. This is due to the risk of losing votes in politician competition. Government did raise revenue through income taxes 30 years ago. While it might probably have reached the peak of Laffer curve. This indicates that we need to look elsewhere, such as expenditure taxes.

5.2 Further Reflections

However, there still exists debate such as the causes and consequences of income inequality induced by differences in capital income and, although the results from this thesis would enrich the inequality literature, more research into these results are needed. The distinction between labor and capital income inequality is first proposed by this thesis, and allows further research to answer questions relevant internationally.

First, further research, utilizing the distinction between labor and capital income inequality, can investigate how inequality stemming from capital income affects growth as well as the size of government within-country. As mentioned in the introduction of this thesis, income concentration is high and growing in the US. Using US states data, both capital income inequality and labor income inequality can be constructed by using micro individual data of income from the Panel Study of Income Dynamics (PSID) dataset, and I will be able to see whether an increase in capital income inequality has a significant positive relationship with subsequent economic growth in the US. Moreover, I can also test how the sources of inequality affect the size of US state government.

A second question which I can address, employing this distinction, is to analyze the case of China. As mentioned in the introduction, inequality levels in China are currently close to US levels. It is usually not possible to access the city-level data in China which are public-access constrained. However, I have already obtained the internal data set, covering 223 cities in Chia over 2001-2010. Augmenting capital income inequality and labor income inequality constructed by micro individual data of income from the China Household Finance Survey (CHFS), I will be able to see how income inequality affects growth as well as the size of city government in China. Moreover, leaders affect policy outcomes and growth in autocratic settings as less constraints on a leader's power (Jones and Olken, 2005), the econometric analysis thus potentially needs to include the characteristics of mayors in order to analyze the relationship between inequality and growth in China.

The results in Chapter 4 also shed new light on the relationship between population aging and economic growth. Several theories document the negative effect of population aging on economic growth, either due to smaller size of labor force and productivity (see Gordon, 2016) or due to excess of savings over desired investment (see also Teulings and Baldwin, 2014). However, there is no such negative relationship in evidence. Whilst countries experiencing more rapid aging have grown more in recent years (Acemoglu and Restrepo, 2017). Hence, another further research can theorize that if income taxes are distortionary whilst expenditure taxes are non-distortionary as in Kneller et al. (1999), then population aging leads to reduced demand for distortionary rather than non-distortionary taxes, and therefore promoting growth.

Appendix A

Appendices to Chapter 2

A.1 Derivation of Equations (2.12) and (2.13)

The problem of the median voter m is to choose the tax rate so as to maximize

$$u^m(c^m, l^m) = u^m\Big[(1-t)x^m n^m + R^m + t\bar{y}, 1 - n^m\Big], \tag{A.1}$$

and the first-order condition for the median voter with respect to the tax rate is

$$\left(\bar{y} - y^m + t\frac{d\bar{y}}{dt}\right)u_c + \Big[(1-t)x^m u_c - u_l\Big]\left(\frac{dn^m}{dt}\right) = 0. \tag{A.2}$$

Thus, making use of equation (2.3), the tax rate chosen by the median voter must satisfy

$$\bar{y} - y^m + t\left(\frac{d\bar{y}}{dt}\right) = 0. \tag{A.3}$$

Changes in the tax rate t affect average income via two channels: its effect on the opportunity cost of leisure, and its effect on transfers (from the government's budget constraint $r = t\bar{y}$). In particular, we have that

$$\begin{aligned} \frac{d\bar{y}}{dt} &= \frac{\partial \bar{y}}{\partial r}\frac{dr}{dt} - \frac{\partial \bar{y}}{\partial \tau}, \\ &= \frac{\partial \bar{y}}{\partial r}\left(\bar{y} + t\frac{d\bar{y}}{dt}\right) - \frac{\partial \bar{y}}{\partial \tau}. \end{aligned} \tag{A.4}$$

with $\tau = 1 - t$. Thus, the total derivative of average income with respect to changes in the tax rate is given by

$$\frac{d\bar{y}}{dt} = \frac{\bar{y}_r \bar{y} - \bar{y}_\tau}{1 - t\bar{y}_r} < 0, \tag{A.5}$$

with $\bar{y}_r = \frac{\partial \bar{y}}{\partial r}$ and $\bar{y}_\tau = \frac{\partial \bar{y}}{\partial \tau}$.

Finally, making use of (A.5) to substitute in (A.3), we obtain

$$\begin{aligned} 0 &= \bar{y} - y^m + t\left(\frac{\bar{y}_r \bar{y} - \bar{y}_\tau}{1 - t\bar{y}_r}\right), \\ &= (\bar{y} - y^m)(1-t) + \left[\frac{\eta_r \bar{y}(1-t) - \eta_\tau \bar{y} t}{1 - \eta_r}\right], \end{aligned} \tag{A.6}$$

where $\eta_r = \bar{y}_r (r/\bar{y})$ and $\eta_\tau = \bar{y}_\tau (\tau/\bar{y})$ are the partial elasticities of average income. Solving the above equation for t, yields

$$t = \frac{m - 1 + \eta_r}{m - 1 + \eta_r + m\eta_\tau}, \tag{A.7}$$

with $m = \bar{y}/y^m$.

A.2 Proof of Proposition 2.1

We begin with the following decomposition of average income

$$\bar{y} = p(\mathscr{K})\bar{y}(\mathscr{K}) + (1 - p(\mathscr{K}))\bar{y}(\sim \mathscr{K}), \tag{A.8}$$

where $\bar{y}(\mathscr{K})$ is the average income of the individuals in set $\mathscr{K}$ and $\bar{y}(\sim \mathscr{K})$ is the average income of the individuals not in set $\mathscr{K}$. From Assumption 2.2 we have that $\bar{y}^{\mathscr{K}} > y^m$.

Taking the total derivative of $\bar{y}$ with respect to $R(\mathscr{K})$, the capital income of the individuals in set $\mathscr{K}$ in equation (A.8) we obtain

$$\begin{aligned}\frac{d\bar{y}}{dR(\mathscr{K})} &= p(\mathscr{K})\left(\frac{\partial\bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} + \frac{\partial\bar{y}(\mathscr{K})}{\partial r}\frac{d\bar{y}}{dR(\mathscr{K})}t\right) + (1-p(\mathscr{K}))\left(\frac{\partial\bar{y}(\sim\mathscr{K})}{\partial r}\frac{d\bar{y}}{dR(\mathscr{K})}t\right),\\ &= p(\mathscr{K})\frac{\partial\bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} + \frac{\partial\bar{y}}{\partial r}\frac{d\bar{y}}{dR(\mathscr{K})}t,\\ &= p(\mathscr{K})\frac{\partial\bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} + \eta_r\frac{d\bar{y}}{dR(\mathscr{K})},\end{aligned} \tag{A.9}$$

where we used the fact that $\eta_r = \frac{\partial\bar{y}}{\partial r}\frac{r}{\bar{y}} = \frac{\partial\bar{y}}{\partial r}\frac{t\bar{y}}{\bar{y}} = \frac{\partial\bar{y}}{\partial r}t$. Using (A.9) to solve for $\frac{d\bar{y}}{dR(\mathscr{K})}$, we obtain

$$\frac{d\bar{y}}{dR(\mathscr{K})} = \frac{p(\mathscr{K})}{1-\eta_r}\frac{\partial\bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} < 0, \tag{A.10}$$

since leisure is a normal good. Thus, average income $\bar{y}$ must fall.

In turn, we have that

$$\frac{dy^m}{dR(\mathscr{K})} = \frac{\partial y^m}{\partial r}\frac{\partial\bar{y}}{\partial R(\mathscr{K})}t > 0. \tag{A.11}$$

Thus, we have established that $\bar{y}$ must fall and y^m must increase following an increase in the capital-income going to the top capital-income recipients. Therefore, $m = \bar{y}/y^m$ falls and the increase in capital income inequality lowers labor income inequality. The upshot is that the increase in the capital income going to the top capital-income recipients results in a lower t, the labor income tax chosen by the median voter.

Appendix B

Appendices to Chapter 3

B.1 Derivation of Equations (3.13) and (3.14)

The problem of the median voter m is to choose the tax rate so as to maximize

$$v_t^m = U\Big[\frac{\gamma}{\gamma+D}\Big((1-\tau_t)n^m e^m k_t + \tau_t \bar{y}_t + R^m k_t\Big), 1-n^m, \\ \frac{\gamma D}{\gamma+D}\Big((1-\tau_t)n^m e^m k_t + \tau_t \bar{y}_t + R^m k_t\Big)\Big], \tag{B.1}$$

and the first-order condition for the median voter with respect to the tax rate is

$$\Big(\bar{y}_t - y_t^m + \tau_t \frac{d\bar{y}_t}{d\tau_t}\Big)\Big(\frac{\gamma}{\gamma+D}U_c + \frac{\gamma D}{\gamma+D}U_d\Big) \\ + \Big[\frac{\gamma}{\gamma+D}(1-\tau_t)e^m k_t U_c - U_l + \frac{\gamma D}{\gamma+D}(1-\tau_t)e^m k_t U_d\Big]\frac{dn^m}{d\tau_t} = 0. \tag{B.2}$$

Thus, making use of equation (3.4), the tax rate chosen by the median voter must satisfy

$$\bar{y}_t - y_t^m + \tau_t \frac{d\bar{y}_t}{d\tau_t} = 0. \tag{B.3}$$

For a given labor income inequality, the political equilibrium τ is constant over time, so that the time subscript t is suppressed henceforth. Changes in the tax rate τ affect average income via two channels: its effect on the opportunity cost of leisure, and its effect on transfers (from the government's budget constraint, $r = \tau\bar{y} + \vartheta R^i k$, and fixed capital income taxation, $\vartheta R^i k$). In particular, I have that

$$\begin{aligned}\frac{d\bar{y}}{d\tau} &= \frac{\partial\bar{y}}{\partial r}\frac{dr}{d\tau} + \frac{\partial\bar{y}}{\partial\theta}\frac{d\theta}{d\tau}, \\ &= \frac{\partial\bar{y}}{\partial r}\left(\bar{y} + \tau\frac{d\bar{y}}{d\tau}\right) - \frac{\partial\bar{y}}{\partial\theta}\end{aligned} \tag{B.4}$$

with $\theta = 1 - \tau$. Thus, the total derivative of average labor income with respect to changes in the tax rate is given by

$$\frac{d\bar{y}}{d\tau} = \frac{\bar{y}_r\bar{y} - \bar{y}_\theta}{1 - \tau\bar{y}_r} < 0, \tag{B.5}$$

with $\bar{y}_r = \frac{\partial\bar{y}}{\partial r}$ and $\bar{y}_\theta = \frac{\partial\bar{y}}{\partial\theta}$. Finally, substituting (B.5) into (B.3) I have

$$\begin{aligned}0 &= \bar{y} - y^m + \tau\frac{\bar{y}_r\bar{y} - \bar{y}_\theta}{1 - \tau\bar{y}_r}, \\ &= (\bar{y} - y^m)(1 - \tau) + \frac{\eta_r\bar{y}(1 - \tau) - \eta_\theta\bar{y}\tau}{1 - \eta_r},\end{aligned} \tag{B.6}$$

where $\eta_r = \bar{y}_r\frac{r}{\bar{y}}$ and $\eta_\theta = \bar{y}_\theta\frac{\theta}{\bar{y}}$ are the partial elasticities of average income. Solving the above equation for τ, yields

$$\tau = \frac{m - 1 + \eta_r}{m - 1 + \eta_r + m\eta_\theta} \tag{B.7}$$

with $m = \frac{\bar{y}}{y^m}$.

B.2 Proof of Lemma 3.1

Although I impose almost no restrictions on the joint distribution $f(e^i, R^i)$, as in chapter two I wish to guarantee that: i) the chosen tax rate is positive; and that ii) the individuals that are in the top of the capital income distribution are never the decisive voter. Thus, in the sequel I make the following two assumptions:

Assumption B.1. *The joint distribution $f(e^i, R^i)$ is such that the labor income distribution is right-skewed. Thus, $y^m < \bar{y}$ and the chosen tax rate is positive.*

From (3.13) I see that Assumption B.1 guarantees that the chosen tax rate is positive.

Assumption B.2. *The joint distribution $f(e^i, R^i)$ is such that the set of individuals $i \in \mathscr{K}$ with capital income R^i above the 99% percentile of the capital income distribution has productivity e^i which is sufficiently high so that $y^i = e^i n^i k > y^m$ for all $i \in \mathscr{K}$.*

Figure 2.2 illustrates the condition imposed by Assumption B.2. The locus denoted $y = y^m$ represents productivity and capital income pairs, $\left(e^i, R^i\right)$, for which labor income y is equal to the median voter's labor income, y^m. To the right of this locus, $y > y^m$, since $\frac{\partial y^i}{\partial e^i} > 0$ and $\frac{\partial y^i}{\partial R^i} < 0$. The dashed line denoted $\mathscr{Q}_{99\%}$ represents the 99% quantile of the capital income marginal density function. Assumption B.2 is a condition requiring that the set $\mathscr{K}$ of all individuals with capital income above $\mathscr{Q}_{99\%}$ is located to the right of the locus $y = y^m$, as shown in Figure 2.2.

Now consider an increase in the capital income earned by the individuals in the set $\mathscr{K}$ of all individuals with capital income above $\mathscr{Q}_{99\%}$. This is represented in Figure 2.3: the individuals in the set $\mathscr{K}$ that correspond to the original individuals in the top 1% of the capital income distribution receive an exogenous increase in capital income; thus, the set $\mathscr{K}$ shifts upwards in the space $\left(e^i, R^i\right)$, but still satisfying the restriction imposed by Assumption B.2, that guarantees that none of the members of the set $\mathscr{K}$ are the median voter (the new set is represented by the triangle above, in Figure 2.3). Notice that this experiment constitutes an increase in capital income inequality, since I maintain the capital income of all the other individuals unchanged and, hence, the capital income share of the top 1% is increased.[1] Under a right-skewed labor income distribution $y^m < \bar{y}$, and given (3.14) above then $\tau > 0$. As with Meltzer and Richard (1981) demand for redistribution stems from changes in the labor income distribution. However, the labor income distribution may now change depending on the distribution of capital income as well as the productivity distribution.

To see the consequences of higher capital income inequality, notice that all the individuals in the set $\mathscr{K}$ will choose to work less, because they enjoy an increase in their capital income and leisure is a normal good. This will tend to lower the average labor income $\bar{y}$, since I have

[1] It is not, however, a mean preserving spread in capital income. But lowering the capital income of the bottom 99% capital income earners in order to preserve the mean capital income would only reinforce our results.

that

$$\bar{y} = p(\mathscr{K})\,\bar{y}(\mathscr{K}) + (1 - p(\mathscr{K}))\,\bar{y}(\sim \mathscr{K}), \tag{B.8}$$

where $\bar{y}(\mathscr{K})$ is the average income of the individuals in set $\mathscr{K}$ and $\bar{y}(\sim \mathscr{K})$ is the average income of the individuals not in set $\mathscr{K}$. From Assumption B.2 I have that $\bar{y}^{\mathscr{K}} > y^m$.

Taking the total derivative of $\bar{y}$ with respect to $R(\mathscr{K})$, the capital income of the individuals in set $\mathscr{K}$ in equation (B.8) I obtain

$$\begin{aligned}\frac{d\bar{y}}{dR(\mathscr{K})} &= p(\mathscr{K})\left(\frac{\partial \bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} + \frac{\partial \bar{y}(\mathscr{K})}{\partial r}\frac{d\bar{y}}{dR(\mathscr{K})}\tau\right) + (1 - p(\mathscr{K}))\left(\frac{\partial \bar{y}(\sim \mathscr{K})}{\partial r}\frac{d\bar{y}}{dR(\mathscr{K})}\tau\right), \\ &= p(\mathscr{K})\frac{\partial \bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} + \frac{\partial \bar{y}}{\partial r}\frac{d\bar{y}}{dR(\mathscr{K})}\tau, \\ &= p(\mathscr{K})\frac{\partial \bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} + \eta_r \frac{d\bar{y}}{dR(\mathscr{K})},\end{aligned} \tag{B.9}$$

where I used the fact that $\eta_r = \frac{\partial \bar{y}}{\partial r}\frac{r}{\bar{y}} = \frac{\partial \bar{y}}{\partial r}\frac{\tau \bar{y}}{\bar{y}} = \frac{\partial \bar{y}}{\partial r}\tau$. Using (B.9) to solve for $\frac{d\bar{y}}{dR(\mathscr{K})}$, I obtain

$$\frac{d\bar{y}}{dR(\mathscr{K})} = \frac{p(\mathscr{K})}{1 - \eta_r}\frac{\partial \bar{y}(\mathscr{K})}{\partial R(\mathscr{K})} < 0, \tag{B.10}$$

since leisure is a normal good. Thus, average income $\bar{y}$ must fall.

In turn, I have that

$$\frac{dy^m}{dR(\mathscr{K})} = \frac{\partial y^m}{\partial r}\frac{\partial \bar{y}}{\partial R(\mathscr{K})}\tau > 0. \tag{B.11}$$

Thus, I have established that $\bar{y}$ must fall and y^m must increase following an increase in the capital-income going to the top capital-income recipients. Therefore, $m = \bar{y}/y^m$ falls and the increase in capital income inequality lowers labor income inequality. The upshot is that the increase in the capital income going to the top capital-income recipients results in a lower τ, the labor income tax chosen by the median voter.

B.3 Derivation of Equation (3.16)

For the average individual in (3.1) and (3.2), I have

$$\begin{aligned} k_{t+1} &= y_t + Rk_t - c_t, \\ &= y_t + Rk_t - \frac{d_{t+1}}{D}, \\ &= y_t + Rk_t - \frac{\gamma k_{t+1}}{D}. \end{aligned} \tag{B.12}$$

Solving the above equation for k_{t+1}, yields

$$k_{t+1} = \frac{D(y_t + Rk_t)}{\gamma + D}. \tag{B.13}$$

Combining the above equation and (3.8), the growth rate of k can be obtained

$$\begin{aligned} g_t &= \frac{k_{t+1} - k_t}{k_t}, \\ &= \frac{D\left(\int_0^\infty \int_0^\infty e^i n[(1-\tau_t)e^i, r_t, R^i] f(e^i, R^i) de^i dR^i + R\right)}{\gamma + D} - 1. \end{aligned} \tag{B.14}$$

Again for a given labor income inequality, the political equilibrium τ and g are constant over time, so that the time subscript t is suppressed henceforth. Thus, the effect of taxation on growth, making use of (B.5), yields

$$\begin{aligned} \frac{dg}{d\tau} &= \frac{D}{\gamma + D} \frac{d\left(\int_0^\infty \int_0^\infty e^i n[(1-\tau)e^i, r, R^i] f(e^i, R^i) de^i dR^i + R\right)}{d\tau}, \\ &= \frac{D}{\gamma + D} \frac{\frac{1}{k} dy}{d\tau} < 0. \end{aligned} \tag{B.15}$$

B.4 Labor Income Inequality and Growth (without R_i)

B.4.1 Economic Environment

Different individuals have different incomes. The budget constraints, common to all individuals, are

$$c_t^i + k_{t+1}^i = (1-\tau_t)y_t^i + r_t \tag{B.16}$$

$$d_{t+1}^i = \gamma k_{t+1}^i \tag{B.17}$$

where y^i is the individual's labor income when young, and is taxed at a linear rate τ, k^i is the individual accumulation of asset, r is lump-sum redistribution, and γ is the exogenous rate of return on asset. Individuals make decision between consumption and investment when young, financed by disposable labor income and lump-sum redistribution, and benefit from the return on that investment when old. The labor income when young is defined as

$$y_t^i = n^i e^i k_t \tag{B.18}$$

where e^i is productivity, and the stock of k accumulated on average by the previous generation has a positive externality on the income of the newborn generation as in Persson and Tabellini (1994). Note that the stock of aggregate capital is accumulated as average productivity of all individuals increases. With homothetic preferences, the ratio of consumption in the two periods is independent of wealth and labor income taxation, $\frac{d_{t+1}^i}{c_t^i} = D$. Equivalently, every individual has the same "saving rate".

Each individual chooses labor supply so as to maximize

$$v_t^i = U\left[\frac{\gamma}{\gamma+D}\Big((1-\tau_t)n^i e^i k_t + r_t\Big), 1-n^i, \frac{\gamma D}{\gamma+D}\Big((1-\tau_t)n^i e^i k_t + r_t\Big)\right]. \tag{B.19}$$

The first-order condition is

$$\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t U_c - U_l + \frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t U_d = 0 \tag{B.20}$$

which determines the labor supply, $n[(1-\tau)e^i, r]$, for those who wish to work. Note that k is given due to accumulation by the previous generation. The choice depends only on the size of redistribution, r, and after-tax wage, $(1-\tau)e^i$.

Let F denote the distribution function for individual productivity, so that $F(e^i)$ is the fraction of the population with productivity less than e^i. Average labor income is obtained by integrating

$$\bar{y}_t = k_t \int_0^{\infty} e^i n\Big[(1-\tau_t)e^i, r_t\Big] dF(e^i). \tag{B.21}$$

Finally, the budget of government is balanced and all government spending is for redistribution of income. If per capita income is $\bar{y}$, then

$$\tau_t \bar{y}_t = r_t. \tag{B.22}$$

For the average individual, $k_{t+1} = \bar{y}_t - c_t$. By use of (B.17) and (B.21) I can therefore solve for the growth rate of k

$$g_t = \frac{k_{t+1} - k_t}{k_t} = \frac{D \int_0^{\infty} e^i n[(1-\tau_t)e^i, r_t] dF(e^i)}{\gamma + D} - 1. \tag{B.23}$$

Since leisure is a normal good, I have

$$\frac{\partial n^i}{\partial r_t} = \frac{(\frac{\gamma}{\gamma+D})^2(1-\tau_t)e^i k_t U_{cc} + (\frac{\gamma D}{\gamma+D})^2(1-\tau_t)e^i k_t U_{dd} - \frac{\gamma}{\gamma+D}U_{cl} + 2\frac{\gamma^2 D}{(\gamma+D)^2}(1-\tau_t)e^i k_t U_{cd} - \frac{\gamma D}{\gamma+D}U_{dl}}{-\Omega} < 0, \tag{B.24}$$

with $\frac{\partial^2 v}{\partial n^2} \equiv \Omega = \big(\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t\big)^2 U_{cc} + U_{ll} + \big(\frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t\big)^2 U_{dd} - 2\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t U_{cl} + 2(\frac{\gamma}{\gamma+D}(1-\tau_t)e^i k_t\big)^2 D U_{cd} - 2\frac{\gamma D}{\gamma+D}(1-\tau_t)e^i k_t U_{dl} < 0$, given the assumption that v is strictly concave. Hence for given productivity endowment, individual labor supply falls with increased redistribution. Therefore

$$\frac{\partial \bar{y}_t}{\partial r_t} = k_t \int_0^{\infty} e^i \frac{\partial n^i}{\partial r_t} dF(e^i) < 0. \tag{B.25}$$

This establishes that the left-hand side of (B.22) is strictly decreasing with r. Moreover, $\tau\bar{y}$ is non-negative and bounded above by τe, where e is average productivity. In turn, the right-hand side of (B.22) is strictly increasing with r. Thus, there is a unique value of r to satisfy (B.22) for any τ.

B.4.2 Political-Economic Equilibrium

In order to characterize the political economic equilibrium, the median voter m sets taxes to maximize utility subject to the budget constraints (B.16) and (B.17), and the government budget constraint (B.22):

$$v_t^m = U\left[\frac{\gamma}{\gamma+D}\Big((1-\tau_t)n^m e^m k_t + \tau_t \bar{y}_t\Big), 1-n^m, \frac{\gamma D}{\gamma+D}\Big((1-\tau_t)n^m e^m k_t + \tau_t \bar{y}_t\Big)\right], \tag{B.26}$$

and the first-order condition for the median voter with respect to the tax rate is

$$\begin{aligned}\left(\bar{y}_t - y_t^m + \tau_t \frac{d\bar{y}_t}{d\tau_t}\right)\left(\frac{\gamma}{\gamma+D}U_c + \frac{\gamma D}{\gamma+D}U_d\right) \\ + \left[\frac{\gamma}{\gamma+D}(1-\tau_t)e^m k_t U_c - U_l + \frac{\gamma D}{\gamma+D}(1-\tau_t)e^m k_t U_d\right]\frac{dn^m}{d\tau_t} = 0.\end{aligned} \tag{B.27}$$

Thus, making use of equation (B.20), the tax rate chosen by the median voter must satisfy

$$\bar{y}_t - y_t^m + \tau_t \frac{d\bar{y}_t}{d\tau_t} = 0. \tag{B.28}$$

For a given labor income inequality, the political equilibrium τ is constant over time, so that the time subscript t is suppressed henceforth. Let $\theta = 1-\tau$ be the fraction of earned income retained. From (B.21), $\bar{y}$ depends on r and θ. The total derivative of average labor income, I have

$$\begin{aligned} \frac{d\bar{y}}{d\tau} &= \frac{\partial \bar{y}}{\partial r}\frac{dr}{d\tau} + \frac{\partial \bar{y}}{\partial \theta}\frac{d\theta}{d\tau}, \\ &= \frac{\partial \bar{y}}{\partial r}\Big(\bar{y} + \tau\frac{d\bar{y}}{d\tau}\Big) - \frac{\partial \bar{y}}{\partial \theta}. \end{aligned} \tag{B.29}$$

Thus, the total derivative of average labor income with respect to changes in the tax rate is given by

$$\frac{d\bar{y}}{d\tau} = \frac{\bar{y}_r\bar{y} - \bar{y}_\theta}{1 - \tau\bar{y}_r} < 0, \tag{B.30}$$

with $\bar{y}_r = \frac{\partial \bar{y}}{\partial r}$ and $\bar{y}_\theta = \frac{\partial \bar{y}}{\partial \theta}$. Finally, substituting (B.30) into (B.28) I have

$$\begin{aligned} 0 &= \bar{y} - y^m + \tau\frac{\bar{y}_r\bar{y} - \bar{y}_\theta}{1 - \tau\bar{y}_r}, \\ &= (\bar{y} - y^m)(1-\tau) + \frac{\eta_r\bar{y}(1-\tau) - \eta_\theta\bar{y}\tau}{1-\eta_r}, \end{aligned} \tag{B.31}$$

where $\eta_r = \bar{y}_r\frac{r}{\bar{y}}$ and $\eta_\theta = \bar{y}_\theta\frac{\theta}{\bar{y}}$ are the partial elasticities of average income. Solving the above equation for τ, yields

$$\tau = \frac{m-1+\eta_r}{m-1+\eta_r+m\eta_\theta} \tag{B.32}$$

with $m = \frac{\bar{y}}{y^m}$. Identical to the spirit of Meltzer and Richard (1981), equation (B.32) yields that an increase in labor income inequality raises taxation

$$\frac{d\tau}{dm} > 0. \tag{B.33}$$

B.4.3 Labor Income Inequality and Growth

For the average individual in (B.16) and (B.17), I have

$$\begin{aligned} k_{t+1} &= \bar{y}_t - c_t, \\ &= \bar{y}_t - \frac{d_{t+1}}{D}, \\ &= \bar{y}_t - \frac{\gamma k_{t+1}}{D}. \end{aligned} \tag{B.34}$$

Solving the above equation for k_{t+1}, yields

$$k_{t+1} = \frac{D\bar{y}_t}{\gamma + D}. \tag{B.35}$$

Combining the above equation and (B.21), the growth rate of k can be obtained

$$\begin{aligned} g_t &= \frac{k_{t+1} - k_t}{k_t}, \\ &= \frac{D\int_0^\infty e^i n[(1-\tau_t)e^i, r_t]dF(e^i)}{\gamma + D} - 1. \end{aligned} \tag{B.36}$$

Again for a given labor income inequality, the political equilibrium τ and g are constant over time, so that the time subscript t is suppressed henceforth. Thus, the effect of taxation on growth, Combining (B.23) and making use of the total derivative of $\bar{y}$ (B.30), yields

$$\begin{aligned} \frac{dg}{d\tau} &= \frac{D}{\gamma + D} \frac{d\left[\int_0^\infty e^i n[(1-\tau)e^i, r]dF(e^i)\right]}{d\tau}, \\ &= \frac{D}{\gamma + D} \frac{\frac{1}{k}d\bar{y}}{d\tau} < 0. \end{aligned} \tag{B.37}$$

Thus all else equal, the higher is the labor income taxation, the lower is the growth rate.

Combining (B.33), therefore, the effect of labor income inequality on growth yields

$$\frac{dg}{dm} = \frac{dg}{d\tau}\frac{d\tau}{dm} < 0. \tag{B.38}$$

If labor income inequality increases such that divergence between mean and median labor income increases, then the preferred labor income tax rate (or redistribution) rises, and hence less growth because redistributive policies are coming from distortionary taxes that affect capital accumulation. This indicates that labor income inequality is harmful for growth which is identical in spirit to Persson and Tabellini (1994).

B.5 Redistribution and Growth

In the model, I theorize that greater capital income inequality leads to smaller tax burden on labor and thus higher subsequent economic growth. However, this implicitly indicates a negative relationship between government size and subsequent growth. Therefore, I test this relationship in Table B.2 which presents regressions of average annual per capita growth rate on the lagged total government outlays. Column 1 includes the initial income level on the right-hand side, and I can see a negative and significant relationship. The rest of the table investigates the robustness of this relationship. Column 2 includes the initial male and female education, while column 3 instead adds initial market distortions on the right-hand side. Column 4 includes all controls mentioned above, and with these controls the relationship between total outlays and growth remains negative and statistically significant at 10 percent. Column 5 in addition includes $TOPINC$ and $UTIP$, thus allowing for robustness check of how inequality affects growth. In this case, controlling for the lagged total outlays again yields positive relationship between capital income inequality and growth, in support of the mechanism proposed in this chapter. Moreover, columns 6-10 mimic columns 1-5 and show that the broad picture is also similar when I focus on the ten-year panel data.

Table B.1 Sensitivity Analysis

	1960-2010					1960-2010				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
L.OUTLAYS	-0.0781*	-0.0812**	-0.0714	-0.0747*	-0.0658	-0.0718**	-0.0763***	-0.0652**	-0.0690**	-0.0708**
	(0.0403)	(0.0383)	(0.0419)	(0.0403)	(0.0457)	(0.0266)	(0.0262)	(0.0262)	(0.0261)	(0.0321)
L.y	-0.408***	-0.415***	-0.389**	-0.396***	-0.617***	-0.363**	-0.368***	-0.353**	-0.357**	-0.321**
	(0.140)	(0.134)	(0.139)	(0.134)	(0.115)	(0.135)	(0.128)	(0.132)	(0.125)	(0.121)
L.MEDU		-0.485		-0.507	-0.425		0.222		0.249	-0.257
		(0.607)		(0.636)	(0.477)		(0.686)		(0.744)	(0.507)
L.FEDU		0.580		0.640	0.769		0.0642		0.0558	0.678
		(0.633)		(0.639)	(0.513)		(0.642)		(0.686)	(0.858)
L.PPPI			-0.00886	-0.00888	0.00616			-0.00578	-0.00681	-0.00257
			(0.00670)	(0.00727)	(0.00760)			(0.00497)	(0.00577)	(0.00875)
L.TOPINC					0.412**					0.00314
					(0.180)					(0.114)
L.UTIP					-0.142					0.0378
					(0.170)					(0.0891)
Obs	179	179	177	177	135	89	89	88	88	61
Countries	19	19	19	19	19	19	19	19	19	18
Periods	5-year	5-year	5-year	5-year	5-year	10-year	10-year	10-year	10-year	10-year
Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.591	0.593	0.593	0.595	0.655	0.699	0.704	0.699	0.705	0.741

Notes: Panel regressions of average annual per capita growth rate including fixed effects, *L.OUTLAYS*, *L.y*, *L.MEDU*, *L.FEDU*, *L.PPPI*, and robust standard errors clustered by country in parentheses. Year dummies are included in all regressions. Column (5) includes *L.TOPINC* and *L.UTIP*. Columns (6)-(10) test (1)-(5) using 10-year panel data. *, **, and *** respectively denote significance levels at 10%, 5%, and 1%.

Appendix C

Appendices to Chapter 4

C.1 Derivation of Equations (4.11) and (4.15)

The problem of the decisive voter is to maximize:

$$u\Big[(1-\tau_{c,t})\Big((1-\tau_{y,t})y_d+(\tau_{c,t}+\tau_{y,t})\frac{1+n}{2+n}\bar{y}_t\Big),(1-\tau_{c,t+1})r_{t+1}\Big]. \tag{C.1}$$

The properties of (4.9) indicate that:

$$\frac{d\bar{y}_t}{d\tau_{y,t}}=-\delta_y\bar{y}_t \tag{C.2}$$

$$\frac{d\bar{y}_t}{d\tau_{c,t}}=-\delta_c\bar{y}_t. \tag{C.3}$$

The first-order condition for the pivotal voter with respect to the labor income tax rate is:

$$(1-\tau_{c,t})\Big(-y_d+\frac{1+n}{2+n}\bar{y}_t-(\tau_{c,t}+\tau_{y,t})\frac{1+n}{2+n}\delta_y\bar{y}_t\Big)\frac{\partial u}{\partial c_{1,t}}=0. \tag{C.4}$$

and the first-order condition for the pivotal voter with respect to the expenditure tax rate is:

$$\Big[(1-\tau_{c,t})\Big(\frac{1+n}{2+n}\bar{y}_t-(\tau_{c,t}+\tau_{y,t})\frac{1+n}{2+n}\delta_c\bar{y}_t\Big)-\Big((1-\tau_{y,t})y_d+(\tau_{c,t}+\tau_{y,t})\frac{1+n}{2+n}\bar{y}_t\Big)\Big]\frac{\partial u}{\partial c_{1,t}}=0. \tag{C.5}$$

For a given n, the political equilibrium τ_y is constant over time, so that the time subscript t is suppressed henceforth. Therefore, the labor income tax rate chosen by the decisive voter yields equation (4.11) in the text

$$-y_d + \frac{1+n}{2+n}\bar{y} - (\tau_c + \tau_y)\frac{1+n}{2+n}\delta_y \bar{y} = 0. \tag{C.6}$$

and the expenditure tax rate chosen by the decisive voter yields equation (4.15) in the text

$$(1-\tau_c)\Big[\frac{1+n}{2+n}\bar{y} - (\tau_c + \tau_y)\frac{1+n}{2+n}\delta_c \bar{y}\Big] - (1-\tau_y)y + (\tau_c + \tau_y)\frac{1+n}{2+n}\bar{y} = 0. \tag{C.7}$$

C.2 Proof of Proposition 4.1

From (4.15), I have

$$(1-\tau_c)\Big[\frac{1+n}{2+n}\bar{y} - (\tau_c + \tau_y)\frac{1+n}{2+n}\delta_c \bar{y}\Big] = (1-\tau_y)y + (\tau_c + \tau_y)\frac{1+n}{2+n}\bar{y}. \tag{C.8}$$

Given that $y = \bar{y}$, and dividing through by y this yields

$$(1-\tau_c)\Big[\frac{1+n}{2+n} - (\tau_c + \tau_y)\frac{1+n}{2+n}\delta_c\Big] = (1-\tau_y) + (\tau_c + \tau_y)\frac{1+n}{2+n}. \tag{C.9}$$

From (4.11), substituting for $(\tau_c + \tau_y)$ and τ_y using $\tau_c + \tau_y = \frac{\frac{1+n}{2+n}-1}{\frac{1+n}{2+n}\delta_y}$ implies

$$(1-\tau_c)\Big[\frac{1+n}{2+n} - \frac{\frac{1+n}{2+n}-1}{\delta_y}\delta_c\Big] = 1 + \tau_c - \frac{\frac{1+n}{2+n}-1}{\frac{1+n}{2+n}\delta_y} + \frac{\frac{1+n}{2+n}-1}{\delta_y}. \tag{C.10}$$

Solving for τ_c yields

$$\tau_c = \frac{(\frac{1+n}{2+n}-1)\Big[\frac{1+n}{2+n}(\delta_y - \delta_c) - (\frac{1+n}{2+n}-1)\Big]}{\frac{1+n}{2+n}\Big[(\frac{1+n}{2+n}+1)\delta_y - (\frac{1+n}{2+n}-1)\delta_c\Big]}. \tag{C.11}$$

Substituting into $\tau_y = \frac{\frac{1+n}{2+n}-1}{\frac{1+n}{2+n}\delta_y} - \tau_c$ yields

$$\tau_y = \frac{(\frac{1+n}{2+n}-1)\Big[(\frac{1+n}{2+n}+1)\delta_y - (\frac{1+n}{2+n}-1)\delta_c\Big] - \delta_y(\frac{1+n}{2+n}-1)\Big[\frac{1+n}{2+n}(\delta_y-\delta_c) - (\frac{1+n}{2+n}-1)\Big]}{\delta_y \frac{1+n}{2+n}\Big[(\frac{1+n}{2+n}+1)\delta_y - (\frac{1+n}{2+n}-1)\delta_c\Big]}. \tag{C.12}$$

Simplifying the above two equations yields

$$T \equiv \frac{\tau_y}{\tau_c} = \frac{(3+2n)\delta_y + \delta_c}{\delta_y\Big((1+n)(\delta_y-\delta_c)+1\Big)} - 1 \tag{C.13}$$

The first-order condition with respect to n yields:

$$\begin{aligned} \frac{dT}{dn} &= \frac{2\delta_y^2\Big((1+n)(\delta_y-\delta_c)+1\Big) - \Big((3+2n)\delta_y+\delta_c\Big)\delta_y(\delta_y-\delta_c)}{\Big[\delta_y\Big((1+n)(\delta_y-\delta_c)+1\Big)\Big]^2} \\ &= \frac{\delta_y^2(2-\delta_y)+\delta_y\delta_c^2}{\Big[\delta_y\Big((1+n)(\delta_y-\delta_c)+1\Big)\Big]^2} > 0 \end{aligned} \tag{C.14}$$

given that $0 < \delta_y < 1$ and $0 < \delta_c < 1$.

References

[1] Acemoglu, Daron, and Pascual Restrepo (2017). Secular stagnation? The effect of aging on economic growth in the age of automation. *American Economic Review: Papers and Proceedings* 107(5): 174-79.

[2] Aidt, Toke S., and Peter S. Jensen (2009). The taxman tools up: An event history study of the introduction of the personal income tax. *Journal of Public Economy* 93(1-2): 160-175.

[3] Alesina, Alberto, and Dani Rodrik (1994). Distributive politics and economic growth. *Quarterly Journal of Economics* 109(2): 465-490.

[4] Alstadsæter, Annette, Niels Johannesen, and Gabriel Zucman (2017). Tax evasion and inequality. National Bureau of Economic Research Working Paper 23772.

[5] Alvaredo, Facundo, Anthony Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2011-present). *The World Wealth and Income Database*. online at http://www.wid.world.

[6] Arellano, Manuel, and Stephen R. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58(2): 277-297.

[7] Arellano, Manuel, and Olympia Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68(1): 29-51.

[8] Atkinson, Anthony B., and Christoph Lakner (2013). Wages, capital and top incomes: The factor income composition of top incomes in the USA, 1960-2005. Mimeo, Nuffield College.

[9] Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez (2011). Top incomes in the long run of history. *Journal of Economic Literature* 49(1): 3-71.

[10] Azmat, Ghazala, Alan Manning, and John Van Reenen (2012). Privatization and the decline of labour's share: International evidence from network industries. *Economica* 79: 470-492.

[11] Banerjee, Abhijit V., and Esther Duflo (2003). Inequality and growth: What can the data say?. *Journal of Economic Growth* 8(3): 267-299.

[12] Barro, Robert J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth* 5(1): 5-32.

[13] Barro, Robert J., and Jong Wha Lee (1996). International measures of schooling years and schooling quality. *American Economic Review: Papers and Proceedings* 86(2): 218-223.

[14] Barro, Robert J., and Xavier Sala-I-Martin (2003). *Economic Growth*. New York: McGraw-Hill.

[15] Bassett, William F., John P. Burkett, and Louis Putterman (1999). Income distribution, government transfers, and the problem of unequal influence. *European Journal of Political Economy* 15(2): 207-228.

[16] Baunsgaard, Thomas, and Michael Keen (2010). Tax revenue and (or?) trade liberalization. *Journal of Public Economy* 94(9-10): 563-577.

[17] Bazzi, Samuel, and Michael A. Clemens (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. *American Economic Journal: Macroeconomics* 5(2): 152-186.

[18] Benabou, Roland (1996). Inequality and growth. *National Bureau of Economic Research Macroeconomics Annual* 11: 11-74.

[19] Benabou, Roland (2000). Unequal societies: Income distribution and the social contract. *American Economic Review* 90(1): 96-129.

[20] Bergh, Andreas, and Magnus Henrekson (2011). Government size and growth: A survey and interpretation of the evidence. *Journal of Economic Surveys* 25(5): 872-897.

[21] Bertola, Giuseppe (1993). Factor shares and savings in endogenous growth. *American Economic Review* 83: 1184-1199.

[22] Besley, Timothy, and Torsten Persson (2014). Why do developing countries tax so little?. *Journal of Economic Perspectives* 28(4): 99-120.

[23] Blundell, Richard, and Stephen Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87(1): 115-143.

[24] Castello-Climent, Amparo (2010). Inequality and growth in advanced economies: An empirical investigation. *Journal of Economic Inequality* 8(3): 293-321.

[25] Chinn, Menzia D., and Hiro Ito (2006). What matters for financial development? Capital controls, institutions, and interactions. *Journal of Development Economics* 81(1): 163-192.

[26] Dabla-Norris, Era, Kalpana Kochhar, Frantisek Ricka, Nujin Suphaphiphat, and Evridiki Tsounta (2015). Causes and consequences of income inequality: A global perspective. International Monetary Fund.

[27] Deininger, Klaus, and Lyn Squire (1996). A new data set measuring income inequality. *World Bank Economic Review* 10(3): 565-591.

[28] Deininger, Klaus, and Lyn Squire (1998). New ways of looking at old issues: Inequality and growth. *Journal of Economic Growth* 57(2): 259-287.

[29] Diamond, Peter, and James Mirrlees (1971). Optimal taxation and public production, i: Production efficiency. *American Economic Review* 61(1): 8-27.

[30] Disney, Richard (2007). Population ageing and the size of the welfare state: Is there a puzzle to explain?. *European Journal of Political Economy* 23: 542-553.

[31] Facchini, Francois, Mickael Melki, and Andrew Pickering (2017). Labor costs and the size of government. *Oxford Bulletin of Economics and Statistics* 79(2): 251-275.

[32] Foellmi, Reto, and Josef Zweimuller (2006). Income distribution and demand-induced innovations. *Review of Economic Studies* 73(4): 941-960.

[33] Forbes, Kristin J. (2000). A reassessment of the relationship between inequality and growth. *American Economic Review* 90(4): 869-887.

[34] Frank, Mark W. (2009). Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measure. *Economic Inquiry* 47(1): 55-68.

[35] Frydman, Carola, and Raven E. Saks (2010). Executive compensation: A new view from a long-term perspective, 1936-2005. *Review of Financial Studies* 23(5): 2099-2138.

[36] Galasso, Vincenzo, and Paola Profeta (2007). How does ageing affect the welfare state?. *European Journal of Political Economy* 23: 554-563.

[37] Galbraith, James K., and Hyunsub Kum (2005). Estimating the inequality of household incomes: A statistical approach to the creation of a dense and consistent global data set. *Review of Income and Wealth* 51(1): 115-143.

[38] Galor, Oded, and Omer Moav (2004). From physical to human capital accumulation: Inequality and the process of development. *Review of Economic Studies* 71(4): 1001-1026.

[39] Garbinti, Bertrand, Jonathan Goupille, and Thomas Piketty (2017). Inequality dynamics in France, 1900–2014: Evidence from distributional national accounts (DINA). WID.world Working Paper.

[40] Goldin, Claudia, and Lawrence Katz (2009). *The Race between Education and Technology*. Harvard University Press, Cambridge MA.

[41] Gordon, Robert (2016). *The Rise and Fall of American Growth*. Princeton University Press, Princeton NJ.

[42] Gordon, Roger, Laura Kalambokidis, and Joel Slemrod (2004). Do we now collect any revenue from taxing capital income?. *Journal of Public Economics* 88(5): 981-1009.

[43] Gordon, Roger, and Joel Slemrod (1988). Do we collect any revenue from taxing capital income?. *Tax Policy and the Economy* 2: 89-130.

[44] Halter, Daniel, Manuel Oechslin, and Josef Zweimuller (2014). Inequality and growth: The neglected time dimension. *Journal of Economic Growth* 19: 81-104.

[45] Jones, Benjamin F., and Benjamin A. Olken (2005). Do leaders matter? National leadership and growth since World War II. *Quarterly Journal of Economics* 120(3): 835-864.

[46] Karabarbounis, Loukas, and Brent Neiman (2013). The global decline of the labor share. *Quarterly Journal of Economics* 129(1): 61-103.

[47] Kau, James B., and Paul H. Rubin (2002). The growth of government: Sources and limits. *Public Choice* 113(3-4): 389-402.

[48] Kaymak, Baris, and Markus Poschke (2016). The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics* 77: 1-25.

[49] Keen, Michael, and Ben Lockwood (2010). The value added tax: Its causes and consequences. *Journal of Development Economics* 92(2): 138-151.

[50] Kneller, Richard, Michael F. Bleaney, and Norman Gemmell (1999). Fiscal policy and growth: Evidence from OECD countries. *Journal of Public Economics* 74(2): 171-190.

[51] Kraay, Aart (2015). Weak instruments in growth regressions: Implications for recent cross-country evidence on inequality and growth. No. 7494. The World Bank.

[52] Krusell, Per, and Jose-Victor Rios-Rull (1999). On the size of US government: Political economy in the neoclassical growth model. *American Economic Review* 89(5): 1156-1181.

[53] Kuznets, Simon (1953). Shares of upper income groups in income and savings. In *Shares of Upper Income Groups in Income and Savings*: 171-218. National Bureau of Economic Research.

[54] Li, Hongyi, and Heng-fu Zou (1998). Income inequality is not harmful for growth: Theory and evidence. *Review of Development Economics* 2(3): 318-334.

[55] Luo, Weijie (2017). Inequality and growth in the 21st century. Discussion Papers 17/18, Department of Economics and Related Studies, University of York.

[56] Mello, Luiz d., and Erwin R. Tiongson (2006). Income inequality and redistributive government spending. *Public Finance Review* 34(3): 282-305.

[57] Meltzer, Allan H., and Scott F. Richard (1981). A rational theory of the size of government. *Journal of Political Economy* 89(5): 914-927.

[58] Muinelo-Gallo, Leonel, and Oriol Roca-Sagales (2013). Joint determinants of fiscal policy, income inequality and economic growth. *Economic Modelling* 30: 814-824.

[59] Perotti, Roberto (1996). Growth, income distribution, and democracy: What the data say. *Journal of Economic Growth* 1(2): 149-187.

[60] Persson, Mats (1995). Why are taxes so high in egalitarian societies?. *Scandinavian Journal of Economics* 97(4): 569-580.

[61] Persson, Torsten, and Guido Tabellini (1994). Is inequality harmful for growth?. *American Economic Review* 84(3): 600-621.

[62] Persson, Torsten, and Guido Tabellini (2003). *The Economic Effects of Constitutions*. Massachusetts Institute of Technology Press, Cambridge MA.

[63] Pickering, Andrew, and Sheraz Rajput (2018). Inequality and the composition of taxes. *International Tax and Public Finance*, forthcoming.

[64] Pickering, Andrew, and James Rockey (2011). Ideology and the growth of government. *Review of Economics and Statistics* 93(3): 907-919.

[65] Piketty, Thomas (2014). *Capital in the Twenty-First Century*. Harvard University Press, Cambridge MA.

[66] Piketty, Thomas, and Emmanuel Saez (2006). The evolution of top incomes: A historical and international perspectives. *American Economic Review: Papers and Proceedings* 96(2): 200-205.

[67] Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics*, forthcoming.

[68] Piketty, Thomas, Li Yang, and Gabriel Zucman (2017). Capital accumulation, private property and rising inequality in China, 1978-2015. National Bureau of Economic Research Working Paper 23368.

[69] Ram, Rati (1987). Wagner's hypothesis in time-series and cross-section perspectives: Evidence from real data for 115 countries. *Review of Economics and Statistics* 69(2): 194-204.

[70] Razin, Assaf, Efraim Sadka, and Phillip Swagel (2002). The aging population and the size of the welfare state. *Journal of Political Economy* 110(4): 900-918.

[71] Rodriguez, Francisco (2004). Inequality, redistribution, and rent-seeking. *Economics and Politics* 16(3): 287-320.

[72] Rodrik, Dani (1998). Why do more open economies have bigger governments?. *Journal of Political Economy* 106(5): 997-1037.

[73] Saez, Emmanuel, and Gabriel Zucman. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics* 131(2): 519-578.

[74] Saint-Paul, Gilles (1994). Unemployment, wage rigidity, and the returns to education. *European Economic Review* 38(3-4): 535-543.

[75] Sanz, Ismael, and Francisco J. Velazquez (2007). The role of ageing in the growth of government and social welfare spending in the OECD. *European Journal of Political Economy* 23: 917-931.

[76] Shelton, Cameron A. (2007). The size and composition of government expenditure. *Journal of Public Economics* 91(11-12): 2230-2260.

[77] Shelton, Cameron A. (2008). The aging population and the size of the welfare state: Is there a puzzle?. *Journal of Public Economics* 92(3-4): 647-651.

[78] Simonovits, Adras (2007). Can population ageing imply a smaller welfare state?. *European Journal of Political Economy* 23: 534-541.

[79] Teulings, Coen, and Richard Baldwin (2014). *Secular Stagnation: Facts, Causes and Cures*. Centre for Economic Policy Research Press.

[80] Voitchovsky, Sarah (2005). Does the profile of income inequality matter for economic growth?. *Journal of Economic Growth* 10(3): 273-296.

[81] Winer, Stanley L., Michael W. Tofias, Bernard Grofman, and John H. Aldrich (2008). Trending economic factors and the structure of congress in the growth of government, 1930-2002. *Public Choice* 135(3): 415-448.