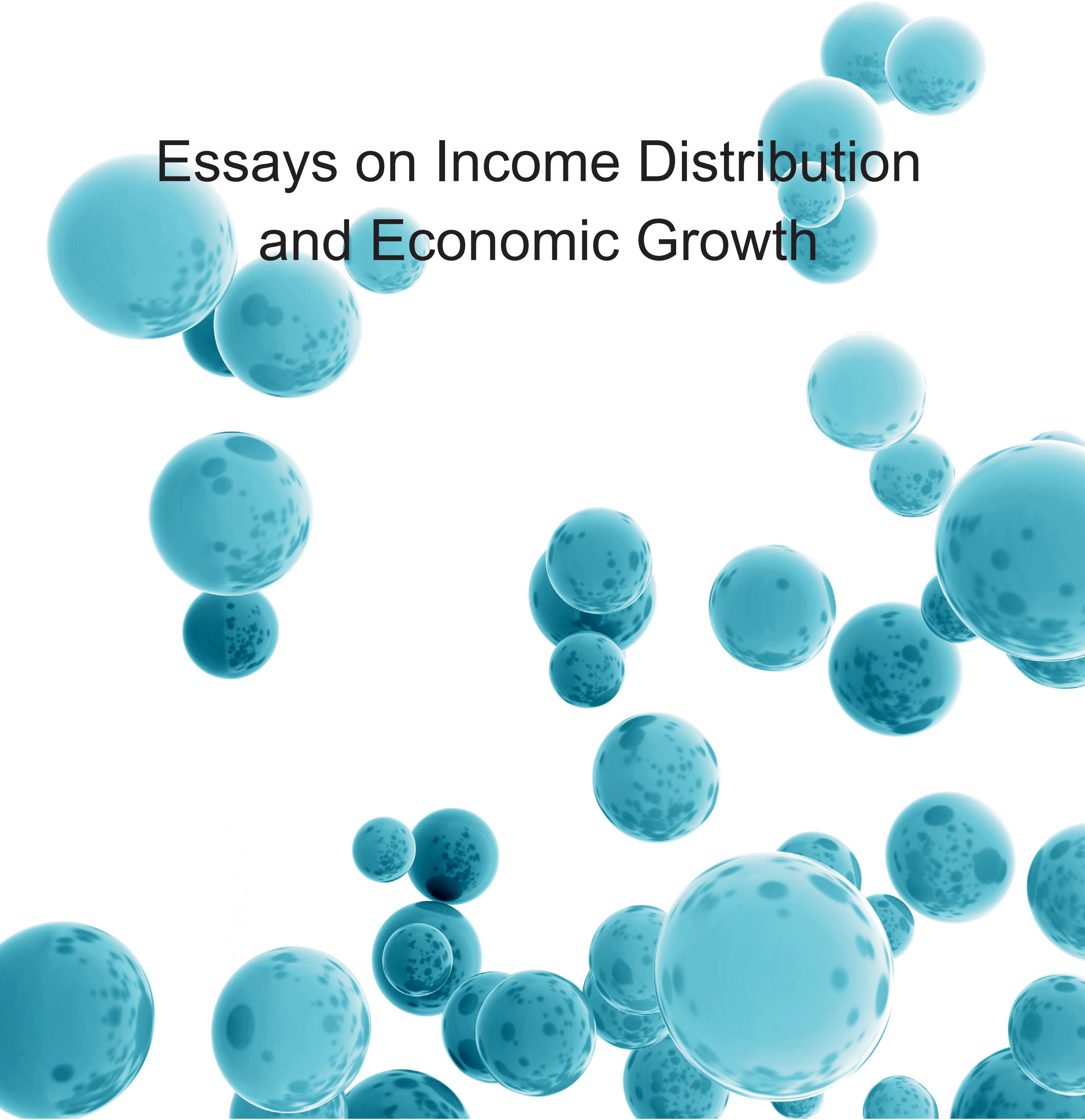
Acta Universitatis Tamperensis 2119

ELINA TUOMINEN

Essays on Income Distribution and Economic Growth



ELINA TUOMINEN

Essays on Income Distribution and Economic Growth

ACADEMIC DISSERTATION
To be presented, with the permission of
the Board of the School of Management of the University of Tampere,
for public discussion in the Paavo Koli auditorium,
Kanslerinrinne 1, Tampere,
on 18 December 2015, at 12 o'clock.

UNIVERSITY OF TAMPERE

ELINA TUOMINEN

Essays on Income Distribution and Economic Growth

Acta Universitatis Tamperensis 2119
Tampere University Press
Tampere 2015

ACADEMIC DISSERTATION
University of Tampere
School of Management
Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Cover design by
Mikko Reinikka

Distributor:
verkkokauppa@juvenesprint.fi
https://verkkokauppa.juvenes.fi

Acta Universitatis Tamperensis 2119
ISBN 978-951-44-9982-1 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1616
ISBN 978-951-44-9983-8 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2015

441 729
Painotuote

Acknowledgments

A number of people have motivated and helped me over the years as I was writing this thesis. I am deeply grateful to my supervisor, Prof. Matti Tuomala. His support has been crucial throughout this process. Further, I wish to thank the pre-examiners, Prof. Markus Jäntti and Dr. Hannu Tanninen, whose constructive comments helped me revise this thesis. I also am very pleased that Prof. Jäntti agreed to act as the opponent. He was one of the organizers of the Luxembourg Income Study (LIS) Summer Workshop in 2010; the workshop was a memorable and inspiring experience to me. I also owe my thanks to Prof. Kaisa Kotakorpi for motivating me over the years. I highly value her opinions and friendship.

At the University of Tampere, I have been lucky enough to have encouraging people around me. I appreciate the kind help and advice of Prof. Jari Vainiomäki, Prof. Jukka Pirttilä, and Dr. Sinikka Hämäläinen. I also specifically want to thank Prof. Hannu Laurila for the nice and useful conversations. Further, I have enjoyed discussions with many other colleagues; thank you, Matti Hovi, Dr. Jari Hännikäinen, Anna Kork, Prof. Jani-Petri Laamanen, Harri Nikula, and Terhi Ravaska—just to name some of you. Moreover, I want to express that the statistics courses given by Dr. Tapio Nummi and Dr. Arto Luoma have proven to be of great value; I also appreciate the helpful discussions I have had with them.

This study has received funding from the Finnish Doctoral Programme in Economics, University of Tampere, the Finnish Cultural Foundation, and the Science Foundation of the City of Tampere. During the final stages of this thesis, I have been a researcher in a project funded by the Academy of Finland. I am grateful for all this financial support. With the help of this funding, I have also participated in a number of seminars and conferences, and the essays of the thesis have benefited from the comments of the other participants in these events.

I also thank my caring parents, Merja and Harri, as well as my brother, grandparents, family-in-law, dear friends, and two lovely godchildren for supporting me and bringing joy to my life. Finally, I thank Vilppu, who is not only my loving husband but also my best friend.

Tampere, November 2015

Elina Tuominen

Abstract

Questions related to income distribution have been a popular topic in public and academic debates over the past few years. Recently published and continuously expanding data on top income shares have had a significant role in these discussions. These unprecedentedly long series on top incomes have opened up a new possibility to investigate one of the most basic questions in economics, namely, the association between income distribution and economic growth. Views about this relationship have varied over time, and empirical results have been conflicting.

This thesis is composed of four parts: an introduction and three empirical essays. The essays examine the relationship between the top 1% income shares and economic growth from different perspectives, and flexible methods are used to allow for nonlinearities. The introduction begins with a discussion of the concept of economic inequality and the background of the top income shares data. Theoretical and empirical literature on inequality and economic growth is also introduced.

In economics literature, different theories describe different mechanisms through which inequality can both invigorate and hamper economic growth. However, it is not obvious which mechanisms are more powerful than others, and empirical evidence has been mixed. The first two essays of the thesis explore the relationship between top-end inequality and subsequent economic growth. The main observation in the first essay is the negative medium- to long-run relationship between the level of top 1% income share and subsequent growth; however, this negative association is likely to become weaker in the course of economic development. The second essay extends the analysis and explores whether we should focus on changes instead of levels when we are interested in the relationship between top incomes and subsequent growth. The second essay demonstrates that the association between the level of top 1% share and growth is more evident in the data than the relationship between the change in top 1% share and growth. However, most of the data are from advanced economies, which limits the possibility of discussing these associations in less-advanced economies.

Economic development may also affect income distribution. The Kuznets hypothesis suggests that during the process of economic development, inequality first increases and then declines; this results in an inverted U-shaped relationship between inequality and economic development. This association has been explored in numerous empirical studies, but the results have not been uniform. The last essay of the thesis considers the relationship between the level of economic development and the top 1% income shares. The data show a reversal of the Kuznets curve after a certain level of development is reached. Thus, a positive association between top-end inequality and development is now observed at the highest levels of economic development.

Keywords:
inequality, top incomes, growth, development, nonlinearity, longitudinal data

Contents

Introduction

1. Background

The focus of this thesis is the examination of the relationship between income distribution and economic growth. After the introductory chapter, there are three empirical essays that employ recently published top income shares data. The first two essays involve the study of the association between top-end inequality and subsequent growth, and the third essay involves the study of the relationship between the level of economic development and top incomes.

The rest of this introductory chapter is organized as follows. This section continues with a discussion of the concept of economic inequality and the background of the top income shares data. A short discussion of the concept of economic development is also provided. Section 2 introduces the theoretical and empirical literature on inequality/growth issues. Finally, in section 3, the essays are summarized.

1.1. Why take interest in economic inequality?

Economic inequality is a widely discussed theme in sociopolitical debates. However, in economics the interest in studying this topic has varied over time. After World War II, advanced countries faced a phase of sustained, fairly stable economic growth and inequality was not considered an interesting topic. The issue of inequality was brought back into discussion in the 1970s. Publications by Sen (1973) and Atkinson (1975) have had a substantial role in building a whole branch of inequality research within the economic literature. Since the 1970s, economic inequality has risen in many developed economies, which has motivated researchers to investigate issues related to inequality. Recently, the Organisation for Economic Co-operation and Development has taken active part in discussions about the pervasive rise in income inequality (see, e.g., OECD, 2008, 2011, 2015).

Salverda et al. (2009) discuss why economists should care about inequality. The first reason is a purely scientific interest in the matter—the aspiration to understand the world that surrounds us. The second motivation is normative in nature. A researcher might be motivated by issues of social justice. But this is not the sole reason to study this topic. Namely, economic agents and decision makers often take a strong stance on inequality, and this gives reason to study the matter. The third reason stems from the desire to understand other phenomena. Many researchers may not be primarily concerned with inequality, but instead with other issues that it relates to or represents. For example, if the transmission of poverty from one generation to another and political power can be linked to inequality, understanding these relations is a significant area of research. The fourth motivation (for economists especially) is the association between economic efficiency and inequality. The standard neoclassical approach shows a tradeoff between efficiency and equality. However, understanding the conditions in which this tradeoff takes place is an important question for both theoretical and empirical research, and also for policy debates. A highly important insight following this debate is that some policies may improve both efficiency and equity—thus, avoiding the issue of a tradeoff altogether. In the Welfare State model, public spending in education and healthcare can be seen as arrangements that support growth instead of hindering it. (Salverda et al., 2009)

Studying the questions related to economic inequality from many aspects will hopefully lead to a better understanding of the dependencies. But as it is hard to deny that some degree of inequality is needed in a functional economy, it is even harder (if not impossible) to answer the question about the "optimal" or "right" level of inequality. Salverda et al. (2009) also point out that many developed countries have faced long periods of stable economic growth while having very different income distributions and social security.

1.2. On inequality and its measurement

Inequality can arise from economic processes, but inequality can also be seen as an input in many economic processes. Differences in individuals, and thus differences in incomes, are an important part of theoretical economic models where the income distribution provides incentives to work, save or take entrepreneurial risks (Welch, 1999). But the wider effects of inequality are hard to identify. Inequality can weaken some dynamics in the economy, and support others. Inequality can also be linked to ideas of *fairness* or *justice*, but these concepts cannot be described using a unique or comprehensive definition. Moreover, the idea of *equal opportunities* has been brought into discussion, and access to education and economic resources can be seen as key factors in this topic.[1]

Economic inequality can be described using various different measures. Nobel laureate Amartya Sen (1973) fits these measures into two broad categories. The first category is for measures that attempt to describe inequality in some objective way, usually using a statistical measure to describe income distribution. Examples are variance and income shares. The other category is for measures that aim to assess inequality using some normative position of welfare. For example, the Atkinson index is a normative measure. The first approach has the advantage of being able to separate observing inequality from "giving value" to inequality. The second approach encompasses ethical evaluation. However, in practice the question of objectivity becomes difficult. Even taking interest in inequality could be taken as a normative concern. Further, Atkinson (1975) states that a researcher has to recognize that summary measures of inequality, such as the Gini coefficient, include features from both categories.[2]

There are also many practical questions that the researcher must consider. For example, the researcher needs to determine, *"Inequality among whom?"* Are we talking about inequality between citizens (no matter where they live) or between countries? The overall inequality in the world would then consist of two components, namely, inequality between countries and within countries. Moreover, the researcher needs to answer the question, *"Inequality of what?"* One can talk about income or wealth inequality. In discussions over income inequality, the chosen income concept also matters. Furthermore, the definition of the time period under investigation needs to be chosen—and the length of the period depends on the research question.[3]

In addition to the conceptual issues discussed above, the unavailability of data and differences in measurement bring challenges in empirical studies. Atkinson and Brandolini (2001) illustrate that different data sources can give very different pictures of economic inequality. There can be differences

[1] *Income mobility* is also a closely related concept. Income mobility can be investigated over a person's lifetime or across generations. Ideas of high income mobility and equal opportunities are related to societies with lower income inequality. (Björklund & Jäntti, 2009; Burkhauser & Couch, 2009; Chetty et al., 2014)

[2] The Gini coefficient is often considered a statistical measure. However, the implicit welfare function attached to Gini is a rank-order weighted sum of different persons' income shares (see, e.g., Sen, 1973).

[3] For further discussion see, for example, Atkinson (1975).

in how the data have been collected or what the coverage is. In addition, it is not always evident that the income concept stays the same over time or that the data are comparable across countries. Jenkins and Micklewright (2007) emphasize that the availability of high-quality data is limited, and this has detained the evolution of empirical research on inequality. As an example one can take the Gini coefficient, which is presumably the most commonly used inequality measure. However, high-quality Gini series are hard to find. Investigating the evolution of (income) inequality over long periods of time and across countries is, thus, complicated.

1.3. Top income shares data

Recent advances in the inequality literature include a large-scale collective project that utilizes tax and population statistics in providing data on top incomes. The first book on these series, edited by Atkinson and Piketty (2007), contrasts the evidence from Continental Europe and the English-speaking countries. The second volume, also edited by Atkinson and Piketty (2010), starts to a build a global picture. Owing to this project, the World Top Incomes Database was initiated (Alvaredo et al., 2011).

Often in inequality research, the focus is in the lower part of the distribution. However, it is worth noting that changes in the upper part of the distribution affect the distribution as a whole:

> *"...understanding the concentration of incomes at the top of the distribution can tell us something about the bottom of the distribution."* (Leigh, 2009, p. 151)

Another view related to top incomes as a measure of inequality links to power. Leigh (2009) notes that concentration of incomes at the top of the distribution can have noteworthy effects on political and economic power. If a small elite has a large share of the resources in the economy, it may influence political outcomes.

Piketty (2001, 2003) generalized the ideas of Kuznets (1953) to produce top income shares data. After the example of Piketty, top income share series have been constructed by different researchers. Naturally, using tax registers as a basis for computations has its limits. For example, tax avoidance and tax evasion are problems that may be present in the data. However, it is unlikely that the overall trend is affected in a significant way (for further discussion, see, e.g., Atkinson et al., 2011). Top income shares data have advantages compared to other inequality series. These series cover longer time periods than any other income distribution data, and the series have been constructed applying the same methodology. Leigh (2007, 2009) and Roine and Waldenström (2015) also find that top income shares correlate with many other inequality measures, although top income shares focus on the upper part of the distribution. Leigh (2009) concludes that

> *"...for periods where other inequality measures are unavailable, top income shares may help fill in the gaps."* (Leigh, 2009, p. 164)

Further, Roine and Waldenström (2015) conclude that top income shares are useful as a general measure of inequality over time.[4]

Progressive income tax systems were created in most industrial countries at the beginning of the twentieth century. In countries that collected income taxes, the tax authorities started to collect and publish statistics based on income tax data. These tax statistics reported the number of taxpayers in a specific income bracket, their total income, and their tax liability. Usually this information was divided into capital income, wage income, business income, and so on. Before World War II, in most countries, there was at most 10–15% of the population under income taxation. This is why it is possible to calculate the top income shares only for the top decile (or its upper part). (Atkinson, 2007)

Piketty (2001, 2003), Piketty and Saez (2003), and Atkinson et al. (2011) have highlighted the composition of top incomes over the twentieth century. During the first half of the twentieth century, top incomes consisted mainly of capital incomes. As an example, consider the series of the United States: The biggest fall in top incomes happened during the war years and depression; the capital incomes fell dramatically under the crises and did not rise back to their previous level. One explanation for the extended fall in top income shares is progressive taxation. In contrast, during the last two or three decades, we have observed an increase in the top income shares. This growth

[4]Moreover, Alvaredo (2011) shows that when the richest group in income distribution owns a share S of total income, the Gini coefficient G can be approximated by $G^{*}(1-S)+S$, where G^{*} is the Gini coefficient for the rest of the people in this population. Alvaredo (2011) also argues that survey-based Gini coefficients could be improved by using top income shares coming from other sources because survey data usually suffer from underreporting at the top.

in top incomes started first in the United States in the 1970s, and similar development has taken place in many other countries since the 1980s. Growth in top incomes has been explained by growth in top wages, especially in the English-speaking countries. As the top wages have increased, top executives have joined capital owners at the top of the income distribution. However, top income shares have not increased substantially in the Continental European countries or Japan.

Alvaredo et al. (2013) suggest factors that would explain the recent surge in top income shares. One example of these factors is tax policy. The top pre-tax income shares have evolved in the opposite direction as the top tax rates.[5] Another example of these factors relates to the possibility of increased bargaining power and greater individualization of pay. In this case, increasing managerial remunerations may have taken place at the expense of employment and enterprise growth. Moreover, Alvaredo et al. discuss the role of capital income and inheritance.

The World Top Incomes Database project is ongoing, and new countries have been added to the database during the process of writing this thesis.[6] In the first volume on top incomes, Thomas Piketty (2007) states that the main motivation for the project was the lack of high-quality, long-spanning income distribution data. Without long-run data, it is very questionable to test for economic mechanisms that span over many years or decades. On behalf of the project, he writes:

> *"We very much hope that [...] our data will contribute to renew the literature on cross-country inequality/growth regressions."*
> (Piketty, 2007, p. 2)

It is clear that this citation has inspired this thesis work.

1.4. Measurement issues in economic development

The adequacy of commonly used measures of economic performance have been challenged, especially those based on gross domestic product (GDP), which is the most widely used measure of economic activity. GDP focuses on market production, and there is a growing concern over the relevance of

[5] Roine et al. (2009) provide empirical evidence for a negative link between top tax rates and top income shares in 16 countries.

[6] For this reason, the number of countries increases from 23 in the first essay to 25 in the second essay, and then to 26 in the last essay.

these figures as measures of economic and environmental sustainability and societal well-being.

Many of the problems with GDP statistics are well known. For example, there is the problem of the measurement of government services that are not sold on the market. In addition, changes in quality are hard to assess, and all home production is not included in GDP accounting. Moreover, due to globalization, citizens of a country may experience their own well-being very differently from the output that is produced within that country. Thus, Stiglitz et al. (2010) recommend broadening the definition. They suggest adding information about the distribution of income, consumption, and wealth into an indicator for living standards.

Despite the issues mentioned above, GDP measures have been used in the inequality–growth literature.[7] One of the main reasons for the use of these measures is the fact that alternative measures are not available over long periods of time across different countries. There are also international standards for the calculation of GDP.

2. Income inequality and economic growth

Questions related to income inequality and economic growth (or development) have been under debate for decades, and studying these issues has proven to be challenging. The direction of causality is one of the most intriguing questions because causality can run in both directions. This section discusses the literature from both aspects. The first subsection deals with several links from distribution to subsequent growth. Then, the second subsection discusses the association between the level of development and distribution in the spirit of Kuznets (1955).

2.1. The association between inequality and subsequent growth

2.1.1. From the classical approach to the modern perspective

The classical economists put forward that inequality enhances economic development (Keynes, 1920; Kaldor, 1956). This approach suggests that since the marginal propensity to save increases with wealth, more unequal distribution represents an economy where resources are directed to individuals with a higher marginal propensity to save. Thus, inequality can increase

[7] Some empirical studies have also used gross national product (GNP).

aggregate savings and capital accumulation, which leads to higher economic growth. In contrast to the classical approach, the subsequent school of neoclassical economics emphasized the view that income distribution is of no interest in the growth process. The standard neoclassical approach assumes representative, homogeneous agents. Within this view the relationship between inequality and growth is seen only as the effect of the growth process on the distribution. (Galor, 2009)

Over the past two or three decades, the role of income distribution has been brought back into discussion. Both theoretical and empirical studies have shown that income distribution has a significant role in the growth process. The modern approach includes various research papers that illustrate the detrimental effect of inequality on economic development. These studies are often classified into two approaches, namely credit market imperfection approach and the political economy approach.[8] (Galor, 2009)

The credit market imperfection channel between distribution and growth is demonstrated by Galor and Zeira (1993), who allow heterogeneous agents. In their set-up, inequality can hinder investment in human capital if the interest rate for borrowers is noticeably higher than that for lenders.[9,10] Further, Banerjee and Newman (1993) analyze the effect of inequality on occupational choices and show that inequality may deter investment in entrepreneurial activity, and thus also economic development. As an extension to this literature, Aghion and Bolton (1997) demonstrate that redistribution can enhance the efficiency of the economy because it improves the so-called trickle-down process from the rich to the poor and equality of opportunity.

Moreover, the political economy approach illustrates the notion that inequality has an adverse effect on economic development. Some early studies argued that inequality creates pressure for redistribution, but the distortions introduced by the policies hinder growth. Often this approach is called the

[8] Other issues that have been studied within this literature include questions related to gender inequality and fertility. These questions have been studied in light of industrialization and increased participation of females in the labor force. Further, issues related to ethnic and genetic diversity can be related to growth. (Galor & Weil, 1996; de la Croix & Doepke, 2003; Galor, 2009)

[9] Galor (2009) notes that publicly provided education may alleviate part of the adverse effect of inequality.

[10] However, in very poor economies, only the rich may be able to invest in education, and thus inequality may be positively associated with investment in human capital (Perotti, 1993).

fiscal policy hypothesis. Using the median voter approach, studies provided results that taxation on physical capital and human capital would be lower in more equal economies, thus decreasing the distortions in investments and improving economic growth (Perotti, 1993; Alesina & Rodrik, 1994; Persson & Tabellini, 1994). However, this political channel has lacked empirical support (e.g., Perotti, 1996). Some pursuant studies suggest that inequality may introduce an incentive for the wealthy to lobby against redistribution, and thus efficient redistribution policies may be prevented (e.g., Bénabou, 2000).

2.1.2. Unified theory and the modern perspective

The different channels described above illustrate conflicting effects. However, these theories do not explain which effect dominates another. A unified hypothesis was introduced by Galor and Moav (2004) to explain the role of inequality in the process of development. This theory includes both classical and modern perspectives in a broader framework. The unified hypothesis describes a development process in which the main engine of growth changes from physical to human capital accumulation. During this replacement process, the effect of inequality changes:

> *"In early stages of industrialization, as physical capital accumulation is a prime source of economic growth, inequality enhances the process of development by channeling resources towards individuals whose marginal propensity to save is higher. In later stages of development, however, as physical capital accumulates, the demand for human capital increases (due to capital–skill complementarity) and human capital becomes the prime engine of economic growth. [...] A more equal distribution of income, in the presence of credit constraints, stimulates investment in human capital and promotes economic growth. Lastly, as economies become wealthier and credit constraints [become] less binding while the differences in the marginal propensity to save decline, the aggregate effect of income distribution on the growth process becomes less significant."* (Galor, 2009, p. xiv)

The central idea behind the unified approach is built on the notion that human capital and physical capital accumulation processes are asymmetric. Human capital is an inherent characteristic that has diminishing returns because of physiological constraints. Thus, a widely spread human capital accumulation (education) would imply a larger aggregate stock of human capital. As long as credit constraints are binding, inequality hinders human capital accumulation. In comparison, the accumulation of the stock of physical capital is not very dependent on who owns it. Assuming that marginal propensity to save increases with income, inequality improves physical capital accumulation. (Galor, 2009)

The importance of human capital accumulation is highlighted in the second stage of the unified hypothesis. However, Galor et al. (2009) provide an economic mechanism that explains why all sectors of the economy might not benefit from human capital accumulation. The process of industrialization aroused a conflict between the interests of the landed aristocracy and the emerging capitalists—the return to land decreased. The landowners wanted to curb the mobility of the rural labor force and did not encourage education, whereas the capitalists needed new labor force and supported widely-spread education policies. In this setting, inequality in land ownership can hamper human capital accumulation, industrialization, and economic growth if the landowners can influence decision-making. In addition, Sokoloff and Engerman (2000) discuss the power of political elite who may want to maintain the existing inequality, which delays the implementation of public education and thus also economic development. (Galor, 2009)

Furthermore, inequality has been linked to sociopolitical instability, which is assumed to have an adverse effect on economic growth. Studies suggest that redistribution and educational reforms reduce sociopolitical unrest, and these policies may improve investment and economic growth (see, e.g., Alesina & Perotti, 1996; Acemoglu & Robinson, 2000; Gradstein, 2007).

2.1.3. Empirical literature on the inequality–growth relationship

Empirical studies have provided mixed evidence for the inequality–growth association. The availability and quality of data, estimation techniques, and used empirical specifications are all issues that have been raised. The empirical evidence is next discussed in relation to the challenges faced by researchers in this area.

Earlier studies in the literature applied cross-sectional data and found a negative link between the level of inequality and economic growth. These studies were usually based on ordinary least squares (OLS) analyses of cross-country data, and it was typical that the average growth rate of per capita GDP over some long period was regressed on initial inequality and several control variables, including the initial level of per capita GDP to account for the possibility of convergence.[11] For example, results by Alesina and Rodrik (1994) and Persson and Tabellini (1994) are in accordance with the fiscal policy hypothesis. Perotti (1996) studies various channels through which inequality may influence the development process. His results support the educational attainment hypothesis of Galor and Zeira (1993) and the link between income distribution and sociopolitical instability, but his results are not in line with the fiscal policy hypothesis. A summary of the early literature can be found in Bénabou (1996). However, the results of the early cross-sectional studies have been found to be sensitive to the inclusion of regional dummies or other explanatory variables, or to sample composition (see Voitchovsky, 2009, for further discussion).

The lack of data has been an obstacle for the empirical examination of the dependency. An important contribution was the introduction of the Deininger and Squire (1996) (DS) panel data set. This data set has been widely used in the literature since its release, despite its shortcomings. The quality of the DS data has been criticized, but many data sets are based on these data (for example, World Income Inequality Database, WIID). However, the impact of the data quality problem is likely to diminish as more reliable data become available. (Atkinson & Brandolini, 2001; Voitchovsky, 2009)

The DS panel data set opened up new possibilities, as it allowed more advanced estimation techniques in studying the relatioship between inequality and growth. Following the development in the growth literature, empirical studies started to use panel estimation methods. It has been argued that traditional OLS estimates are biased because of omitted country-specific effects. This view has motivated investigation of the association using fixed-effect (FE) specifications. One way to eliminate fixed country-specific effects in the estimation is to take first differences. However, because the estimation equation includes a lagged dependent variable on the right-hand side, the OLS estimate of the differenced equation (and also the FE estimate of the non-differenced equation) is likely to be biased. In addition, other explanatory variables in these models may be endogenous.[12] The generalized method

[11]It is also possible to think of sources for reverse causality, which complicates interpretations. However, using lagged right-hand-side variables in growth regressions should at least diminish this problem. Moreover, some two-stage least squares regressions were reported in the early literature.

[12]For example, literature on Kuznets relationship investigates how economic develop-

of moments (GMM) estimator based on first differences became common in the empirical literature because this technique should correct for the bias introduced by the lagged endogenous variable and it allows endogeneity in other regressors.[13] However, the first-difference GMM estimator may not be suitable in cases when variables are persistent, like inequality variables tend to be.

The DS data are exploited in a widely-known study by Forbes (2000). Her study includes both FE and first-difference GMM results. In summary, Forbes suggests that inequality has a significant positive relationship with growth in the short or medium run.[14] However, Banerjee and Duflo (2003) argue that it is not warranted that the problem related to omitted variables could be solved by including a fixed country effect in a linear specification.

The effect of measurement error has also been discussed in the empirical literature. For example, Barro (2000) argues that fixed-effects regressions that are based on differencing the data, exacerbate the measurement error problem for inequality variables. Barro considers that the variation across countries is more important than the variation over time, and he uses a three-stage least squares estimator with random country-specific effects. It turns out that Barro's results with the DS data are not in line with Forbes's results. However, Banerjee and Duflo (2003) suggest that measurement errors alone do not explain the conflicting results in the literature.

Banerjee and Duflo (2003) challenge the tradition of using linear specifications. They study the DS data using various specifications with random effects, and they also apply kernel regression. They conclude that the imposed linearity may have caused the conflicting results in the empirical studies. Contrary to previous empirical results, Banerjee and Duflo find that changes in the Gini coefficient, in any direction, are linked with lower growth rates. However, subsequent studies have continued to focus on linear specifications.

Voitchovsky (2009) points out that different mechanisms linking inequality to growth involve different definitions of inequality. Empirically, it may not be negligible which income concept is used as a basis for the inequality indicator (gross income, net income, or expenditure).[15] However, again

ment might influence inequality. This literature will be discussed in subsection 2.2.

[13]Lagged values of each of the variables are used as instruments.

[14]Further, Li and Zou (1998) estimate linear specifications with fixed and random country effects. They argue that inequality is not harmful for growth.

[15]For example, if the preferred level of redistribution is investigated, then pre-tax income

the unavailability of all types of income data limits empirical studies. The chosen inequality statistic may also be of relevance. The Gini coefficient is commonly used due to its availability and comparability to existing literature. However, as different mechanisms may relate differently to different parts of the distribution, it may be that different inequality statistics capture different mechanisms.

Voitchovsky (2005) uses the system GMM technique, which is an extended version of the first-differenced GMM procedure. Voitchovsky notes that the system GMM estimator is of interest, particularly with persistent variables such as inequality.[16] Voitchovsky finds that the upper part of the distribution is positively related to growth, but inequality further down the distribution is adversely linked to growth. For example, credit constraints on education may influence those lower down the distribution. If different parts of the distribution are differently related to growth, then one measure might not suffice to capture the whole inequality–growth relationship. Unfortunately, this approach is significantly limited by the lack of data.[17]

It has also been noted that the short lag structure of panel estimations and the long lag structure of cross-sectional studies could capture different effects of inequality on growth: the former referring to the short-term effects and the latter to the long-term effects. These effects can be different. The time dimension is discussed in a recent study by Halter et al. (2014), who use system GMM techniques and find that higher inequality may help growth in the short term, but it is harmful in the long run.[18]

Some studies indicate that the inequality–growth association varies between countries and samples. For example, Barro (2000) reports opposite effects of inequality for poor and rich countries: a positive relationship for rich countries and an adverse relationship for less-wealthy countries. In comparison, the unified growth theory gains some empirical support in a study by

---

inequality is of interest.

[16]The system GMM technique uses lagged variables as instruments in the first-differenced equations and lagged differences as instruments in the equations in levels.

[17]Voitchovsky (2005) exploits the Luxembourg Income Study (LIS) data, which is of high quality for cross-country comparisons. However, the data cover only selected years. Moreover, the inequality measures used by Voitchovsky (2005) do not reflect the very top of the distribution.

[18]Moreover, according to Berg and Ostry (2011) and Berg et al. (2012), growth duration is positively associated with income equality.

Chambers and Krause (2010), who use Gini coefficients from the WIID data. Chambers and Krause use semiparametric methods and find that, generally, inequality reduces growth in the subsequent 5-year period.

There are some previous studies that examine the empirical association between top income shares and economic growth.[19] Andrews et al. (2011) discuss the relationship using data for 12 advanced countries and suggest that inequality may foster subsequent growth when inequality is measured by the top 10% income share (after 1960). But when they use the top 1% share as their inequality measure, their results are not statistically significant in many of their specifications. Andrews et al. rely primarily on traditional linear specifications, and their preferred specifications include fixed country-specific effects.[20] Moreover, additional results by Andrews et al. do not support the idea that all changes in top income shares are related to lower growth (compare to Banerjee & Duflo, 2003). The result of a positive association between the top 10% share and growth has been challenged by Herzer and Vollmer (2013), who use modern panel cointegration techniques and argue that the long-run effect of the top 10% income share on growth is negative in nine high-income countries. However, Herzer and Vollmer also rely on prespecified functional forms.[21]

The studies on top income shares and growth can now be extended to cover a larger sample of countries, and preceding inequality–growth literature suggests that nonlinearities should be studied. The first two essays of this thesis focus on issues related to nonlinearities and time dimension in the distribution–growth regressions. The first essay investigates the link between the level of top income shares and subsequent growth. The second essay studies whether we should be interested in changes, instead of levels, when we discuss the association between top incomes and subsequent growth.

[19] Note that Roine et al. (2009) study top income shares and economic growth, but they discuss determinants of top-end inequality.

[20] Andrews et al. (2011) also report some pooled models and models with random country effects.

[21] According to Herzer and Vollmer (2013), their heterogeneous panel cointegration estimator is robust to problems such as omitted variables, slope heterogeneity, and endogenous variables.

2.2. The link between the level of economic development and inequality

2.2.1. Theoretical literature inspired by Kuznets

The analysis of inequality and development by Simon Kuznets (1955) has inspired a whole branch of literature. According to his hypothesis, as a country develops, inequality increases first and then declines after a certain development level is achieved.

Kuznets described the role of urbanization (or modernization) in the development process, and this is probably the best-known message of his paper.[22] But in his paper, he identified a number of additional factors that may bring out the famous inverted U-shaped curve between inequality and economic development. One of these additional factors was the concentration of savings among the rich, which promotes inequality as a country reaches higher income levels. Among other suggested factors was, for example, political pressure for redistribution, which would reinforce the reduction of inequality during the process of development.[23]

Various theoretical papers have studied the Kuznets-type relation. An early example of these studies is by Robinson (1976) who demonstrates that the (inverted) U relation between income (in)equality and economic development can be derived using a fairly simple model. There are also more recent theoretical papers that are related. For example, Greenwood and Jovanovic (1990) describe a process with a shift from unorganized financial structures to the modern financial system. Further, Galor and Tsiddon (1997) describe that the technological progress may drive the evolution of inequality, as the economy shifts toward using more advanced technologies. Other studies suggesting a Kuznets-type association include, for example, Anand and Kanbur (1993a), Galor and Tsiddon (1996), Aghion and Bolton (1997), and Dahan and Tsiddon (1998).

---

[22]Kuznets (1955) illustrated the effect of urbanization and industrialization using numerical examples. He did this by holding within-rural and within-urban distributions and the between-sectors income ratio constant, and then providing calculations with a population shift from the rural to the urban sector. Assuming that the rural sector incomes and inequality are lower compared to the urban sector, the population shift produced an inverted U-shaped curve.

[23]Further, Lewis (1954) discussed sectoral shifts in his study on the impact of development on distribution. Discussion on the studies by Lewis (1954) and Kuznets (1955), and their influence, can be found in Kanbur (2000).

2.2.2. *Empirical literature on the Kuznets curve*

The Kuznets hypothesis has been investigated in many empirical studies, but the results have not been uniform. The inequality data problems described earlier remain in this branch of the empirical literature, and discussion related to data quality is kept to minimum to avoid repetition. This subsection provides a brief overview of the empirical literature.[24]

Particularly within this branch of the literature, the chosen functional forms have been called into question. For example, a cross-sectional study by Ahluwalia (1976) supports the inverted-U link, but Anand and Kanbur (1993b) challenge the results with respect to chosen functional forms and data quality. Studies by Huang (2004), Lin et al. (2006), and Huang and Lin (2007) are examples of more recent cross-sectional studies that address the problem of the predetermined functional form. These studies use modern flexible methods, and the results are fairly consistent with the Kuznets hypothesis.[25]

Panel studies have become more common with the development of new inequality data sets such as that of Deininger and Squire (1996). This data set is used by Deininger and Squire (1998) and Barro (2000), who rely on prespecified functional forms.[26] The inverse-U shape holds in the cross-section or pooled results. However, Deininger and Squire (1998) reject the Kuznets curve for the fixed country effects specification.[27]

More recent studies have applied flexible methods to panel data. Frazer (2006) studies the relationship between the Gini coefficient and economic development, and in his pooled models he discovers an association that is more complex than a second-degree polynomial. Moreover, Zhou and Li (2011) conduct a nonparametric investigation with country fixed effects on the inequality–development association. They discover an inverse-U relation

[24]Further, Fields (2001) and Frazer (2006) provide overviews of the empirical literature on the Kuznets curve.

[25]To be more precise, the specification in Huang (2004) is a combination of a linear part and a stochastic nonlinear part. Moreover, Lin et al. (2006) apply penalized spline approach in semiparametric partially linear regression, and Huang and Lin (2007) provide semiparametric Bayesian inferences using a partially linear regression.

[26]Deininger and Squire (1998) use GDP per capita and 1/(GDP per capita), whereas Barro (2000) uses the logarithm of GDP per capita and its square.

[27]Moreover, Li et al. (1998) have argued that Kuznets hypothesis works better between countries at a point in time than over time within countries.

between Gini coefficients and development, but only after a certain development level is reached. In addition, Desbordes and Verardi (2012) use the semiparametric fixed effects regression estimator with Gini data and provide empirical support for the latter stages of the Kuznets relation.[28]

Various inequality indices have shown an upward trend in many countries during the past two or three decades, and the inverse-U association has been called into question. Atkinson (1995, pp. 25–26) also suspects that Kuznets would not have been surprised if the inverse-U shape no longer emerged. The shift away from manufacturing toward services has been suggested as one explanation for the rise in inequality, thus indicating a new sectoral shift in high-income economies (e.g., List & Gallet, 1999). In addition, globalization and the new role of information technology have been suggested as reasons for the recent increase in inequality. Roine and Waldenström (2015) discuss the explanations based on globalization and technological change—however, they suspect that other factors are important in explaining the evolution of top-end inequality, as discussed earlier in subsection 1.3.

Kanbur (2000) notes that the role of policy has been neglected in most inequality–development studies:

> *"The Kuznetsian literature's drive for deriving and estimating an aggregative, reduced form relationship between inequality and development has a strong tendency to minimize the role of policy—indeed, to treat the distribution/development relationship as a law. For example, this tendency is always present, no matter how hedged, in both supporters and critics of the inverted-U relationship. Supporters of the inverted-U relationship draw one of two inferences. The more left-leaning commentators view it as a warning that growth will have disruptive short-run distributional effects, with increasing inequality and perhaps even poverty. The more conservative commentators view the relationship as vindicating a drive for growth—since inequality will eventually fall, all the better to accelerate growth and get over the "hump" of the inverted-U. Those who do not find an inverted-U in the data use this finding typically to argue against those who are seen as warn-*

[28] All three of these studies use different inequality data and different methods. Frazer (2006) and Zhou and Li (2011) apply kernel-based methods whereas Desbordes and Verardi (2012) use spline-based methods.

ing against growth because of its distributional consequences—since there is no systematic relationship, no law which decrees that inequality must increase as growth accelerates, policies for accelerating growth can safely be followed (and these policies [...] may well entail inducing greater equity).” (Kanbur, 2000, p. 811)

The third essay of this thesis investigates the association between the level of economic development and top-end inequality. The top income share series provide a new possibility of exploring the distribution–development association with very long time series.

3. Summary of the essays

The three empirical essays of this thesis share a common theme of top income shares and economic growth (or development). Furthermore, all essays study nonlinearities in a flexible way. Penalized cubic regression splines are exploited within the additive model framework to allow for nonlinearities. Complex interaction structures can also be studied.[29] The estimation method is described in all essays, but, to avoid repetition, the reader may skip the description of the method in the last two essays.[30] The reader may also want to read the data description sections selectively after reading the first essay because the top 1% income share data are used in all essays.

The main contributions of this thesis are in using new data and flexible methods in studying the controversial question of the relationship between income inequality and economic growth (or development). The essays demonstrate that nonlinearities and sample composition are worth studying while exploring these associations. Instead of focusing on one specific estimate that should be able to characterize a complex relationship, a broader view is emphasized. In many cases, graphical illustrations are used to describe the discovered associations.

[29] In previous empirical literature, it has been typical to assume that control variables enter the estimation equation linearly, although the variable of interest may enter nonlinearly. In the estimated models of this thesis, the control variables' functional forms are not predetermined to be linear.

[30] Detailed information about the method can be found in Wood (2006).

3.1. Essay I. Top-end inequality and growth: Empirical evidence

This essay investigates the relationship of top 1% income shares to subsequent growth. This question has previously been studied by Andrews et al. (2011), but their sample consists of only 12 wealthy countries and they rely mainly on standard linear specifications. Moreover, many of their results on top 1% shares are not statistically significant. The data used in this essay consist of 23 countries; most of these countries are "advanced," but some "less-advanced" countries are included as well. The earliest data start from the 1920s and the latest data span to the 2000s, but the data set is not balanced. The inequality–growth association is studied using different time-period specifications, with a focus on data averaged over 5-year and 10-year periods to address the issue of time dimension. Penalized regression spline methods are utilized to allow for nonlinearities. Two different approaches are taken in the empirical analysis: the first specifications exploit the very long inequality series and are very parsimonious; the second specifications include some typical growth regression variables, but the time series are shorter. There are two reasons behind the decision to report results in two different ways. First, all data are not available for the long period. Second, there is no concensus on the "right" set of control variables in the literature.

The main results lay emphasis on "advanced" countries and their development process: the discovered negative association between top-end inequality and subsequent growth is likely to become weaker in the course of economic development. This association is observed in the medium and long term. This "fading relationship" may also explain why many of the results on top 1% shares are not significant in Andrews et al. (2011). The essay refrains from making conclusions about "less-advanced" economies due to sparse data, but the tentative findings indicate that one should not generalize the above-stated result to all types of economies. "Less-advanced" economies need to be studied further when more data become available. In summary, this essay finds a nonpositive medium- or long-run association between top-end inequality and future growth in "advanced" economies.

3.2. Essay II. Changes or levels? Reassessment of the relationship between top-end inequality and growth

The second essay is motivated by Banerjee and Duflo (2003), who discover that changes in the Gini coefficient, in any direction, are associated with lower growth in the subsequent period (that is, they find an inverse-U relationship of changes in inequality to growth). They also argue that nonlinearity may explain why the formerly reported estimates have varied greatly in the inequality–growth literature. This essay reinvestigates the linkages between top-end inequality and growth, but now the question is whether the changes in top incomes are related to subsequent growth. Previously, Andrews et al. (2011) studied top incomes in 12 wealthy countries, and their results do not support the inverse-U association between changes in top-end inequality and growth. The small number of countries and predetermined functional forms in the study by Andrews et al. motivate further analysis. Thus, this essay exploits the top 1% income share series in 25 countries from the 1920s to the 2000s. Most of these countries are "advanced." Again, penalized regression splines are used in estimation to allow for nonlinearities. As in the first essay, two different approaches are taken in the main analysis: the first models span the whole period but are very parsimonious; the second specifications investigate data from the 1950s onward but include several control variables. Moreover, both 5- and 10-year average data are studied to investigate whether the chosen period length affects the main findings.

The first discovery is that the relationship between the *level* of top 1% share and growth is more evident in the data than the association between the *change* in top-end inequality and growth. Second, the main results relate primarily to currently "advanced" countries (as in the first essay); the results demonstrate that a negative association of the *level* of top 1% shares to growth is likely to become weaker in the course of economic development. This nonpositive linkage is suggested for these countries in the medium or long run. Finally, the essay provides tentative results for "less-advanced" countries; there are no strong grounds for believing that the association between top-end inequality and growth would be similar in all types of countries. In general, the sensitivity checks illustrate that sample composition should be given attention in inequality–growth studies.

3.3. Essay III. Reversal of the Kuznets curve: Study on the inequality–development relation using top income shares data

The last essay exploits the top 1% income share series (1900–2010) in 26 countries to study the inequality–development relationship. The recent empirical inequality–development literature has challenged the use of prespecified functional forms and, thus, this study applies penalized regression splines. An important inspiration for this essay is a study by Frazer (2006). He applies nonparametric methods to Gini data and discovers a nonlinear inequality–development association that is more complex than a second-degree polynomial. It turns out that there are similarities in the overall shape of the inequality–development relationship when one compares the pooled Gini results in Frazer's study to this essay, althoug different distributional measures are used in these studies.

Various specifications in the essay show a negative association between top-end inequality and economic development after a certain level of GDP per capita has been reached. The results also demonstrate that the relationship experiences a reversal at the highest levels of economic development and, thus, a positive link is now observed in many "advanced" economies. However, earlier stages of the development process need to be studied further when more data become available. The results are also checked using data over a shorter time period (1980–2009) while controlling for urbanization and service sector. This additional analysis is motivated by the discussion about sectoral shifts—an idea that can be linked back to Kuznets. Although the essay is descriptive in nature, the empirical findings indicate that these sectoral shifts are not a sufficient explanation for changes in top-end inequality in the course of economic development. This is in line with the previous discussion within the top income literature that emphasizes other factors such as taxation (Alvaredo et al., 2011, 2013; Roine & Waldenström, 2015).

References

Acemoglu, D., Robinson, J.A., 2000. Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. Quarterly Journal of Economics 115(4), 1167–1199.

Aghion, P., Bolton, P., 1997. A Theory of Trickle-Down Growth and Development. Review of Economic Studies 64(2), 151–172.

Ahluwalia, M.S., 1976. Inequality, Poverty and Development. Journal of Development Economics 3(4), 307–342.

Alesina, A., Perotti, R., 1996. Income Distribution, Political Instability, and Investment. European Economic Review 40(6), 1203–1228.

Alesina, A., Rodrik, D., 1994. Distributive Politics and Economic Growth. Quarterly Journal of Economics 109(2), 465–490.

Alvaredo, F., 2011. A Note on the Relationship between Top Income Shares and the Gini Coefficient. Economics Letters 110(3), 274–277.

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2011. The World Top Incomes Database. Website: `http://topincomes.g-mond.parisschoolofeconomics.eu/`

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2013. The Top 1 Percent in International and Historical Perspective. Journal of Economic Perspectives 27(3), 3–20.

Anand, S., Kanbur, S.M.R., 1993a. The Kuznets process and the inequality–development relationship. Journal of Development Economics 40(1), 25–52.

Anand, S., Kanbur, S.M.R., 1993b. Inequality and development: A critique. Journal of Development Economics 41(1), 19–43.

Andrews, D., Jencks, C., Leigh, A., 2011. Do Rising Top Incomes Lift All Boats? B.E. Journal of Economic Analysis & Policy 11(1), Article 6.

Atkinson, A.B., 1995. Incomes and the Welfare State: Essays on Britain and Europe. Cambridge University Press, Cambridge.

Atkinson, A.B., 1975. The Economics of Inequality. Clarendon Press, Oxford.

Atkinson, A.B., 2007. Measuring Top Incomes: Methodological Issues, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 18–42.

Atkinson, A.B., Brandolini, A., 2001. Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study. Journal of Economic Literature 39(3), 771–799.

Atkinson, A.B., Piketty, T. (Eds.), 2007. Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T. (Eds.), 2010. Top Incomes: A Global Perspective. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T., Saez, E., 2011. Top incomes in the Long Run of History. Journal of Economic Literature 49(1), 3–71.

Banerjee, A.V., Duflo, E., 2003. Inequality and Growth: What Can the Data Say? Journal of Economic Growth 8(3), 267–299.

Banerjee, A.V., Newman, A.F., 1993. Occupational Choice and the Process of Development. Journal of Political Economy 101(2), 274–298.

Barro, R.J., 2000. Inequality and Growth in a Panel of Countries. Journal of Economic Growth 5(1), 5–32.

Bénabou, R., 1996. Inequality and Growth, in: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual. The MIT Press, Cambridge. pp. 11–74.

Bénabou, R., 2000. Unequal Societies: Income Distribution and the Social Contract. American Economic Review 90(1), 96–129.

Berg, A.G., Ostry, J.D., 2011. Inequality and Unsustainable Growth: Two Sides of the Same Coin? International Monetary Fund (IMF) Staff Discussion Note, SDN/11/08, April 8, 2011.

Berg, A., Ostry, J.D., Zettelmeyer, J., 2012. What Makes Growth Sustained? Journal of Development Economics 98(2), 149–166.

Björklund, A., Jäntti, M., 2009. Intergenerational Income Mobility and the Role of Family Background, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 491–521.

Burkhauser, R.V., Couch, K.A., 2009. Intragenerational Inequality and Intertemporal Mobility, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 522–545.

Chambers, D., Krause, A., 2010. Is the Relationship between Inequality and Growth Affected by Physical and Human Capital Accumulation? Journal of Economic Inequality 8(2), 153–172.

Chetty, R., Hendren, N., Kline, P., Saez, E., Turner, N., 2014. The Equality of Opportunity Project. Website: `http://www.equality-of-opportunity.org/`

Dahan, M., Tsiddon, D., 1998. Demographic Transition, Income Distribution, and Economic Growth. Journal of Economic Growth 3(1), 29–52.

de la Croix, D., Doepke, M., 2003. Inequality and Growth: Why Differential Fertility Matters. American Economic Review 93(4), 1091–1113.

Deininger, K., Squire, L., 1996. A New Data Set Measuring Income Inequality. World Bank Economic Review 10(3), 565–591.

Deininger, K., Squire, L., 1998. New ways of looking at old issues: Inequality and growth. Journal of Development Economics 57(2), 259–287.

Desbordes, R., Verardi, V., 2012. Refitting the Kuznets curve. Economics Letters 116(2), 258–261.

Fields, G.S., 2001. Distribution and Development: A New Look at the Developing World. The MIT Press, Cambridge, Massachusetts.

Forbes, K.J., 2000. A Reassessment of the Relationship between Inequality and Growth. American Economic Review 90(4), 869–887.

Frazer, G., 2006. Inequality and Development Across and Within Countries. World Development 34(9), 1459–1481.

Galor, O., 2009. Inequality and Economic Development: An Overview, in: Galor, O. (Ed.), Inequality and Economic Development: The Modern Perspective. Edward Elgar Publishing, Cheltenham. pp. xi–xxiv.

Galor, O., Moav, O., 2004. From Physical to Human Capital Accumulation: Inequality and the Process of Development. Review of Economic Studies 71(4), 1001–1026.

Galor, O., Moav, O., Vollrath, D., 2009. Inequality in Landownership, Human Capital Promoting Institutions and the Great Divergence. Review of Economic Studies 76(1), 143–179.

Galor, O., Tsiddon, D., 1996. Income Distribution and Growth: the Kuznets Hypothesis Revisited. Economica 63(250), S103–S117.

Galor, O., Tsiddon, D., 1997. Technological progress, Mobility, and Economic Growth. American Economic Review 87(3), 363–382.

Galor, O., Weil, D.N., 1996. The Gender Gap, Fertility, and Growth. American Economic Review 86(3), 374–387.

Galor, O., Zeira, J., 1993. Income Distribution and Macroeconomics. Review of Economic Studies 60(1), 35–52.

Gradstein, M., 2007. Inequality, Democracy, and the Protection of Property Rights. Economic Journal 117(516), 252–269.

Greenwood, J., Jovanovic, B., 1990. Financial Development, Growth and the Distribution of Income. Journal of Political Economy 98(5), 1076–1107.

Halter, D., Oechslin, M., Zweimüller, J., 2014. Inequality and growth: the neglected time dimension. Journal of Economic Growth 19(1), 81–104.

Herzer, D., Vollmer, S., 2013. Rising top incomes do not raise the tide. Journal of Policy Modeling 35(4), 504–519.

Huang, H.–C.R., 2004. A flexible nonlinear inference to the Kuznets hypothesis. Economics Letters 84(2), 289–296.

Huang, H.–C.R.. Lin, S.–C., 2007. Semiparametric Bayesian inference of the Kuznets hypothesis. Journal of Development Economics 83(2), 491–505.

Jenkins, S.P., Micklewright, J., 2007. New Directions in the Analysis of Inequality and Poverty, in: Jenkins, S.P., Micklewright, J. (Eds.), Inequality and Poverty Re-examined. Oxford University Press, Oxford. pp. 3–33.

Kaldor, N., 1956. Alternative Theories of Distribution. Review of Economic Studies 23(2), 83–100.

Kanbur, R., 2000. Income Distribution and Development, in: Atkinson, A.B., Bourguignon F. (Eds.), Handbook of Income Distribution Vol. 1. North-Holland, Amsterdam. pp. 791–841.

Keynes, J.M., 1920. The Economic Consequences of the Peace. Macmillan and Co. Ltd, London.

Kuznets, S., 1953. Shares of Upper Income Groups in Income and Saving. NBER Publication No. 55, New York.

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

Leigh, A., 2007. How Closely Do Top Income Shares Track Other Measures of Inequality? Economic Journal 117(524), F589–F603.

Leigh, A., 2009. Top Incomes, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 150–174.

Lewis, W.A., 1954. Economic Development with Unlimited Supplies of Labour. Manchester School 22(2), 139–191.

Li, H., Squire, L., Zou, H.–f., 1998. Explaining international and intertemporal variations in income inequality. Economic Journal 108(446), 26–43.

Li, H., Zou, H.–f., 1998. Income Inequality is not Harmful for Growth: Theory and Evidence. Review of Development Economics 2(3), 318–334.

Lin, S.–C., Huang, H.–C.R., Weng, H.–W., 2006. A semi-parametric partially linear investigation of the Kuznets' hypothesis. Journal of Comparative Economics 34(3), 634–647.

List, J.A., Gallet, C.A., 1999. The Kuznets Curve: What Happens After the Inverted-U? Review of Development Economics 3(2), 200–206.

OECD, 2008. Growing Unequal? Income Distribution and Poverty in OECD Countries. OECD Publishing, Paris. `http://dx.doi.org/10.1787/9789264044197-en` (June 5, 2015).

OECD, 2011. Divided We Stand: Why Inequality Keeps Rising. OECD Publishing, Paris. `http://dx.doi.org/10.1787/9789264119536-en` (June 5, 2015).

OECD, 2015. In It Together: Why Less Inequality Benefits All. OECD Publishing, Paris. `http://dx.doi.org/10.1787/9789264235120-en` (June 5, 2015).

Perotti, R., 1993. Political Equilibrium, Income Distribution, and Growth. Review of Economic Studies 60(4), 755–776.

Perotti, R., 1996. Growth, Income Distribution, and Democracy: What the Data Say. Journal of Economic Growth 1(2), 149–187.

Persson, T., Tabellini, G., 1994. Is Inequality Harmful for Growth? American Economic Review 84(3), 600–621.

Piketty, T., 2001. Les Hauts revenus en France au 20e siècle: inégalités et redistribution, 1901–1998. B. Grasset, Paris.

Piketty, T., 2003. Income Inequality in France 1901–1998. Journal of Political Economy, 111(5), 1004–1042.

Piketty, T., 2007. Top Incomes Over the Twentieth Century: A Summary of Main Findings, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 1–17.

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

Robinson, S., 1976. A Note on the U Hypothesis Relating Income Inequality and Economic Development. American Economic Review 66(3), 437–440.

Roine, J., Vlachos, J., Waldenström, D., 2009. The long-run determinants of inequality: What can we learn from top income data? Journal of Public Economics 93(7–8), 974–988.

Roine, J., Waldenström, D., 2015. Long-Run Trends in the Distribution of Income and Wealth, in: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution Vol. 2A. North-Holland, Amsterdam. pp. 469–592.

Salverda, W., Nolan, B., Smeeding, T.M., 2009. Introduction, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 3–22.

Sen, A., 1973. On Economic Inequality. Clarendon Press, Oxford.

Sokoloff, K.L., Engerman, S.L., 2000. Institutions, Factor Endowments, and Paths of Development in the New World. Journal of Economic Perspectives 14(3), 217–232.

Stiglitz, J.E., Sen, A., Fitoussi, J.–P., 2010. Mismeasuring Our Lives: Why GDP Doesn't Add Up. The New Press, New York.

Voitchovsky, S., 2005. Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution. Journal of Economic Growth 10(3), 273–296.

Voitchovsky, S., 2009. Inequality and Economic Growth, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 549–574.

Welch, F., 1999. In Defense of Inequality. American Economic Review 89(2), 1–17.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton FL.

Zhou, X., Li, K.–W., 2011. Inequality and development: Evidence from semiparametric estimation with panel data. Economics Letters 113(3), 203–207.

Essay I.

Top-end inequality and growth: Empirical evidence

Elina Tuominen

Abstract

New series of the top 1% income shares in 23 countries are used to investigate the relationship between top-end inequality and subsequent economic growth from the 1920s to the 2000s. The association is studied using different time-period specifications, with a focus on data averaged over 5- and 10-year periods. To address the issue related to chosen functional forms, penalized spline methods are exploited to allow for nonlinearities. Empirical evidence suggests that the association between top-end inequality and growth can be linked to the level of economic development. The main findings relate to currently "advanced" countries: the results show a negative relationship between top-end inequality and subsequent growth in many settings, but the findings also suggest that this association may become weaker in the course of economic development. "Less-advanced" countries need to be studied further when more data become available.

Keywords: inequality, top incomes, growth, nonlinearity, longitudinal data
JEL classification: O11, O15

Acknowledgments
The Finnish Doctoral Programme in Economics (FDPE) and the University of Tampere are gratefully acknowledged for financial support. The author thanks Oded Galor, Ravi Kanbur, Omer Moav, Hannu Tanninen, and Matti Tuomala for their thoughts. Comments by Petri Böckerman, Kari Heimonen, Mika Kortelainen, Tapio Nummi, Petra Todd, and Jari Vainiomäki, as well as participants in the FDPE Public Economics Workshop, the IIPF 2012 Congress, the EEA 2012 Congress, and the Swedish Institute for Social Research (SOFI) seminar, have also been beneficial in the earlier stages of the study. Remaining errors are the author's own.

1. Introduction

Theoretical literature has suggested numerous competing channels from income distribution to growth, and empirical studies have provided mixed evidence on the inequality–growth association. The available inequality data and the tradition of using linear specifications have been challenged, and for these reasons the current study applies flexible methods to new inequality data. This study discusses the association between the top 1% income shares and subsequent growth. Although top income shares describe the upper part of the distribution, Leigh (2007) and Roine and Waldenström (2015) provide evidence that these series also reflect changes in many other inequality measures over time. Thus, these data bring new insights into the inequality–growth literature. A brief and selective introduction to this literature is provided next (see, e.g., Voitchovsky, 2009, for a more comprehensive overview).

Theoretical models suggest that inequality can both promote and hamper growth. One of the most common arguments that inequality enhances growth is based on the classical approach: inequality channels resources toward wealthier individuals who are assumed to have a higher propensity to save; increased inequality may increase investment and thus also growth (e.g., Kaldor, 1957). Another widely mentioned mechanism is incentives: inequality encourages skilled individuals to increase their effort, which invigorates economic performance. However, productive investments can be lost if some individuals are not able to use their skills due to limited funds. The credit market imperfection approach brings forward that credit constraints at the lower part of the distribution inhibit growth: inequality reduces investment in human capital, assuming that credit constraints are binding (e.g., Galor & Zeira, 1993).[1]

Furthermore, Galor and Moav (2004) describe a unified theory that combines two contradictory approaches at different stages of the development process. Galor and Moav suggest that the classical channel dominates in the early stages of development, at which time physical capital accumulation is the main engine of growth. However, the credit market imperfection mechanism starts to dominate in the next stages of the process, at which

[1] However, the economy's income level affects this conclusion. Perotti (1993) illustrates that in very poor economies only the rich may be able to attain education, and inequality may correlate positively with investment in human capital.

time human capital is the main source of growth. Finally, Galor and Moav suggest that both mechanisms dim with development.

There are also other arguments that associate higher inequality with lower future growth. As an example, inequality may reflect polarization of power. The wealthy may have incentives to lobby against redistribution, thus preventing efficient policies (Bénabou, 2000).[2] Further, Galor et al. (2009) suggest that inequality may bring out incentives for the wealthy to impede institutional policies and changes that facilitate human capital formation and economic growth. In a more general perspective, Bénabou (1996) argues that high overall inequality may give rise to sociopolitical instability, which in turn reduces growth.

Early empirical inequality–growth studies relied on cross-sectional data, but the focus has shifted to panel studies as new data have become available.[3] In the 1990s, many cross-sectional studies found a negative link between inequality and growth (e.g., Bénabou, 1996; Perotti, 1996). However, many of the early empirical results have been called into question. It has also been suggested that the positive effects of inequality may materialize in the short term, whereas the negative effects may set in more slowly.[4] Some panel estimations, such as Li and Zou (1998) and Forbes (2000), have found a positive short- or medium-run association between inequality and subsequent growth. Recently, Halter et al. (2014) investigated the time dimension and suggest that the long-run (or total) association between inequality and growth is negative. Moreover, Barro (2000) finds that high income inequality can hinder growth in poor countries, whereas it can promote growth in rich countries.

Empirical literature has also suffered from the limited availability of high-quality inequality data. Since its release, the panel data set constructed by Deininger and Squire (1996) has been widely used despite its limitations.[5] The Luxembourg Income Study (LIS) project provides high-quality data for cross-country comparisons; unfortunately, using the LIS data results in a fairly small sample size (as discussed by, e.g., Leigh, 2007). Voitchovsky

[2]Moreover, Aghion and Bolton (1997) suggest that redistribution creates greater equality of opportunity and enhances the the trickle-down process, which is assumed to stimulate growth.

[3]Most results are based on Gini coefficient data.

[4]Many of the negative effects operate via political processes, institutional changes, and human capital formation, all of which take time to materialize.

[5]Atkinson and Brandolini (2001) demonstrate these shortcomings.

(2005) utilizes the panel features of the LIS data primarily for wealthy countries and finds that inequality is positively associated with growth in the upper part of the distribution, whereas inequality is negatively related to growth in the lower part of the distribution.[6]

Studies by Banerjee and Duflo (2003) and Chambers and Krause (2010) challenge, for example, Forbes (2000), who suggests a positive relationship between inequality and growth. Banerjee and Duflo study various specifications, including kernel regression, with the "high quality" subset of the Deininger–Squire data and find that changes in the Gini coefficient, in any direction, relate to lower subsequent growth.[7] Banerjee and Duflo argue that nonlinearity may explain why the reported estimates vary greatly in the literature. Furthermore, Chambers and Krause use semiparametric methods in their study with Gini coefficients from the World Income Inequality Database. They find that higher inequality generally reduces growth in the next 5-year period. They also provide some empirical support for the unified theory of Galor and Moav (2004).

Growth regressions without inequality variables have been studied in non- or semiparametric frameworks (e.g., Liu & Stengos, 1999; Maasoumi et al., 2007; Henderson et al., 2012). These studies highlight that important features of data are likely to be lost if linearity is forced into models. Further, the results by Banerjee and Duflo (2003) and Chambers and Krause (2010) show that linearity assumptions may be too restrictive in modeling the inequality–growth association. The contradictory evidence in the literature may be a consequence of misspecified models and low-quality inequality data. Therefore, this study applies penalized spline methods to high-quality data.

This study exploits new and unprecedentedly long inequality series. The top 1% income shares used in the current study describe top-end inequality in 23 countries from the 1920s to the 2000s. The data are explored with various time frequencies: annual data and data averaged over 5- and 10-year periods. The role of top incomes in explaining growth has previously been studied by Andrews et al. (2011), who exploit an adjusted data set from Leigh (2007). Andrews et al. use the top 10% and top 1% income shares of 12 wealthy countries and rely mainly on standard linear estimation techniques.

[6]However, the inequality measures used by Voitchovsky (2005) do not emphasize the very top of the distribution.

[7]Banerjee and Duflo (2003) also find some evidence for a negative relationship between growth rates and inequality lagged one period.

They find that after 1960, higher inequality may foster growth if inequality is measured by the top 10% income share. Recently, this result was challenged by Herzer and Vollmer (2013), who argue that the long-run effect of the top 10% share is negative. Moreover, in Andrews et al., many results for the top 1% share are not statistically significant. The small number of countries in their sample and the possibility of nonlinearities motivate the current study to investigate the top 1% further.[8,9]

This study finds a negative medium- to long-run relationship between top 1% income shares and subsequent growth, but this association is likely to become weaker in the course of economic development (as the level of per capita GDP increases). This finding relates primarily to currently "advanced" countries and is robust to various specifications. This study refrains from making conclusions about the relationship in "less-advanced" countries due to sparse data—"less-advanced" economies should be studied further when more data become available.

The rest of the paper is organized as follows. The data and methods are described in section 2. The empirical results and sensitivity analysis on the findings are provided in section 3. Finally, section 4 presents conclusions.

2. Data and methods

2.1. Data

Using tax and population statistics, it is possible to compose long and fairly consistent series on top income shares. Kuznets (1953) was the first to use this kind of data to produce top income share estimates, and Piketty (2001, 2003) generalized Kuznets's approach. Following Piketty, different researchers have constructed top income share series using the same principles of calculation. Atkinson et al. (2011) provide a thorough overview of the top income literature.[10] This study focuses on the top 1% income share series (note that this is pre-tax income). Most of the data are from "advanced"

[8]Andrews et al. (2011) also study the relationship of changes in top incomes to growth. Their results are not in line with the finding of Banerjee and Duflo (2003). The association between *changes* in the top 1% income share and subsequent growth is reassessed in a follow-up study to the current paper.

[9]Moreover, Roine et al. (2009) study top incomes and growth, but they discuss determinants of top-end inequality.

[10]In addition, see, for example, Atkinson (2007a) for the methodology. Piketty and Saez (2006), Leigh (2007), and Roine and Waldenström (2015) discuss the advantages and

economies such as Japan, as well as the English-speaking, Nordic, Continental European, and Southern European countries. Some "less-advanced" countries are also included in the total sample of 23 countries. The years from 1920 onward are studied, but the data set is not balanced. Appendix A provides more information.

The debate about how to choose control variables is put aside consciously because this study is not testing a specific channel from distribution to growth. The main goal is to explore possible nonlinearities and the overall association. For this reason and due to data availability, two different approaches are taken in the empirical analysis. First, very long time series are studied in parsimonious (henceforth, "simplified") specifications that include only the per capita GDP as a control variable to account for the level of economic development. Second, shorter time series are exploited in expanded models that include various additional covariates. Obviously, the interpretation is different in these two approaches because the influence of inequality may be channeled (at least to some extent) through some of these variables.[11]

Information from the exceptionally long inequality series is utilized in the simplified models that apply GDP per capita data 1920–2008 from Maddison (2010). In the expanded specifications, most of the data are from the Penn World Table version 7.0 (PWT 7.0) by Heston et al. (2011). The GDP per capita data span 1950–2009, and the other variables are those commonly used in growth regressions: government consumption, investment, price level of investment, and trade openness.[12] Moreover, the expanded models include average years of secondary schooling, the data of which are available every five years (Barro & Lee, 2010). More information on these variables is provided in Appendix B. Table 1 shows summary statistics with the data averaged over 5-year periods.

---

limitations of these series. The top income data are described in two volumes edited by Atkinson and Piketty (2007, 2010). The updated data are also available in the World Top Incomes Database initiated by Alvaredo et al. (2011), and the data project is ongoing.

[11]Investment is an example of this kind of variable.

[12]Price level of investment is a commonly used proxy for market distortions. Openness measure is defined as the ratio of imports plus exports to GDP.

Table 1: Descriptive statistics.

Simplified models (data from the 1920s onward)	N	min	mean	max
$top1_t$	291	3.0	10.1	23.4
$ln(GDP\ p.c.)_t$	291	6.4	8.9	10.3
$growth_{t+1}$	291	-15.2	2.3	16.1
Expanded models (data from the 1950s onward)	N	min	mean	max
$top1_t$	204	3.0	8.6	16.9
$ln(GDP\ p.c.)_t$	204	6.4	9.5	10.7
$government\ consumption_t$	204	4.0	9.5	18.6
$investment_t$	204	10.7	23.9	54.4
$price\ level\ of\ investment_t$	204	18.9	87.8	294.0
$openness_t$	204	8.0	61.6	386.3
$schooling_t$	204	0.1	2.2	5.4
$growth_{t+1}$	204	-2.7	2.4	9.6

Data averaged over 5-year periods are used in the calculations.
The 5-year periods t are defined as 1920–24, 1925–29, ..., and 2000–04.
Growth refers to average annual log growth. See footnotes 18 and 22 for more details.
Sources: see Appendix A for the top 1% shares and Appendix B for other variables.

2.2. Methods

Additive models provide a flexible framework for investigating the link between inequality and growth.[13,14] This study follows the approach presented in Wood (2006). The basic idea is that the model's predictor is a sum of linear and smooth functions of covariates:

$$\mathbb{E}(Y_i) = \boldsymbol{X}_i^* \boldsymbol{\theta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \ldots$$

In the above presentation, Y_i is the response variable (here: average annual future growth), $\boldsymbol{X}_i^*$ is a row of the model matrix for any strictly parametric model components, $\boldsymbol{\theta}$ is the corresponding parameter vector, and the $f_\bullet$ are smooth functions of the covariates, $x_\bullet$.

[13] Additive models are a special case of generalized additive models (GAMs). GAMs were introduced by Hastie and Tibshirani (1986, 1990). They present a GAM as a generalized linear model with a linear predictor that involves a sum of smooth functions of covariates. Some of the covariates can enter the model in linear form. Note here the analogy to "generalized linear models" and "linear models." This study is restricted to a special case: it uses an identity link and assumes normality in errors, which leads to additive models.

[14] In a recent study on determinants of top incomes, Roine et al. (2009) discuss the problems of using a long and narrow panel data set. For example, GMM procedures are not designed for settings where the number of countries is small but the series are long. Roine et al. run their regressions without instrumentation, which is also done here.

The flexibility of these models comes at the cost of two problems. First, one needs to represent the smooth functions $f_\bullet$ in some manner. One way to represent these functions is to use cubic regression splines, which is the approach adopted in this study. A cubic regression spline is a curve constructed from sections of cubic polynomials that are joined together so that the resulting curve is continuous up to the second derivative. The points at which sections are joined (and the end points) are the knots of the spline, and these locations must be chosen. The spline can be represented in terms of its values at the knots.[15] Second, the amount of smoothness that functions $f_\bullet$ will have needs to be chosen. Overfit is to be avoided and, thus, departure from smoothness is penalized. The appropriate degree of smoothness for $f_\bullet$ can be estimated from the data by, for example, maximum likelihood.

Illustration

Consider a model containing only one smooth function of one covariate: $y_i = f(x_i) + \epsilon_i$, where ϵ_i are i.i.d. $N(0, \sigma^2)$ random variables. To estimate function f here, f is represented so that the model becomes a linear model. This is possible by choosing a basis, defining the space of functions of which f (or a close approximation to it) is an element. In practice, one chooses basis functions, which are treated as known.

Assume that the function f has a representation $f(x) = \sum_{j=1}^{k} b_j(x)\beta_j$, where β_j are unknown parameters and $b_j(x)$ are known basis functions. Using a chosen basis for f implies that we have a linear model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where the model matrix $\boldsymbol{X}$ can be represented using basis functions such as those in the cubic regression spline basis. The departure from smoothness can be penalized with $\int f''(x)^2 dx$. The penalty $\int f''(x)^2 dx$ can be expressed as $\boldsymbol{\beta}^T \boldsymbol{S} \boldsymbol{\beta}$, where $\boldsymbol{S}$ is a coefficient matrix that can be expressed in terms of the known basis functions.

Accordingly, the penalized regression spline fitting problem is to minimize $\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + \lambda \boldsymbol{\beta}^T \boldsymbol{S} \boldsymbol{\beta}$, with respect to $\boldsymbol{\beta}$. The problem of estimating the degree of smoothness is a problem of estimating the smoothing parameter λ.[16] The penalized least squares estimator of $\boldsymbol{\beta}$, given λ, is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T \boldsymbol{X} +$

[15] There are usually two extra conditions that specify that the second derivative of the curve should be zero at the two end knots.

[16] In the estimation, one faces a bias–variance tradeoff: on the one hand, the bias should be small, but on the other hand, the fit should be smooth. One needs to compromise between the two extremes. $\lambda \longrightarrow \infty$ results in a straight line estimate for f, and $\lambda = 0$

$\lambda \boldsymbol{S})^{-1}\boldsymbol{X}^T\boldsymbol{y}$. Thus, the expected value vector is estimated as $\widehat{\mathbf{E}(\boldsymbol{y})} = \hat{\boldsymbol{\mu}} = \boldsymbol{A}\boldsymbol{y}$, where $\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{S})^{-1}\boldsymbol{X}^T$ is called an influence matrix.

This setting can be augmented to include several covariates and smooths. Given a basis, an additive model is simply a linear model with one or more associated penalties. Smooths of several variables can also be constructed. In this paper, tensor product smooths are used in cases of smooths of two variables (see Appendix C for more information).

Practical notes

The size of basis dimension for each smooth is usually not critical in estimation, because it only sets an upper limit on the flexibility of a term. Smoothing parameters control the effective degrees of freedom (*edf*). Effective degrees of freedom are defined as trace($\boldsymbol{A}$), where $\boldsymbol{A}$ is the influence matrix. The effective degrees of freedom can be used to measure the flexibility of a model. It is also possible to divide the effective degrees of freedom into degrees of freedom for each smooth. For example, a simple linear term would have one degree of freedom, and *edf*=2.3 can be thought of as a function that is slightly more complex than a second-degree polynomial.

Confidence (credible) intervals for the model terms can be derived using Bayesian methods, and approximate *p*-values for model terms can be calculated. Models can be compared using information criteria such as the Akaike information criterion (AIC). When using the AIC for penalized models (models including smooth terms), the degrees of freedom are the effective degrees of freedom, not the number of parameters. Moreover, random effects can be included in these models. For further details, see Wood (2006).[17]

---

leads to an unpenalized regression spline estimate.

[17] The results presented in this study are obtained using the R software package "mgcv" (version 1.7-21), which includes a function "gam." Basis construction for cubic regression splines is used (the knots are placed evenly through the range of covariate values by default). The maximum likelihood method is used in the selection of the smoothing parameters. The identifiability constraints (due to, for example, the model's additive constant term) are taken into account by default. The function "gam" also allows for simple random effects: it represents the conventional random effects in a GAM as penalized regression terms. More details can be found in Wood (2006) and the R project's web pages (`http://cran.r-project.org/`).

3. Results

The new top income share series allow for the overall relationship between top-end inequality and growth to be studied in various ways. First, this section reports simplified models for very long series using three different time-period specifications. Second, findings based on shorter series are reported, but these specifications include some usual growth regression variables. The section ends with additional sensitivity checks.

3.1. Simplified models: long series from the 1920s onward

The simplified models include the top 1% income share (*top1*) and ln(GDP per capita) as covariates, and the dependent variable is the future log growth of GDP per capita. The GDP per capita data of Maddison (2010) are used in these models. The relationship is investigated using annual, 5-year, and 10-year average data. The averaged data are used to mitigate the potential problems related to short-run disturbances.

The specifications in Table 2 are of the form:

$$growth_{i,t+1} = \alpha + f_1(top1_{it}) + f_2(ln(GDP\ p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it}, \quad \text{and}$$
$$growth_{i,t+1} = \alpha + f_{12}(top1_{it}, ln(GDP\ p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it},$$

where i refers to a country and t to a time period, α is a constant, functions $f_\bullet$ refer to smooth functions, δ_{decade} refers to a fixed decade effect (one decade is the reference category), u_i refers to a simple country-specific random effect ($u_i \sim \mathrm{N}(0, \sigma_u^2)$), and $\epsilon_{it} \sim \mathrm{N}(0, \sigma^2)$ is the error term; inequality and GDP variables are used as period averages (except for annual data).[18] The

[18]In the annual data (t refers to 1920, 1921, ..., 2007), future growth corresponds to the difference of $ln(GDP\ p.c.)$ values at $t+1$ and t multiplied by 100. In the 5-year average data the time periods t are 1920–24, 1925–29, ..., 2000–04. The averages of the covariates in 1920–24 are used with the subsequent period's average annual log growth (calculated using $ln(GDP\ p.c.)$ values in 1925–30); the averages of the covariates in 1925–29 are used with the following period's average annual log growth (calculated using $ln(GDP\ p.c.)$ values in 1930–35), and so on. The only exception is the future growth for the last 5-year period (2000–04): $growth_{t+1}$ is calculated using $ln(GDP\ p.c.)$ values in 2005–08 (i.e., average growth is based on three, not five, growth rates due to data unavailability in Maddison, 2010). Correspondingly, in the 10-year average data, the periods t are 1920–29, 1930–39, ..., 1990–99. The only exception to the logic is the future growth for the last 10-year period (1990–99): average growth is calculated using $ln(GDP\ p.c.)$ values in 2000–08 (i.e., $growth_{t+1}$ is not an average of ten annual growth rates but eight). Thus, in the averaged

random-effect specification allows for correlation over time within countries, and the results reflect both variations over time within countries and cross-sectional differences among countries. The random-effect approach is also used by Banerjee and Duflo (2003) who investigate nonlinearities in various specifications.[19]

Univariate smooth functions of the top 1% share and ln(GDP per capita) are studied in models (1), (3), and (5) of Table 2. Initially, the top 1% share and ln(GDP per capita) were allowed to enter in a flexible form, but $f(\mathit{top1}_t)$ had effective degrees of freedom equal to one in models (3) and (5). The models in question were then re-estimated with the assumption that *top1* enters in linear form: the coefficient for the top 1% share is negative and statistically significant in the 5- and 10-year data. Plot (a) of Figure 1 provides an illustration of the smooth $f(\mathit{top1}_t)$ with the annual data: the smooth function shows a negative slope (or possibly some U shape). Moreover, plots (b)–(d) of Figure 1 show an inverse-U shape for the smooth $f(ln(GDP\ p.c.)_t)$.

The bivariate smooths $f(\mathit{top1}_t, ln(GDP\ p.c.)_t)$ in models (2), (4), and (6) of Table 2 are visualized in Figure 2. In plots (a1)–(a2) of Figure 2, the annual data show that although the relationship between top-end inequality and growth is U-shaped at "medium" levels of economic development, the negative slope part of the U dominates.[20] The U shape is no longer evident at "high" levels of ln(GDP per capita). Plots (b1)–(b2) and (c1)–(c2) of Figure 2 show clear similarities in the overall relationship in the 5- and 10-year average data. In general, the 5- and 10-year data suggest a negative overall association between top-end inequality and future growth; however, the negative correlation seems to get weaker at the highest levels of ln(GDP per capita), as can be seen in a comparison of the slope at different levels

data models, the data points of the dependent and the explanatory variables do not overlap in the estimation equation. This should rule out direct reverse causation and reduce the endogeneity problem related to using a (lagged) GDP variable as a regressor.

[19]Further, Barro (2000) prefers random effects. He points out that differencing in the fixed-effects approach exacerbates the measurement error problem, particularly for an inequality variable, for which the variation across countries is important (Barro, 2000). In addition, Banerjee and Duflo (2003) state that there are no strong grounds for believing that the omitted variable problem could be solved by adding a fixed effect for each country, especially in a linear specification (as in, e.g., Forbes, 2000).

[20]For example, in plot (a1), look at the shape of f at $ln(GDP\ p.c.) \approx 8$ ($GDP\ p.c. \approx$ 3000 in 1990 int. GK\$) or at $ln(GDP\ p.c.) \approx 8.5$ ($GDP\ p.c. \approx$ 4900 in 1990 int. GK\$). The negative slope part of the U is more evident.

Table 2: Simplified models for 23 countries (data from the 1920s onward; GDP data from Maddison, 2010): the effective degrees of freedom for each smooth and the coefficients for the linear terms. The dependent variable is the average annual log growth in the subsequent period, where one period is 1, 5, or 10 years. See Figure 1 for illustrations of the univariate smooths with $edf > 1$, and Figure 2 for illustrations of the bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$.

	1-year data (N=1269)		5-year average data (N=291)		10-year average data (N=144)	
	(1)	(2)	(3)	(4)	(5)	(6)
$f(top1_t)$	[$edf \approx 1.7^{a}$]’	-	[linear^{a}] -0.137**	-	[linear^{a}] -0.203***	-
$f(ln(GDP\ p.c.)_t)$	[$edf \approx 2.6^{a}$]***	-	[$edf \approx 2.6^{a}$]***	-	[$edf \approx 2.7^{a}$]***	-
$f(top1_t, ln(GDP\ p.c.)_t)$	-	[$edf \approx 12.7^{b}$]***	-	[$edf \approx 5.1^{b}$]***	-	[$edf \approx 5.0^{b}$]***
adjusted r^2	0.04	0.07	0.17	0.17	0.44	0.43
AIC	7435	7409	1391	1394	520	523

***, **, *, ’ denote significance at the 1, 5, 10, and 15% levels, respectively.
The p-values for parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are reported. The smooth terms’ significance levels are based on approximate p-values.
All specifications include decade dummies and random country-specific effects.

[a]Basis dimension k for the smooth before imposing identifiability constraints is $k = 5$.
[b]Basis dimension k for the smooth before imposing identifiability constraints is $k = 5^2 = 25$ (tensor product smooth using rank 5 marginals).

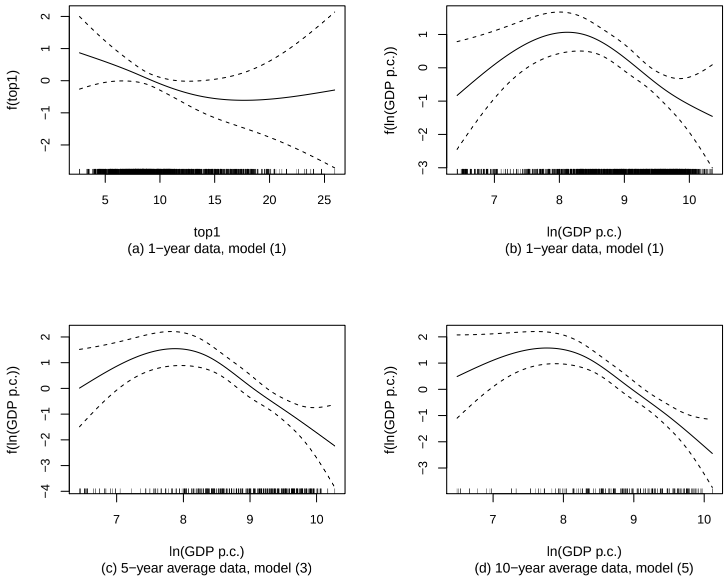
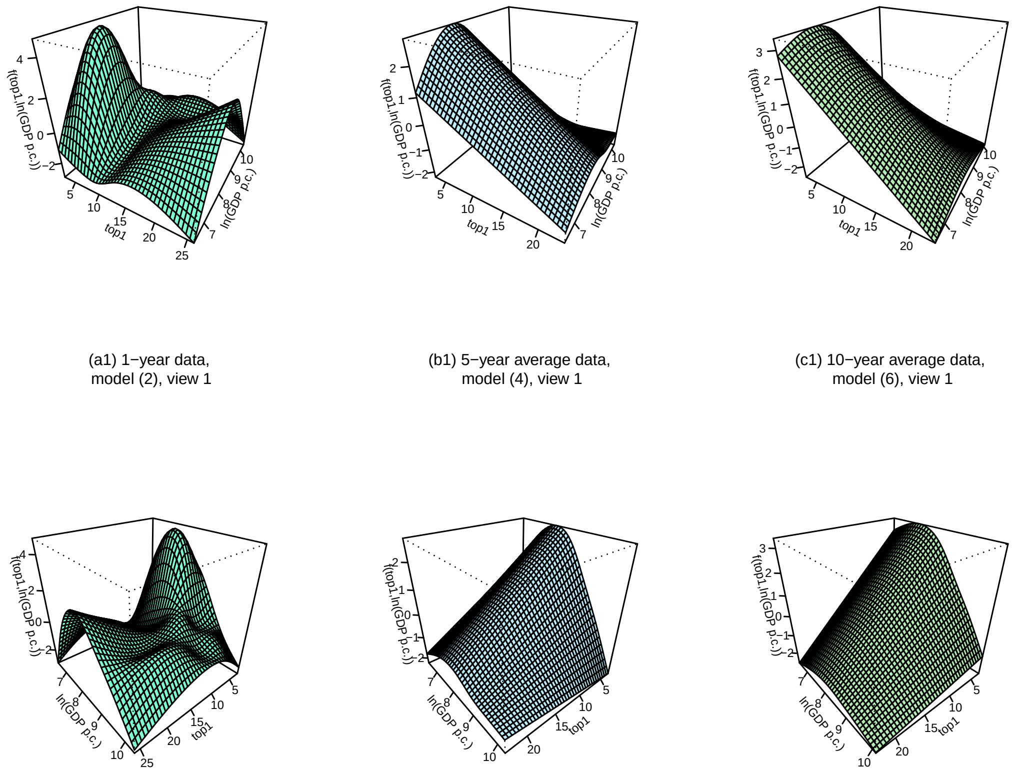


Figure 1: Visualization of the univariate smooths provided in Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as the dashed lines and the covariate values as a rug plot along the horizontal axis.



(a1) 1−year data, model (2), view 1

(b1) 5−year average data, model (4), view 1

(c1) 10−year average data, model (6), view 1

(a2) 1−year data, model (2), view 2

(b2) 5−year average data, model (4), view 2

(c2) 10−year average data, model (6), view 2

Figure 2: Visualization of the simplified models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2), (4), and (6) of Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth function f. Each smooth is illustrated from two views to clarify the shape of the smooth. For additional illustrations, see Figure D.6 in Appendix D.

of ln(GDP per capita). Furthermore, Figure D.6 in Appendix D provides additional plots that illustrate the regions that are hard to predict with the current data. In summary, there is no indication of a positive association between top-end inequality and growth in the medium or long term.

The subset of 17 "advanced" countries was also studied separately to check that the other six countries in the sample do not drive the main results.[21] The main findings about the *top1*–growth association accorded with the whole-sample results. However, stating mechanisms behind the discovered association is more or less guesswork. For example, the initially negative and then fading association between inequality and growth fits to the latter stages of the unified theory of Galor and Moav (2004). Moreover, the top 1% share may be a reasonable indicator for mechanisms that reflect the concentration of (political and economic) power. Furthermore, the years studied in this subsection also include the Great Depression and World War II. The next subsections report further results using data from the 1950s onward.

3.2. Expanded models covering years from 1950 onward

In this subsection, the models are expanded with several typical growth regression variables. This subsection investigates data averaged over 5 and 10 years because the main interest is in the medium- or long-run relationship, and the schooling data is available every five years. Note that here the used GDP per capita series span the years 1950–2009 and are from PWT 7.0 by Heston et al. (2011). The logic of constructing the averaged data is similar to that for the simplified models in the previous subsection.[22] Before estimating expanded specifications, the results that are discussed next were checked to ensure that they were not driven by the shorter time period (particularly

[21] Australia, Canada, Germany, Finland, France, Ireland, Italy, Japan, the Netherlands, Norway, New Zealand, Portugal, Spain, Switzerland, Sweden, the United Kingdom, and the United States. (The other six countries compose a heterogeneous group.)

[22] Here, the 5-year periods t are 1950–54, 1955–59, ..., 2000–04. The logic of constructing the averaged data is described also in footnote 18. As before, the only exception relates to the future growth for the last 5-year period (2000–04): due to data unavailability in PWT 7.0, $growth_{t+1}$ is calculated using $ln(GDP\ p.c.)$ values in 2005–2009 (i.e., average growth is based on four annual growth rates instead of five). Similarly, in the 10-year average data, the periods t are 1950–59, 1960–69, ..., 1990–99, and here the only exception is the future growth for the last 10-year period (1990–99): $growth_{t+1}$ is based on $ln(GDP\ p.c.)$ values in 2000–09 (i.e., it is based on nine growth rates instead of ten).

excluding the war years) and the change of the GDP data source.[23]

3.2.1. Whole-sample results

Two types of specifications are reported in Table 3. In models (1) and (3), all covariates enter the model having univariate smooths:

$$\begin{aligned} growth_{i,t+1} &= \alpha + f_1(top1_{it}) + f_2(ln(GDP\ p.c.)_{it}) + f_3(schooling_{it}) \\ &+ f_4(government\ consumption_{it}) + f_5(price\ level\ of\ investment_{it}) \\ &+ f_6(openness_{it}) + f_7(investment_{it}) + \delta_{decade} + u_i + \epsilon_{it}, \end{aligned}$$

where i refers to a country and t to a time period, α is a constant, functions $f_\bullet$ refer to smooth functions, δ_{decade} refers to a fixed decade effect (one decade is the reference category), u_i is the country-specific random effect, and ϵ_{it} is the conventional error term; variable values are period averages. Moreover, in models (2) and (4), a flexible interaction between top-end inequality and per capita GDP is allowed with a smooth of two variables: instead of $f_1(top1_{it}) + f_2(ln(GDP\ p.c.)_{it})$, a bivariate smooth $f_{12}(top1_{it}, ln(GDP\ p.c.)_{it})$ enters the specification. Again, linear terms are reported in the models of Table 3 when linearity was suggested in the initial stage of the estimation.

Models (1) and (3) in Table 3 do not allow for interaction between *top1* and the level of economic development. In model (1), the 5-year data suggest that the smooth $f(top1_t)$ is not statistically significant (the relationship between *top1* and growth can be negative or slightly U shaped; see plot (a) of Figure E.8 in Appendix E). In model (3), the 10-year data suggest a linear relationship with a negative coefficient that is statistically significant. However, models (2) and (4) with smooth $f(top1_t, ln(GDP\ p.c.)_t)$ illustrate a more complex relationship.

Figure 3 provides illustrations of the smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (4) of Table 3. In model (2) (see plots (a1)–(a2)), the 5-year data suggest that as the GDP per capita increases toward the "medium"

[23]Simplified specifications were estimated with the data from 1950 onward (i.e., models similar to those in Table 2, but using the GDP data from PWT 7.0). The results with the 5- and 10-year average data were qualitatively similar to those in subsection 3.1. Although the medium or long run is the focus of this study, the results with the annual data were also checked (in this case t refers to 1950, 1951, ..., 2008). The annual data showed a U-shaped (or even J-shaped) association between *top1* and growth at "medium" levels of GDP per capita. Details of these checks are omitted for the sake of brevity.

Table 3: Expanded models for 23 countries (data from the 1950s onward; GDP data from PWT 7.0): the effective degrees of freedom for each smooth and the coefficients for the linear terms. The dependent variable is the average annual log growth in the subsequent period, where one period is 5 or 10 years. The bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$ are illustrated in Figure 3. The univariate smooths with $edf > 1$ are illustrated in Figure E.8 in Appendix E.

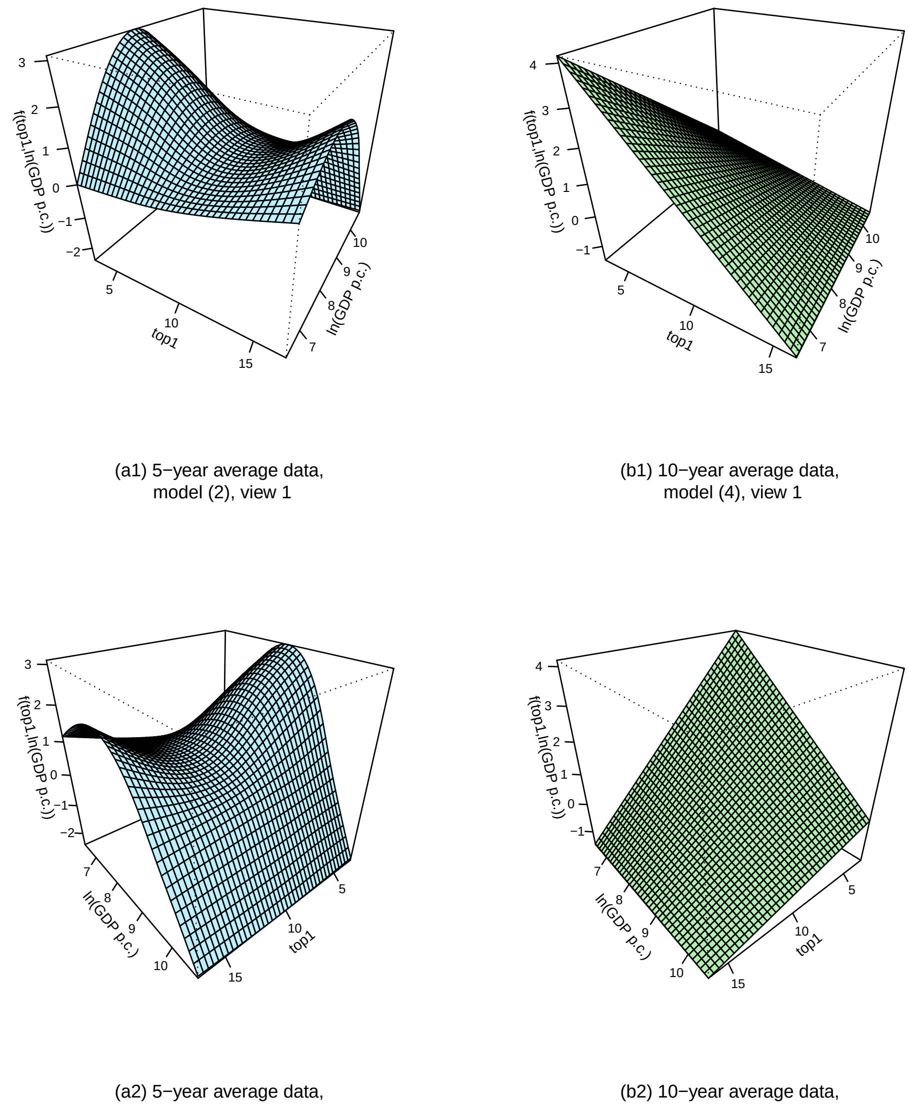
	5-year average data (N=204)		10-year average data (N=96)	
	(1)	(2)	(3)	(4)
$f(top1_t)$	[$edf \approx 2.0^a$]	-	[linear^a] -0.220***	-
$f(ln(GDP\ p.c.)_t)$	[$edf \approx 2.3^a$]***	-	[$edf \approx 1.4^a$]	-
$f(top1_t, ln(GDP\ p.c.)_t)$	-	[$edf \approx 7.2^b$]***	-	[$edf \approx 3.0^{b,c}$]***
$f(gov't\ consumption_t)$	[linear^a] 0.155***	[linear^a] 0.158***	[linear^a] 0.108**	[linear^a] 0.097*
$f(schooling_t)$	[linear^a] 0.093	[linear^a] 0.180	[$edf \approx 2.8^a$]	[$edf \approx 2.9^a$]'
$f(price\ of\ investment_t)$	[linear^a] -0.015***	[linear^a] -0.013**	[$edf \approx 2.9^a$]**	[$edf \approx 2.7^a$]'
$f(openness_t)$	[linear^a] 0.003	[linear^a] 0.005*	[$edf \approx 1.7^a$]'	[linear^a] 0.007**
$f(investment_t)$	[$edf \approx 1.7^a$]	[linear^a] 0.031	[linear^a] 0.016	[linear^a] 0.020
adjusted r^2	0.47	0.46	0.55	0.54
AIC	749	752	319	321

***, **, *, ' indicate significance at the 1, 5, 10, and 15% levels, respectively.
The p-values for parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are reported.
The smooth terms' significance levels are based on approximate p-values.
All specifications include decade dummies and random country-specific effects.

[a] Basis dimension k for the smooth before imposing identifiability constraints is $k = 5$.

[b] Basis dimension k for the smooth before imposing identifiability constraints is $k = 5^2 = 25$ (tensor product smooth with rank 5 marginals).

[c] With 3 degrees of freedom, the tensor product smooth refers to $\theta_1 top1_t + \theta_2 ln(GDP\ p.c.)_t + \theta_3 top1_t ln(GDP\ p.c.)_t$, where $\theta_\bullet$ are coefficients. When model (4) is estimated using this specification in place of $f(top1_t, ln(GDP\ p.c.)_t)$, the obtained coefficients are $\hat{\theta}_1 = -0.922$*, $\hat{\theta}_2 = -1.266$**, and $\hat{\theta}_3 = 0.077$. For example, if $GDP\ p.c.$ is 8100 (2005 I\$), then $ln(GDP\ p.c.) \approx 9$, and the slope with respect to $top1$ is approximately -0.23. Correspondingly, if $GDP\ p.c.$ is 22000 (2005 I\$), then $ln(GDP\ p.c.) \approx 10$, and the slope is approximately -0.15. This change in the slope is illustrated in plots (b1)–(b2) of Figure 3.



(a1) 5−year average data,
model (2), view 1

(b1) 10−year average data,
model (4), view 1

(a2) 5−year average data,
model (2), view 2

(b2) 10−year average data,
model (4), view 2

Figure 3: Visualization of the expanded models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (4) of Table 3 (data from the 1950s onward; GDP data from PWT 7.0). The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth f. The smooths are illustrated from two views. For additional illustrations, see Figure E.7 in Appendix E.

levels of economic development, top-end inequality is in a U-shaped relationship to growth; however, the negative slope of this U dominates. The U shape fades at even higher levels of GDP per capita. In model (4) (see plots (b1)–(b2)), the 10-year data show a negative relationship between top-end inequality and growth; however, these data also show that the association may start to fade at the highest levels of GDP per capita. Additional plots of these bivariate smooths are provided in Figure E.7 in Appendix E.

Causal channels are not in the focus of the current study, but it is tempting to speculate about the results of the models in Table 3. Although the models include, for example, investment and education variables, the data still indicate a relationship between top-end inequality and growth, and this association may depend on the country's level of economic development. Some mechanisms related to polarization of power might provide (at least a partial) explanation. Moreover, it is noteworthy that all models of Table 3 suggest a positive association between government consumption and growth.

In summary, this subsection demonstrates that the main findings in the 10-year data are robust to the inclusion of several controls. In comparison, the 5-year data show some discrepancies compared to simplified models. These disparities arise at "medium" levels of economic development: the shape of the smooth $f(top1_t, ln(GDP\ p.c.)_t)$ in plots (a1)–(a2) of Figure 3 differs from the shape shown in plots (b1)–(b2) of Figure 2; a slight U shape arises after including more covariates (see also footnote 23 for further discussion). The next subsection provides sensitivity checks and discusses the discovered U shape at "medium" levels of economic development in the 5-year data.

3.2.2. Sensitivity of the expanded models' results

The sensitivity of the whole-sample results is assessed from different aspects. The first checks relate to the composition of the sample. The subsequent robustness check involves the set of control variables in the expanded models. Finally, an alternative per capita GDP series is tested.

For the first step, 5-year specifications similar to models (1) and (2) of Table 3 were fitted separately for the English-speaking, Nordic, Continental and Southern European, and "less-advanced" countries.[24] Results for the Conti-

[24]English-speaking: Australia, Canada, Ireland, New Zealand, the United Kingdom, and the United States (N=61). Nordic: Finland, Norway, and Sweden (N=33). Continental and Southern European: France, Germany, Italy, the Netherlands, Portugal, Spain, and

nental and Southern European countries suggested a negative link between top-end inequality and growth. A negative (or slightly U-shaped) relationship was found for the Nordic countries. For the English-speaking countries, a negative (or slightly inverse-U-shaped) association between *top1* and growth was discovered. Furthermore, the small and very heterogeneous sample of "less-advanced" countries indicated a positive association between top-end inequality and growth, but the relationship was not statistically significant. These group-wise findings can help explain the U shape between top-end inequality and growth at "medium" levels of economic development (see plots (a1)–(a2) of Figure 3). It is possible that the association between top-end inequality and growth is different in "less-advanced" and "advanced" countries, at least in the medium term (in the 5-year data in this case). However, this result for "less-advanced" economies is tentative and should be tested with a larger sample when new data become available.

For the second step, Japan and the English-speaking, Continental and Southern European, and Nordic countries (17 countries in total) were used to represent "advanced," wealthy countries. The association between top-end inequality and growth was not statistically significant in the 5-year data, but the results indicated that the relationship would be "negative but fading." This is in line with the whole-sample results at the highest levels of ln(GDP per capita). The "fading link" may also provide an explanation for why many results for the top 1% income shares are not significant in Andrews et al. (2011), who study 12 wealthy countries.

For the next step, more parsimonious versions of the specifications in Table 3 were estimated. The so-called Perotti-style specifications are often used in inequality–growth estimations: in addition to inequality and GDP variables, they include schooling and price-of-investment variables. The results of these parsimonious models were in line with the previous findings. The detailed results are not reported for conciseness.

Finally, the robustness was checked with respect to the chosen GDP series, because PWT 7.0 (Heston et al., 2011) provides alternatives. The specifications in columns (2) and (4) of Table 3 were estimated using alternative series, and the overall patterns were similar to those reported above with the

Switzerland (N=55). "Less-advanced:" Argentina, China, India, Indonesia, and South Africa (N=35). Note that Japan (N=11) and Singapore (N=9) are difficult to fit into these categories.

5- and 10-year data.[25] Thus, the main results should not be driven by the choice of the GDP per capita series.

4. Conclusions

Various studies have discussed the relationship between inequality and subsequent growth. However, this study takes a novel approach to this question by exploiting new inequality series on top income shares and focusing on possible nonlinearities. Penalized splines are used to circumvent problems related to prespecified functional forms, and a complex interaction between top-end inequality and economic development is allowed in many specifications.

The main results of this study relate to currently "advanced" economies, for which a pattern is found in data averaged over 5- and 10-year periods; the overall association between top-end inequality and growth appears to be negative, but this relationship becomes weaker in the course of economic development. Although the current study refrains from making causal claims, the findings accord with the growing literature, suggesting that high inequality does not foster growth in the long run. Moreover, the main results of this study should not be generalized to all types of economies—"less-advanced" economies need to be studied further when more data become available. It will also be interesting to see how the recent economic downturn appears in the results of future studies.

[25]The series *"rgdpch"* from PWT 7.0 data was tested. This series refers to "PPP converted GDP per capita (chain series), at 2005 constant prices."

Appendix A. Information on the top 1% income share series

This is a list of the countries and sources for the top 1% income share series used in this study.[26] For better comparability, series "without capital gains" have been selected when possible. See the source for more details on the series. The 5-year average series are presented in Figure A.4 below.

1. **Argentina**: Alvaredo (2010a): Table 6.5, years 1932–2004.
2. **Australia**: Atkinson & Leigh (2007a): Table 7.1, years 1921–2002; Leigh (2010): Excel file, years 2003–2007.
3. **Canada**: Saez & Veall (2007): Table 6B.1, years 1920–2000.
4. **China**: Atkinson et al. (2010): Table 13A.12, years 1986–2003.[27]
5. **Finland**: Jäntti et al. (2010): Table 8A.2 (taxable income/population), years 1920–1992; Riihelä, M. (2011): updated figures, years 1993–2008.[28]
6. **France**: Piketty (2007): Table 3A.1, years 1920–1995; Landais (2008): Excel file, years 1996–2006.
7. **Germany**: Dell (2007): Table 9I.6, years 1925–1998.
8. **India**: Banerjee & Piketty (2010): Table 1A.5, years 1922–1999.
9. **Indonesia**: Leigh & van der Eng (2010): Table 4.1, years 1921–1939, 1982–2004.
10. **Ireland**: Nolan (2007): Table 12.5, years 1938, 1943, 1975–2000.
11. **Italy**: Atkinson et al. (2010): Table 13A.22, years 1974–2005.[29]
12. **Japan**: Moriguchi & Saez (2010): Table 3A.2, years 1920–2005.
13. **Netherlands**: Salverda & Atkinson (2007): Table 10.2, years 1920–1999.
14. **New Zealand**: Atkinson et al. (2010): Table 13A.6, years 1921–2005.[30]
15. **Norway**: Aaberge & Atkinson (2010): Table 9.1, years 1929–2006.
16. **Portugal**: Alvaredo (2010b): Table 11D.1, years 1976–2005. Note: figures 1976–1982 have been updated to equal figures provided on the website by Alvaredo et al. (2011) (Feb 18, 2011).
17. **Singapore**: Atkinson et al. (2010): Table 13A.15, years 1950–2005.[31] (Note. *top1* data also available for 1947–1949, but GDP data not available.)
18. **South Africa**: Alvaredo & Atkinson (2010): Table A.5B & Table A.5C, years 1950–1993 & 2002–2007. (Note. *top1* data also available for 1944–1949, but GDP data not available.)
19. **Spain**: Alvaredo & Saez (2010): Table 10D.2, years 1981–2005.
20. **Sweden**: Roine & Waldenström (2010): Table 7A.2, years 1920–2006.
21. **Switzerland**: Atkinson et al. (2010): Table 13A.9, years 1933–1996.[32]
22. **United Kingdom**: Atkinson et al. (2010): Table 13A.2, years 1937, 1949–2005.[33]
23. **United States**: Saez (2010): Excel file, Table A1, years 1920–2008.[34]

[26] The data correspond to the available series at the end of 2010/beginning of 2011. Most figures are collected from the two volumes edited by Atkinson and Piketty (2007, 2010). The original series in the first volume is referred to where the series had not been updated for the second volume. After collecting the data, the series were published in the World Top Incomes Database initiated by Alvaredo et al. (2011).

[27] For more information and the original series, see Piketty and Qian (2010).

[28] Figures 1993–2008 received directly from Marja Riihelä by email (Feb 11, 2011).

[29] For more information and the original series, see Alvaredo and Pisano (2010).

[30] For more information and the original series, see Atkinson and Leigh (2007b).

[31] For more information and the original series, see Atkinson (2010).

[32] In the original source: For all years except 1933, the estimates relate to income averaged over the year shown and the following year (for more information, see also Dell et al., 2007). Thus, the same value is repeated for two successive years in the current study.

[33] For more information and the original series, see Atkinson (2007b).

[34] For more information and the original series, see Piketty and Saez (2007).

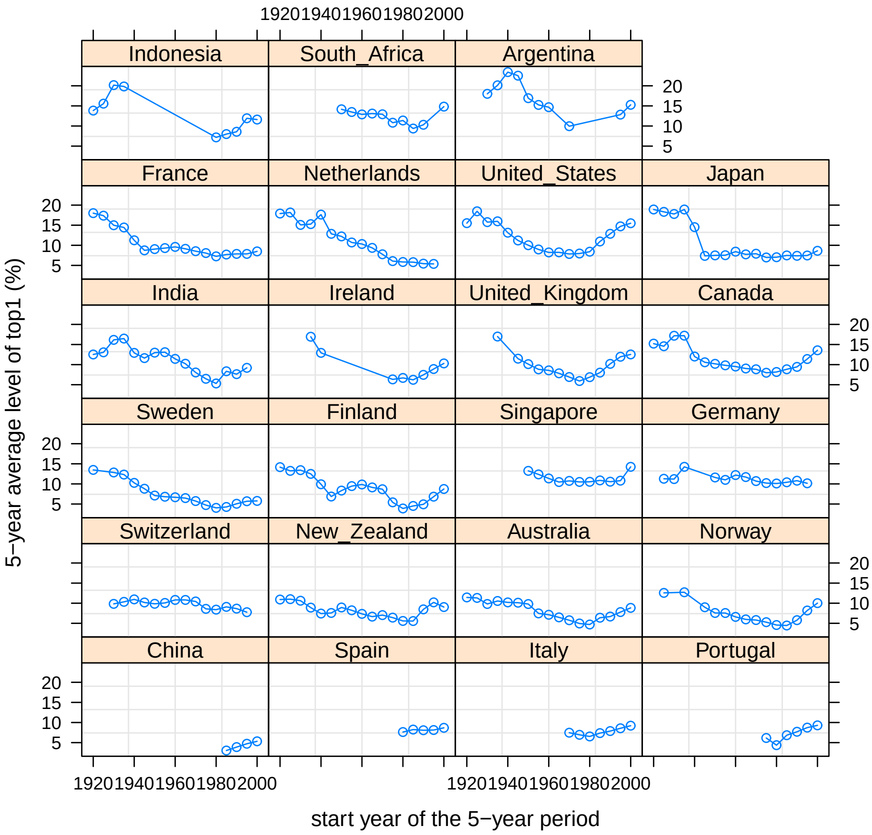


Figure A.4: Top 1% income shares for each country (5-year average data used in models of Table 2; the time periods are 1920–24, 1925–29, ..., and 2000–04). Data sources: see list of countries in this appendix.

Appendix B. Information on other variables

Long series, simplified models (annual observations span 1920–2008):

– GDP per capita, 1990 international GK$; Maddison (2010). See Figure B.5 for illustration.

Expanded models (annual observations span 1950–2009):

– GDP per capita: PPP converted GDP per capita (Laspeyres), derived from growth rates of domestic absorption, at 2005 constant prices (2005 I$/person); PWT 7.0 by Heston et al. (2011) (*"rgdpl2"*)

– Government consumption share of PPP converted GDP per capita at current prices (%); PWT 7.0 by Heston et al. (2011) (*"cg"*)

– Investment share of PPP converted GDP per capita at current prices (%); PWT 7.0 by Heston et al. (2011) (*"ci"*)

– Openness at current prices (%); PWT 7.0 by Heston et al. (2011) (*"openc"*)

– Price level of investment (PPP over investment/XRAT, where XRAT is national currency units per US dollar); PWT 7.0 by Heston et al. (2011) (*"pi"*)

– Average years of secondary schooling for total population (population aged 25 and over); Barro and Lee (2010); available every five years starting from 1950

– Note: "China Version 2" data from PWT 7.0 (Heston et al., 2011) is used.

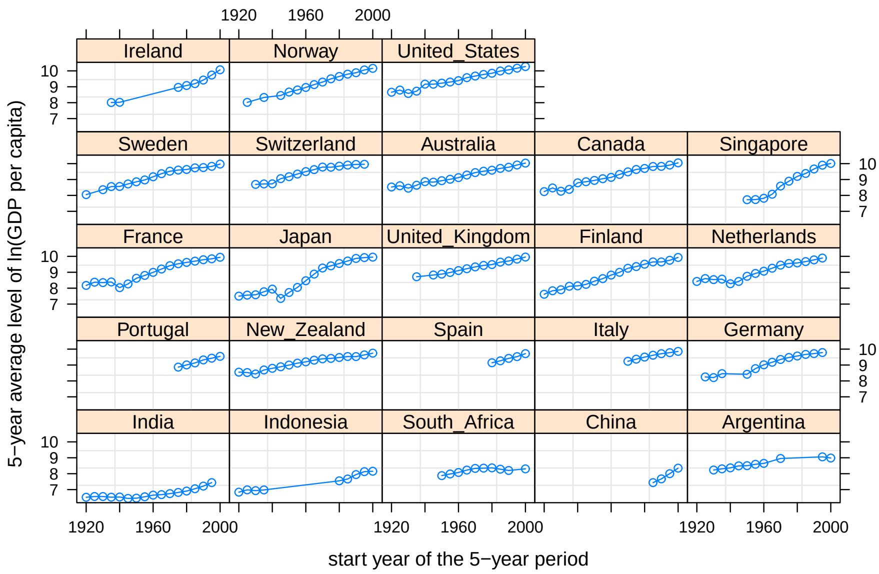


Figure B.5: Level of economic development for each country (5-year average data used in models of Table 2; the time periods are 1920–24, 1925–29, ..., and 2000–04). Data source: Maddison (2010).

Appendix C. Tensor product smooths

This appendix provides additional information to subsection 2.2. Tensor product smooths are recommended if one uses a smooth that contains more than one variable, but the scales of those variables are fundamentally different (i.e., measured in different units). Smooths of several variables are constructed from marginal smooths using the tensor product construction. The basic idea of a smooth function of two covariates is provided as an example.

Consider a smooth comprised of two covariates, x and z. Assume that we have low-rank bases to represent smooth functions f_x and f_z of the covariates. We can then write:

$$f_x(x) = \sum_{i=1}^{I} \alpha_i a_i(x) \quad \text{and} \quad f_z(z) = \sum_{l=1}^{L} \delta_l d_l(z),$$

where α_i and δ_l are parameters, and the $a_i(x)$ and $d_l(z)$ are known (chosen) basis functions such as those in the cubic regression spline basis.

Consider then the smooth function f_x. We want to convert it to a smooth function of both x and z. This can be done by allowing the parameters α_i to vary smoothly with z. We can write:

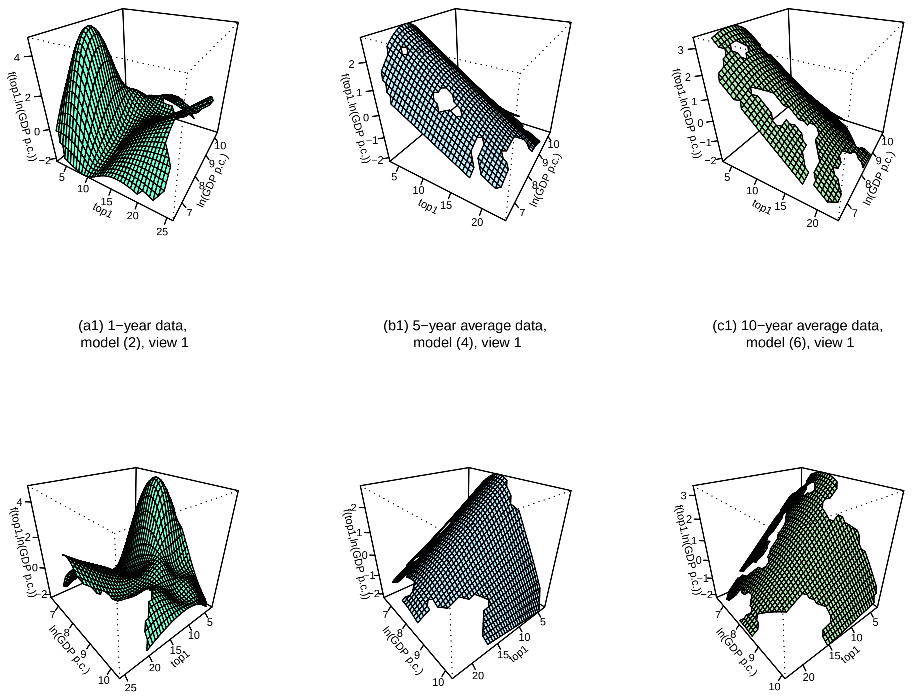
$$\alpha_i(z) = \sum_{l=1}^{L} \delta_{il} d_l(z),$$

and the tensor product basis construction gives:

$$f_{xz}(x, z) = \sum_{i=1}^{I} \sum_{l=1}^{L} \delta_{il} d_l(z) a_i(x).$$

The tensor product smooth has a penalty for each marginal basis. For further technical details, see Wood (2006).

Appendix D. Additional information, simplified models



(a1) 1−year data, model (2), view 1

(b1) 5−year average data, model (4), view 1

(c1) 10−year average data, model (6), view 1

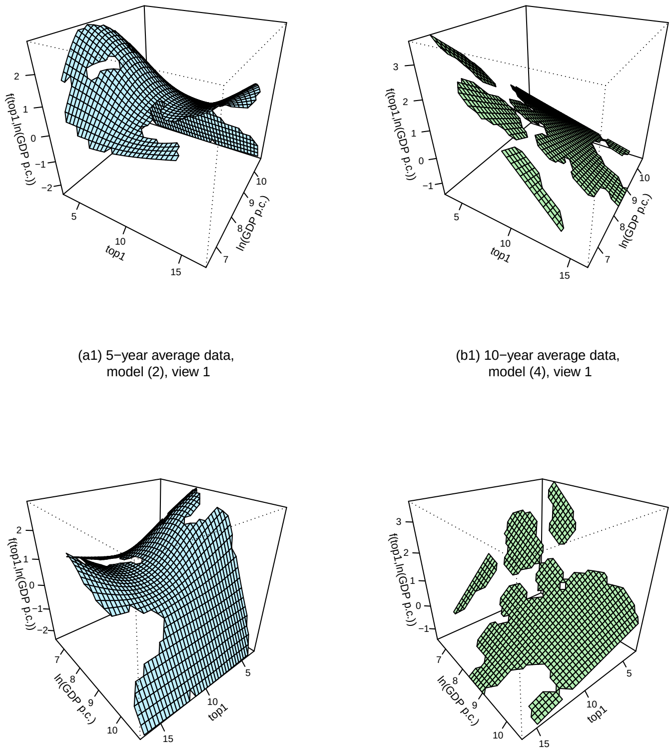
(a2) 1−year data, model (2), view 2

(b2) 5−year average data, model (4), view 2

(c2) 10−year average data, model (6), view 2

Figure D.6: Visualization of the simplified models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2), (4), and (6) of Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth f. Each smooth is illustrated from two views. In all plots, plot grid nodes that are too far from the true data points of the top 1% share and ln(GDP per capita) are excluded: the grid has been scaled into the unit square along with *top1* and GDP variables; grid nodes more than 0.1 from the predictor variables are excluded. Compare to Figure 2.

Appendix E. Additional information, expanded models



(a1) 5−year average data, model (2), view 1

(b1) 10−year average data, model (4), view 1

(a2) 5−year average data, model (2), view 2

(b2) 10−year average data, model (4), view 2

Figure E.7: Visualization of the expanded models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (4) of Table 3 (data from the 1950s onward; GDP data from PWT 7.0). The horizontal axes have the top 1% share and ln(GDP per capita); the vertical axis has the smooth f. The smooths are illustrated from two views. In all plots, plot grid nodes that are too far from the true data points of the top 1% share and ln(GDP per capita) are excluded: the grid has been scaled into the unit square along with *top1* and GDP variables; grid nodes more than 0.1 from the predictor variables are excluded. Compare to Figure 3.

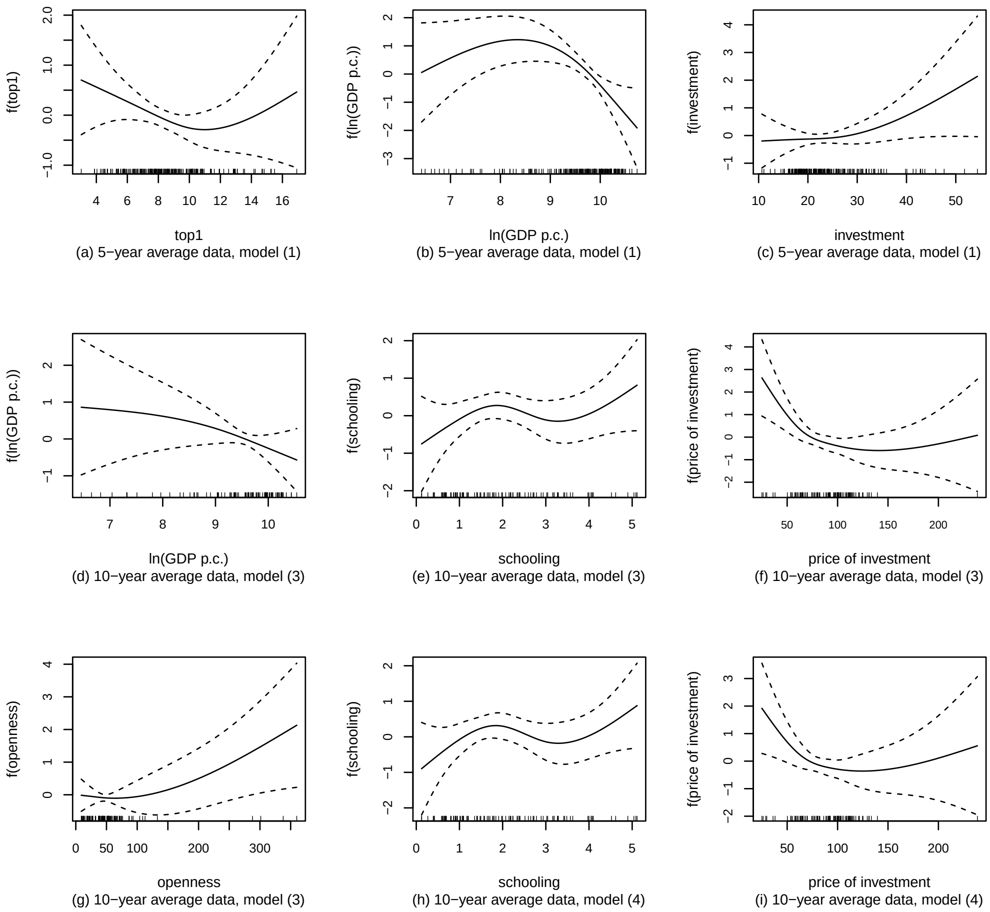


Figure E.8: Visualization of the expanded models: univariate smooths provided in Table 3 (data from the 1950s onward; GDP data from PWT 7.0). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as the dashed lines and the covariate values as a rug plot along the horizontal axis.

References

Aaberge, R., Atkinson, A.B., 2010. Top Incomes in Norway, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 448–481.

Aghion, P., Bolton, P., 1997. A Theory of Trickle-Down Growth and Development. Review of Economic Studies 64(2), 151–172.

Alvaredo, F., 2010a. The Rich in Argentina over the Twentieth century, 1932–2004, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 253–298.

Alvaredo, F., 2010b. Top Incomes and Earnings in Portugal, 1936–2005, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 560–624.

Alvaredo, F., Atkinson, A.B., 2010. Colonial Rule, Apartheid and Natural Resources: Top Incomes in South Africa 1903–2007. CEPR Discussion Paper, No. 8155.

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2011. The World Top Incomes Database. Website: `http://g-mond.parisschoolofeconomics.eu/topincomes`

Alvaredo, F., Pisano, E., 2010. Top Incomes in Italy, 1974–2004, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 625–663.

Alvaredo, F., Saez, E., 2010. Income and Wealth Concentration in Spain in a Historical and Fiscal Perspective, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 482–559.

Andrews, D., Jencks, C., Leigh, A., 2011. Do Rising Top Incomes Lift All Boats? B.E. Journal of Economic Analysis & Policy 11(1), Article 6.

Atkinson, A.B., 2007a. Measuring Top Incomes: Methodological Issues, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 18–42.

Atkinson, A.B., 2007b. The Distribution of Top Incomes in the United Kingdom 1908–2000, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 82–140.

Atkinson, A.B., 2010. Top Incomes in a Rapidly Growing Economy – Singapore, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 220–252.

Atkinson, A.B., Brandolini, A., 2001. Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study. Journal of Economic Literature 39(3), 771–799.

Atkinson, A.B., Leigh, A., 2007a. The Distribution of Top Incomes in Australia, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 309–332.

Atkinson, A.B., Leigh, A., 2007b. The Distribution of Top Incomes in New Zealand, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 333–364.

Atkinson, A.B., Piketty, T. (Eds.), 2007. Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T. (Eds.), 2010. Top Incomes: A Global Perspective. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T., Saez, E., 2010. Top Incomes in the Long Run of History, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 664–759.

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

Banerjee, A.V., Duflo, E., 2003. Inequality and Growth: What Can the Data Say? Journal of Economic Growth 8(3), 267–299.

Banerjee, A.V., Piketty, T., 2010. Top Indian Incomes, 1922–2000, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 1–39.

Barro, R.J., 2000. Inequality and Growth in a Panel of Countries. Journal of Economic Growth 5(1), 5–32.

Barro, R., Lee, J.–W., 2010. A New Data Set of Educational Attainment in the World, 1950–2010. NBER Working Paper No. 15902. Data source: `http://www.barrolee.com/`, Version 2.0, 07/10 (February 15, 2011).

Bénabou, R., 1996. Inequality and Growth, in: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual. The MIT Press, Cambridge. pp. 11–74.

Bénabou, R., 2000. Unequal Societies: Income Distribution and the Social Contract. American Economic Review 90(1), 96–129.

Chambers, D., Krause, A., 2010. Is the Relationship between Inequality and Growth Affected by Physical and Human Capital Accumulation? Journal of Economic Inequality 8(2), 153–172.

Deininger, K., Squire, L., 1996. A New Data Set Measuring Income Inequality. World Bank Economic Review 10(3), 565–591.

Dell, F., 2007. Top Incomes in Germany Throughout the Twentieth Century, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes Over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 365–425.

Dell, F., Piketty, T., Saez, E., 2007. Income and Wealth Concentration in Switzerland over the Twentieth Century, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes Over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 472–500.

Forbes, K.J., 2000. A Reassessment of the Relationship between Inequality and Growth. American Economic Review 90(4), 869–887.

Galor, O., Moav, O., 2004. From Physical to Human Capital Accumulation: Inequality and the Process of Development. Review of Economic Studies 71(4), 1001–1026.

Galor, O., Moav, O., Vollrath, D., 2009. Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. Review of Economic Studies 76(1), 143–179.

Galor, O., Zeira, J., 1993. Income Distribution and Macroeconomics. Review of Economic Studies 60(1), 35–52.

Halter, D., Oechslin, M., Zweimüller, J., 2014. Inequality and growth: the neglected time dimension. Journal of Economic Growth 19(1), 81–104.

Hastie, T., Tibshirani, R., 1986. Generalized additive models (with discussion). Statistical Science 1(3), 297–318.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, New York.

Henderson, D.J., Papageorgiou, C., Parmeter, C.F., 2012. Growth empirics without parameters. Economic Journal 122(559), 125–154.

Herzer, D., Vollmer, S., 2013. Rising top incomes do not raise the tide. Journal of Policy Modeling 35(4), 504–519.

Heston, A., Summers, R., Aten, B., 2011. Penn World Table Version 7.0. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. Version: March, 2011. Data source: `http://pwt.econ.upenn.edu/php_site/pwt_index.php` (April 20, 2011).

Jäntti, M., Riihelä, M., Sullström, R., Tuomala, M., 2010. Trends in Top Income Shares in Finland, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 371–447.

Kaldor, N., 1957. A Model of Economic Growth. Economic Journal 67(268), 591–624.

Kuznets, S., 1953. Shares of Upper Income Groups in Income and Saving. NBER Publication No. 55, New York.

Landais, C., 2008. "Top Incomes in France: booming inequalities?" Mimeo, Paris School of Economics, updated data: `http://www.jourdan.ens.fr/~clandais/index.php?langue=eng&choix=research` (August 31, 2010).

Leigh, A., 2007. How Closely Do Top Income Shares Track Other Measures of Inequality? Economic Journal 117(524), F589–F603.

Leigh, A., 2010. Top incomes in Australia, updated data: `http://people.anu.edu.au/andrew.leigh/pdf/TopIncomesAustralia.xls` (September 15, 2010).

Leigh, A., van der Eng, P., 2010. Top Incomes in Indonesia, 1920–2004, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 171–219

Li, H., Zou, H.–f., 1998. Income Inequality is not Harmful for Growth: Theory and Evidence. Review of Development Economics 2(3), 318–334.

Liu, Z., Stengos, T., 1999. Non-Linearities in Cross-Country Growth Regressions: A Semiparametric Approach. Journal of Applied Econometrics 14(5), 527–538.

Maasoumi, E., Racine, J., Stengos, T., 2007. Growth and convergence: A profile of distribution dynamics and mobility. Journal of Econometrics 136(2), 483–508.

Maddison, A., 2010. Statistics on World Population, GDP and Per Capita GDP, 1–2008 AD. Data on GDP per capita: `http://www.ggdc.net/MADDISON/Historical_Statistics/vertical-file_02-2010.xls` (November 12, 2010).

Moriguchi, C., Saez, E., 2010. The Evolution of Income Concentration in Japan, 1886–2005, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 76–170.

Nolan, B., 2007. Long-Term Trends in Top Income Shares in Ireland, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 501–530.

Perotti, R., 1993. Political Equilibrium, Income Distribution, and Growth. Review of Economic Studies 60(4), 755–776.

Perotti, R., 1996. Growth, Income Distribution, and Democracy: What the Data Say. Journal of Economic Growth 1(2), 149–187.

Piketty, T., 2001. Les Hauts revenus en France au 20e siècle: inégalités et redistribution, 1901–1998. B. Grasset, Paris.

Piketty, T., 2003. Income Inequality in France 1901–1998. Journal of Political Economy 111(5), 1004–1042.

Piketty, T., 2007. Income, Wage and Wealth Inequality in France 1901–98, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 43–81.

Piketty, T., Qian, N., 2010. Income Inequality and Progressive Income Taxation in China and India, 1986–2015, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 40–75.

Piketty, T., Saez, E., 2006. The Evolution of Top Incomes: A Historical and International Perspective. American Economic Review 96(2), 200–205.

Piketty, T., Saez, E., 2007. Income and Wage Inequality in the United States, 1913–2002, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 141–225.

Roine, J., Vlachos, J., Waldenström, D., 2009. The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data? Journal of Public Economics 93(7–8), 974–988.

Roine, J., Waldenström, D., 2010. Top Incomes in Sweden over the Twentieth Century, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes: A Global Perspective. Oxford University Press, Oxford. pp. 299–370.

Roine, J., Waldenström, D., 2015. Long-Run Trends in the Distribution of Income and Wealth, in: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution Vol. 2A. North-Holland, Amsterdam, pp. 469–592.

Saez, E., 2010. USA top incomes, updated data (revised 7/17/2010): http://elsa.berkeley.edu/~saez/TabFig2008.xls (September 15, 2010).

Saez, E., Veall, M.R., 2007. The Evolution of High Incomes in Canada, 1920–2000, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 226–308.

Salverda, W., Atkinson, A.B., 2007. Top Incomes in the Netherlands over the Twentieth Century, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 426–471.

Voitchovsky, S., 2005. Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution. Journal of Economic Growth 10(3), 273–296.

Voitchovsky, S., 2009. Inequality and Economic Growth, in: Salverda, W., Nolan, B., Smeeding, T. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 549–574.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton FL.

Essay II.

Changes or levels? Reassessment of the relationship between top-end inequality and growth

Elina Tuominen

Abstract

This study explores the association between top-end inequality and subsequent economic growth. The motivation stems from the results of Banerjee and Duflo (2003), who study nonlinearities in the inequality–growth relationship and find that changes in the Gini coefficient, in any direction, are associated with lower future growth. The current study addresses the issue of nonlinearity and exploits the top 1% income share series in 25 countries from the 1920s to the 2000s in various specifications. First, this study finds that the association between the level of top 1% share and growth is more evident in the data than the link between the change in top 1% share and growth. Second, the main results on the top 1% shares relate primarily to currently "advanced" economies; a negative association is discovered between the level of top-end inequality and growth, but this relationship is likely to become weaker in the course of economic development. Third, this study illustrates that the sample composition deserves attention in inequality–growth studies.

Keywords: inequality, top incomes, growth, nonlinearity, longitudinal data
JEL classification: O11, O15

Acknowledgments
Financial support from the Finnish Doctoral Programme in Economics (FDPE), the University of Tampere, and the Finnish Cultural Foundation is gratefully acknowledged. The author wishes to thank Olli Ropponen, Jari Vainiomäki, Hannu Tanninen, Jukka Pirttilä, and Matti Tuomala, as well as the participants at the FDPE Public Economics Workshop, the ECINEQ 2013 Conference, and the IIPF 2013 Congress for their comments and conversations. Remaining errors are the author's own.

1. Introduction

Empirical investigation of the relationship between inequality and economic growth has proven to be complex. For example, the diversity of the channels through which the effects may run makes causal inference difficult. Moreover, inequality data sets have suffered from quality issues. Further, the tradition of using linear specifications has been challenged. To address issues related to data and chosen functional forms, this study applies flexible methods to new data on top 1% income share series. Although top income shares best reflect the upper tail of the distribution, Leigh (2007) and Roine and Waldenström (2015) demonstrate that top income shares correlate with many other inequality measures. Thus, these data provide an interesting possibility of studying the inequality–growth association. Next, this section provides a short and selective review of the inequality–growth literature (see, e.g., Voitchovsky, 2009, for a more detailed discussion).

The theoretical literature describes contradictory channels from distribution to growth. According to the classical approach, the savings rate increases with income, and increased inequality may increase investment and thus also growth. Another argument for a positive inequality–growth link is based on incentives: income inequality encourages individuals to increase their effort, which enhances economic growth. In contrast, the imperfect credit market hypothesis describes a channel related to human capital accumulation (Galor & Zeira, 1993). According to this approach, higher inequality reduces growth because inequality reduces investment in human capital, assuming that credit constraints are binding.[1] One attempt to reconcile the conflicting classical and credit market imperfection channels is put forward by Galor and Moav (2004). In their unified growth theory, they argue that the classical channel is dominant in the early stages of development, and that the credit market imperfection channel becomes more important with development.[2] They also propose that both mechanisms fade in the course of development.

There are also many other arguments that inequality has adverse effects on economic performance. For example, Bénabou (2000) suggests that in-

[1] However, inequality might benefit investment in human capital in very poor economies. This is because it is possible that only the rich can invest in education. (Perotti, 1993)

[2] Galor and Moav (2004) propose that physical capital is the main engine of growth in the early stages of development, whereas human capital is the prime source of growth in the later stages of development.

equality may introduce an incentive for the rich to lobby against redistribution, and thus efficient policies may be prevented. Further, Leigh (2009) notes that the concentration of incomes at the top of the distribution can affect political and economic power and decision making.[3] Moreover, inequality may lead to sociopolitical instability, which hampers growth (Bénabou, 1996).

With improvement in the data sets, there has been a shift from cross-sectional to panel studies. In most empirical studies, inequality is measured in terms of the Gini coefficient, but the empirical evidence is mixed. In the 1990s, many cross-sectional studies found a negative relationship between inequality and growth (e.g., Bénabou, 1996; Perotti, 1996). Since then, some panel studies have reported a positive short- or medium-run relationship between inequality and subsequent growth (e.g., Li & Zou, 1998; Forbes, 2000). More recently, Halter et al. (2014) have found that the long-run (or total) association between inequality and growth is negative. It may be that the positive effects can be observed in the short run, but the negative effects take more time to materialize.[4] Furthermore, Barro (2000) suggests that in rich countries the association between inequality and growth is positive, whereas the relation is negative in poor countries. Voitchovsky (2005) exploits the panel features of the Luxembourg Income Study (LIS) data and finds that inequality is positively related to growth in the upper part of the distribution, whereas inequality is negatively associated with growth in the lower part of the distribution.[5]

Studies by Banerjee and Duflo (2003) and Chambers and Krause (2010) have allowed for nonlinearities. These studies also call into question earlier results of a positive association (e.g., Forbes, 2000). Banerjee and Duflo argue that nonlinearity may explain why the previously reported estimates vary greatly in the literature. They study the "high quality" subset of the Deininger and Squire (1996) data and find that *changes* in Gini, in any direc-

[3] Furthermore, Galor et al. (2009) suggest that inequality in the ownership of factors of production can incentivize the wealthy to impede institutional policies and changes that facilitate human capital formation and economic growth.

[4] Political processes, institutional changes, and educational attainment are involved in the channels that describe the negative effects of inequality on growth. It is likely that these mechanisms do not fully materialize in the short term.

[5] However, the inequality indices used by Voitchovsky (2005) do not describe the very top of the distribution.

tion, are associated with reduced subsequent growth—that is, they find an inverse U-shaped association with respect to changes in Gini.[6] In addition, Chambers and Krause find that inequality generally reduces growth in the subsequent 5-year period when they use Gini data from the World Income Inequality Database (WIID); the unified growth theory of Galor and Moav (2004) also gains some empirical support in their study. Thus, the linearity assumption may be too restrictive in modeling the relationship between inequality and growth, and for this reason, the current study applies penalized regression spline methods.

Inequality data sets have suffered from comparability issues over time and across countries (see, e.g., Atkinson & Brandolini, 2001). The recently published top income share series are of high quality compared to many other inequality data. Andrews et al. (2011) use an adjusted data set from Leigh (2007) to study the link between top incomes and growth. They exploit the top income shares of 12 wealthy countries and rely primarily on standard linear estimation methods, finding that after 1960, high inequality may enhance growth if inequality is measured by the top 10% income share. Recently, the conclusion related to the top 10% shares was challenged by Herzer and Vollmer (2013), who argue that the long-run effect of the top 10% share is the opposite. When Andrews et al. use the top 1% share as an inequality measure, many of their results are not statistically significant. Moreover, Andrews et al. report that their results are not in accordance with the inverse U-shaped association that Banerjee and Duflo (2003) find: when Andrews et al. study the relationship of changes in top incomes to growth, they cannot reject a linear association, but they admit that a nonlinear association is still possible.[7] The small number of countries in the study by Andrews et al. and possible nonlinearities in the relationship motivate the current paper.

The relationship between the *level* of top 1% income share and subsequent growth is discussed in a previous study by Tuominen (2015). The current study augments the preceding investigation by exploring the *change* in this

[6] This finding accords with a simple political economy model described by Banerjee and Duflo. However, Banerjee and Duflo (2003, p. 267) note that the inverse U relation "could also reflect the nature of measurement errors."

[7] Andrews et al. (2011, pp. 26–27) write: "...we cannot reject the hypothesis that changes in inequality have linear effects. [...] However, given the size of our standard errors we also cannot reject the existence of nonlinear effects large enough to be of considerable practical importance."

measure. Moreover, the current data include two additional countries compared to the preceding study. The top 1% income share series exploited in the current study describe top-end inequality in 25 countries from the 1920s to the 2000s. Models are fitted using different time-span specifications (data averaged over 5 and 10 years) to investigate the time dimension.

This study finds that future growth is more closely linked to the *level* of top 1% income share than to the *change* in this measure. In line with the preceding study, the association between the *level* of top 1% share and growth appears to depend on the country's level of economic development, and the main results relate primarily to currently "advanced" countries; various specifications show that a negative relationship between the level of top-end inequality and growth fades as the level of per capita GDP increases. However, this finding may not generalize to all kinds of economies—for example, tentative results for "less-advanced" economies provide reasons not to expect a similar relationship. Sensitivity checks illustrate that the sample composition deserves attention in inequality–growth studies.

The remainder of this study is organized in the following manner: Section 2 describes the data and section 3 introduces the estimation method. Section 4 provides the estimation results, including sensitivity checks. Finally, section 5 presents conclusions.

2. Data

Using tax and population statistics, it is possible to compose long series on top income shares. Kuznets (1953) was the first to use this kind of data to produce top income share estimates, and Piketty (2001, 2003) generalized Kuznets's approach. Following Piketty, different researchers have constructed top income share series using the same principles of calculation. Atkinson et al. (2011) provide an overview of the top income literature.[8] This study focuses on the top 1% (note that this is pre-tax income). The top 1% income shares (*top1*) in 25 countries from the 1920s to the 2000s

[8]In addition, for example, Atkinson (2007) provides information on the methodology. Piketty and Saez (2006), Leigh (2007), and Roine and Waldenström (2015) discuss the advantages and limitations of the top income share series. Detailed information on top income shares is published in two volumes edited by Atkinson and Piketty (2007, 2010), and the updated data are available in the World Top Incomes Database by Alvaredo et al. (2012). The top income project is ongoing.

are exploited, but the data set is not balanced. The data include, for example, English-speaking, Continental and Southern European, Nordic, and some "less-advanced" countries. A complete list of countries in the data and a graphical illustration of the top 1% series are provided in Appendix A.

The debate about how to choose control variables is put aside consciously because this study is not testing a specific channel from inequality to growth. The focus is on the overall association and nonlinearities. For this reason and due to data availability, two different approaches are taken in the empirical investigation. First, very long time series are studied in parsimonious (henceforth, "simplified") specifications that control only for the level of GDP per capita. Second, shorter time series are used in expanded specifications that include several additional controls. Naturally, the interpretation of the results is different in these two approaches because inequality may influence growth (at least in part) through some of the control variables.

Table 1: Descriptive statistics.

Simplified models (data from the 1920s onward)	N	min	mean	max
$top1_t$	275	3.9	9.6	23.4
$top1_t - top1_{t-1}$	275	-7.2	-0.2	4.6
$ln(GDP\ p.c.)_t$	275	6.4	8.9	10.3
$growth_{t+1}$	275	-15.2	2.4	16.1
Expanded models (data from the 1950s onward)	N	min	mean	max
$top1_t$	210	3.9	8.5	16.9
$top1_t - top1_{t-1}$	210	-6.9	0.0	3.4
$ln(GDP\ p.c.)_t$	210	6.4	9.5	10.7
$government\ consumption_t$	210	4.0	9.4	18.3
$investment_t$	210	10.6	24.0	54.4
$price\ level\ of\ investment_t$	210	18.9	87.0	294.6
$openness_t$	210	8.0	64.7	386.3
$secondary\ schooling_t$	210	0.1	2.2	5.4
$tertiary\ schooling_t$	210	0.0	0.3	1.7
$growth_{t+1}$	210	-3.1	2.4	9.5

Data averaged over 5-year periods are used in the calculations.
The 5-year periods t are defined as 1925–29, 1930–34, ..., and 2000–04.
Growth refers to average annual log growth; the change in top 1% income share refers to difference of average levels. More details are provided in footnotes 15 and 19.
Sources: see Appendix A for the top 1% shares and Appendix B for other variables.

The exceptionally long inequality series are exploited in the simplified specifications that use GDP per capita data (1920–2008) from Maddison (2010). In the expanded specifications, most of the data are from the Penn World Table version 7.0 (PWT 7.0) by Heston et al. (2011). The GDP per capita data span 1950–2009, and the other variables are those commonly used in growth regressions: government consumption, investment, price level of investment, and trade openness.[9] Furthermore, the expanded models include measures for human capital, namely, average years of secondary schooling and average years of tertiary schooling, the data of which are available every five years (Barro & Lee, 2010). More information on these variables is provided in Appendix B. Table 1 provides summary statistics with the 5-year average data.

3. Estimation method

Additive models provide a flexible framework for investigating the association between inequality and growth.[10,11] This study follows the approach presented in Wood (2006). The basic idea is that the model's predictor is a sum of linear and smooth functions of covariates:

$$\mathbb{E}(Y_i) = \boldsymbol{X}_i^* \boldsymbol{\theta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + ...$$

In the above presentation, Y_i is the response variable (here: average annual log growth in the subsequent period), $\boldsymbol{X}_i^*$ is a row of the model matrix for any strictly parametric model components, $\boldsymbol{\theta}$ is the corresponding parameter vector, and the $f_\bullet$ are smooth functions of the covariates, $x_\bullet$.

The flexibility of these models comes at the cost of two problems. First, one needs to represent the smooth functions $f_\bullet$ in some manner. One way to represent these functions is to use cubic regression splines, which is the approach adopted in this study. A cubic regression spline is a curve constructed from sections of cubic polynomials that are joined together so that the resulting curve is continuous up to the second derivative. The points at

[9] Price level of investment is a commonly used proxy for market distortions. Openness measure is defined as ratio of imports plus exports to GDP.

[10] Additive models are a special case of generalized additive models (GAMs). GAMs were introduced by Hastie and Tibshirani (1986, 1990). They present a GAM as a generalized linear model with a linear predictor that involves a sum of smooth functions of covariates. Some of the covariates can enter the model in linear form. Note here the analogy to "generalized linear models" and "linear models." This study is restricted to a special case: it uses an identity link and assumes normality in errors, which leads to additive models.

[11] In a study on determinants of top incomes shares, Roine et al. (2009) discuss the problems of using a long and narrow panel data set. For example, GMM procedures are not designed for settings with small number of countries and long series. Roine et al. run their regressions without instrumentation, which is also the approach here.

which sections are joined (and the end points) are the knots of the spline, and these locations must be chosen. The spline can be represented in terms of its values at the knots.[12] Second, the amount of smoothness that functions $f_\bullet$ will have needs to be chosen. Overfit is to be avoided and, thus, departure from smoothness is penalized. The appropriate degree of smoothness for $f_\bullet$ can be estimated from the data by, for example, maximum likelihood.

Illustration

Consider a model containing only one smooth function of one covariate: $y_i = f(x_i) + \epsilon_i$, where ϵ_i are i.i.d. $N(0, \sigma^2)$ random variables. To estimate function f here, f is represented so that the model becomes a linear model. This is possible by choosing a basis, defining the space of functions of which f (or a close approximation to it) is an element. In practice, one chooses basis functions, which are treated as known.

Assume that the function f has a representation $f(x) = \sum_{j=1}^{k} \beta_j b_j(x)$, where β_j are unknown parameters and $b_j(x)$ are known basis functions. Using a chosen basis for f implies that we have a linear model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$, where the model matrix $\boldsymbol{X}$ can be represented using basis functions such as those in the cubic regression spline basis. The departure from smoothness can be penalized with $\int f''(x)^2 dx$. The penalty $\int f''(x)^2 dx$ can be expressed as $\boldsymbol{\beta}^T \boldsymbol{S\beta}$, where $\boldsymbol{S}$ is a coefficient matrix that can be expressed in terms of the known basis functions.

Accordingly, the penalized regression spline fitting problem is to minimize $\|\boldsymbol{y} - \boldsymbol{X\beta}\|^2 + \lambda \boldsymbol{\beta}^T \boldsymbol{S\beta}$, with respect to $\boldsymbol{\beta}$. The problem of estimating the degree of smoothness is a problem of estimating the smoothing parameter λ.[13] The penalized least squares estimator of $\boldsymbol{\beta}$, given λ, is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{S})^{-1}\boldsymbol{X}^T\boldsymbol{y}$. Thus, the expected value vector is estimated as $\widehat{\mathbf{E}(\boldsymbol{y})} = \hat{\boldsymbol{\mu}} = \boldsymbol{Ay}$, where $\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{S})^{-1}\boldsymbol{X}^T$ is called an influence matrix.

This setting can be augmented to include several covariates and smooths. Given a basis, an additive model is simply a linear model with one or more associated penalties. Smooths of several variables can also be constructed.

[12] There are usually two extra conditions that specify that the second derivative of the curve should be zero at the two end knots.

[13] In the estimation, one faces a bias–variance tradeoff: on the one hand, the bias should be small, but on the other hand, the fit should be smooth. One needs to compromise between the two extremes. $\lambda \longrightarrow \infty$ results in a straight line estimate for f, and $\lambda = 0$ leads to an unpenalized regression spline estimate.

In this study, tensor product smooths are used in cases of smooths of two variables (Appendix C provides a short description).

Practical notes

The size of basis dimension for each smooth is usually not critical in estimation, because it only sets an upper limit on the flexibility of a term. Smoothing parameters control the effective degrees of freedom (*edf*). Effective degrees of freedom are defined as trace($\boldsymbol{A}$), where $\boldsymbol{A}$ is the influence matrix. The effective degrees of freedom can be used to measure the flexibility of a model. It is also possible to divide the effective degrees of freedom into degrees of freedom for each smooth. For example, a simple linear term would have one degree of freedom, and *edf*=2.3 can be thought of as a function that is slightly more complex than a second-degree polynomial.

Confidence (credible) intervals for the model terms can be derived using Bayesian methods, and approximate *p*-values for model terms can be calculated. Models can be compared using information criteria such as the Akaike information criterion (AIC). When using the AIC for penalized models (models including smooth terms), the degrees of freedom are the effective degrees of freedom, not the number of parameters. Moreover, random effects can be included in these models. For further details, see Wood (2006).[14]

4. Results

This section begins with the results of simplified models for very long series. Then, models with usual growth regression variables are reported using shorter series. The sensitivity checks and an additional example at the end of the section illustrate the importance of investigating the sample composition.

[14]The results presented in this study are obtained using the R software package "mgcv" (version 1.7-21), which includes a function "gam." Basis construction for cubic regression splines is used (the knots are placed evenly through the range of covariate values by default). The maximum likelihood method is used in the selection of the smoothing parameters. The identifiability constraints (due to, for example, the model's additive constant term) are taken into account by default. The function "gam" also allows for simple random effects: it represents the conventional random effects in a GAM as penalized regression terms. More details can be found in Wood (2006) and the R project's web pages (`http://cran.r-project.org/`).

4.1. Long series from the 1920s onward in simplified models

The simplified models include the level of top 1% income share, its change, and ln(GDP per capita) as covariates, and the dependent variable is the future log growth of GDP per capita; the GDP per capita data of Maddison (2010) are exploited. The relationship is investigated using both 5- and 10-year average data to assess whether the choice of period length affects the obtained results. The averaged data are used to mitigate the potential problems related to short-run disturbances.

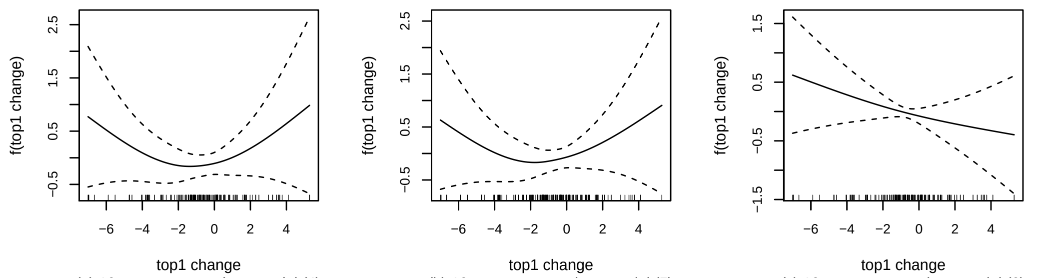
The models in Table 2 are of the form:

$$\begin{aligned}
growth_{i,t+1} &= \alpha + f_1(top1_{it}) + f_2(top1_{it} - top1_{i,t-1}) + f_3(ln(GDP\ p.c.)_{it}) \\
&\quad + \delta_{decade} + u_i + \epsilon_{it}, \\
growth_{i,t+1} &= \alpha + f_{13}(top1_{it}, ln(GDP\ p.c.)_{it}) + f_2(top1_{it} - top1_{i,t-1}) \\
&\quad + \delta_{decade} + u_i + \epsilon_{it}, \quad \text{and} \\
growth_{i,t+1} &= \alpha + f_2(top1_{it} - top1_{i,t-1}) + f_3(ln(GDP\ p.c.)_{it}) \\
&\quad + \delta_{decade} + u_i + \epsilon_{it},
\end{aligned}$$

where i refers to a country and t to a time period, α is a constant, functions $f_\bullet$ refer to smooth functions, δ_{decade} refers to a fixed decade effect (one decade is the reference category), u_i refers to a country-specific random effect ($u_i \sim \mathrm{N}(0, \sigma_u^2)$), and $\epsilon_{it} \sim \mathrm{N}(0, \sigma^2)$ is the error term; inequality and GDP per capita variables are used as period averages.[15] The random-effect spec-

[15] In annual data, growth would refer to the difference of $ln(GDP\ p.c.)$ values at $t+1$ and t multiplied by 100. This idea is also behind the averaged data. In the 5-year average data, the time periods t are 1925–29, 1930–34, ..., 2000–04. For example, the averages of the covariates in 1925–29 (period t) are used with the subsequent period's ($t+1$) average annual log growth (calculated using $ln(GDP\ p.c.)$ values in 1930–35), and the change in *top1* is the difference of the averages in 1925–29 (period t) and 1920–24 (period $t-1$). Then, the same logic applies to the period 1930–34 when it is considered as period t, and so on. The only exception is the future growth for the last 5-year period (2000–04): average growth is calculated using $ln(GDP\ p.c.)$ values in 2005–08 (i.e., $growth_{t+1}$ is based on three, not five, annual growth rates due to data unavailability in Maddison, 2010). Similarly, in the 10-year average data, the periods t are 1930–39, 1940–49, ..., 1990–99. The only exception to the logic is the future growth for the last 10-year period (1990–99): average growth is calculated using $ln(GDP\ p.c.)$ values in 2000–08 (i.e., $growth_{t+1}$ is not an average of ten annual growth rates but eight). Thus, the data points of the dependent and the explanatory variables do not overlap in the estimation equation. This should rule out direct reverse causation and reduce the endogeneity problem related to using a (lagged) GDP variable as a regressor.

ification allows for correlation over time within countries, and the results reflect both cross-sectional differences across countries and variations over time within countries. The random-effect approach is also used by Banerjee and Duflo (2003), who motivate the current study.[16] The second specification with a bivariate smooth $f(top1_t, ln(GDP\ p.c.)_t)$ allows for a very flexible interaction between the level of top-end inequality and the level of economic development—the specification stems from Tuominen (2015). The third specification checks the results when the level of top 1% share is excluded. In Table 2, a linear term is reported when linearity was suggested (that is, smooth's effective degrees of freedom were equal to one) in the estimation.



(a) 10−year average data, model (4) (b) 10−year average data, model (5) (c) 10−year average data, model (6)

Figure 1: Visualization of the simplified models: smooths $f(top1_t - top1_{t-1})$ provided in Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as dashed lines and the covariate values as a rug plot along the horizontal axis.

Table 2 demonstrates that the *change* in top-end inequality (i.e., $f(top1_t - top1_{t-1})$) is not statistically significantly related to subsequent growth. In the 10-year data, the shape of this smooth may even resemble a U (see Figure 1), which is opposite to what Banerjee and Duflo (2003) report with Gini data. Models (1) and (4) of Table 2 suggest that the *level* of top-end inequality is

[16]Barro (2000) points out that differencing in the fixed-effects approach exacerbates the measurement error problem, especially for an inequality variable, for which the variation across countries is important. He prefers using random effects. Moreover, Banerjee and Duflo (2003) state that there are no strong grounds for believing that the omitted variable problem could be solved by adding a fixed effect for each country, especially in a linear specification (as in, e.g., Forbes, 2000).

Table 2: Simplified models for 25 countries (data from the 1920s onward; GDP data from Maddison, 2010): the effective degrees of freedom for each smooth and the coefficients for the linear terms. The dependent variable is the average annual log growth in the next period, where one period is 5 or 10 years. See also Figure 1 and Figure D.8 for the univariate smooths $f(top1_t - top1_{t-1})$ and $f(ln(GDP\ p.c.)_t)$, respectively. The bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$ of models (2) and (5) are illustrated in Figure 2.

	5-year average data (N=275)			10-year average data (N=125)		
	(1)	(2)	(3)	(4)	(5)	(6)
$f(top1_t)$	[linear[a]] -0.146**	-	-	[linear[a]] -0.197**	-	-
$f(top1_t - top1_{t-1})$	[linear[a]] 0.145	[linear[a]] 0.135	[linear[a]] 0.070	[$edf \approx 2.0$[a]]	[$edf \approx 1.9$[a]]	[$edf \approx 1.2$[a]]
$f(ln(GDP\ p.c.)_t)$	[$edf \approx 2.5$[a]]***	-	[$edf \approx 2.5$[a]]***	[$edf \approx 2.6$[a]]***	-	[$edf \approx 2.6$[a]]***
$f(top1_t, ln(GDP\ p.c.)_t)$	-	[$edf \approx 5.1$[b]]***	-	-	[$edf \approx 3.8$[b]]***	-
adjusted r^2	0.15	0.15	0.14	0.45	0.45	0.44
AIC	1325	1327	1329	455	455	456

***, **, *, ' indicate significance at the 1, 5, 10, and 15% levels, respectively.
The p-values for parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are reported. The smooth terms' significance levels are based on approximate p-values.
All specifications include decade dummies and random country-specific effects.

[a]The basis dimension k for the smooth before imposing identifiability constraints is $k = 5$.
[b]The basis dimension k for the smooth before imposing identifiability constraints is $k = 5^2 = 25$ (tensor product smooth using rank 5 marginals).

negatively and statistically significantly associated with growth.[17] Further, Figure 2 illustrates the bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (5): plots (a1)–(a2) and (b1)–(b2) show a negative relationship between the *level* of top-end inequality and growth, but this link becomes weaker with development; the negative slope with respect to *top1* becomes less steep as GDP per capita increases. Additional plots of the bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$ are provided in Figure D.7 in Appendix D.

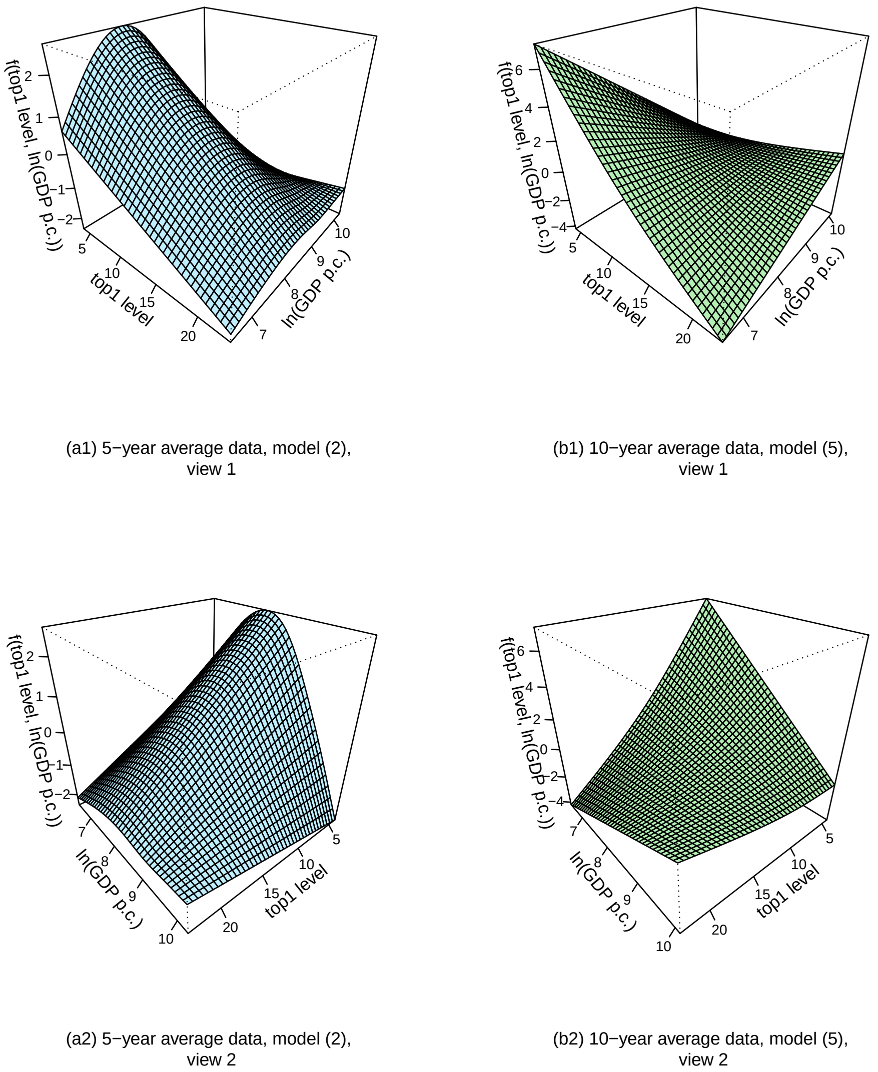
In the current sample, 18 out of the 25 countries are "advanced," and the other countries comprise a heterogeneous group. As a small check, these "advanced" countries were studied separately to see whether the other seven countries affected the main results above. Specifications similar to models (1)–(2) and (4)–(5) of Table 2 were fitted for this subset of the data.[18] The main conclusions about the relationship between the top 1% share and subsequent growth were not affected when the analysis was limited to these 18 countries.

In summary, the *level* of top 1% share appears to be more closely related to growth than the *change* in this measure. The discovered "negative but fading" association may reflect many channels from distribution to growth, but discussing this further would be more or less speculation. Moreover, the data include the Great Depression of the 1930s and the years of World War II, which may affect the findings. The next subsections focus on data from the 1950s onward.

---

[17] In model (1) of Table 2, the coefficient for the linear term $top1_t - top1_{t-1}$ is not significant. However, when the linear terms are written out, the model gives $-0.146 top1_t + 0.145(top1_t - top1_{t-1}) \approx -0.145 top1_{t-1}$. This would favor investigating a longer-run association between top-end inequality and growth, although only the coefficient -0.146 for $top1_t$ is significant. The result appears reasonable in the 5-year data because income distribution (usually) changes fairly slowly. Variables $top1_t$ and $top1_{t-1}$ are likely to reflect very similar information. As a check, a model with two smooths $f(top1_t)$ and $f(top1_{t-1})$ was estimated. In this case, linear terms were suggested, and the corresponding coefficients for $top1_t$ and $top1_{t-1}$ were in line with what model (1) gives when the linear terms are written out; the coefficients were not significant in this specification.

[18] Australia, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States (N=212 in the 5-year data; N=96 in the 10-year data).



(a1) 5−year average data, model (2), view 1

(b1) 10−year average data, model (5), view 1

(a2) 5−year average data, model (2), view 2

(b2) 10−year average data, model (5), view 2

Figure 2: Visualization of the simplified models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (5) of Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Both smooths are illustrated from two views. The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth f. For additional illustrations, see Figure D.7 in Appendix D.

4.2. Series from the 1950s onward in expanded models

The models are expanded with usual growth regression variables in this subsection. Again, data averaged over 5- and 10-year periods are investigated because the medium- and long-term associations are of interest. In this subsection, the GDP per capita series are from PWT 7.0 by Heston et al. (2011).[19] Before estimating the expanded specifications, the findings that are provided next were checked to ensure that they were not driven by the shorter time series and the change of the GDP data source.[20]

4.2.1. Whole-sample results

Results for three types of specifications are provided in Table 3. Models (1) and (4) are of the form:

$$\begin{aligned} growth_{i,t+1} = \alpha &+ f_1(top1_{it}) + f_2(top1_{it} - top1_{i,t-1}) + f_3(ln(GDP\ p.c.)_{it}) \\ &+ f_4(gov't\ consumption_{it}) + f_5(price\ level\ of\ investment_{it}) \\ &+ f_6(openness_{it}) + f_7(investment_{it}) + f_8(sec.\ schooling_{it}) \\ &+ f_9(tert.\ schooling_{it}) + \delta_{decade} + u_i + \epsilon_{it}, \end{aligned}$$

where i refers to a country and t to a time period, α is a constant, functions $f_\bullet$ refer to smooth functions, δ_{decade} refers to a fixed decade effect (one decade is the reference category), u_i is a country-specific random effect, and ϵ_{it} is the conventional error term; variable values are period averages. In comparison, models (2) and (5) include a bivariate smooth $f_{13}(top1_t, ln(GDP\ p.c.)_t)$

[19]The averaged data are constructed in a similar manner as in the case of longer series (see footnote 15). In the 5-year average data, the periods t are 1950–54, 1955–59, ..., 2000–04. For example, the averages of covariates in 1950–54 (period t) are used with the next period's ($t+1$) average annual log growth (calculated using $ln(GDP\ p.c.)$ values in 1955–60), and the change in *top1* variable is the difference of averages in 1950–54 (period t) and 1945–49 (period $t-1$). Then again, the same logic applies to the period 1955–59 when it is considered as period t. The only exception is the future growth for the last 5-year period (2000–04): average growth is calculated using $ln(GDP\ p.c.)$ values in 2005–09 (i.e., $growth_{t+1}$ is based on four, not five, annual growth rates due to data unavailability in PWT 7.0 Heston et al., 2011). Correspondingly, in the 10-year average data, the periods t are 1950–59, 1960–69, ..., 1990–99. The only exception to the logic is the future growth for the last 10-year period (1990–99): $growth_{t+1}$ is based on $ln(GDP\ p.c.)$ values in 2000–09 (i.e., it is not an average of ten annual growth rates but nine).

[20]Simplified specifications that resemble models (1)–(2) and (4)–(5) of Table 2 were estimated with the shorter $ln(GDP\ p.c.)$ series from the PWT 7.0 data. The results were qualitatively similar to those in subsection 4.1. For brevity, the details are not reported.

instead of smooths $f_1(top1_t)$ and $f_3(ln(GDP\ p.c.)_t)$; models (3) and (6) do not include the level of top 1% income share. As in the previous subsection, linear terms are reported only if the smooth's effective degrees of freedom were equal to one during the initial stage of the model fitting.

The models in Table 3 do not support an inverted U relationship between the *change* in top-end inequality and subsequent growth: the (positive) association is not statistically significant in any of the specifications (1)–(6), whereas the *level* of top 1% share appears to be relevant. The negative coefficient for the linear $top1_t$ term in the 10-year data is statistically significant in model (4).[21] Furthermore, models (2) and (5) include bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$ that are illustrated in Figure 3. In plots (a1)–(a2), the 5-year data show a positive or U-shaped *top1*–growth relation at "low" or "medium" levels of ln(GDP per capita); however, the association between the level of top 1% share and growth fades away at "high" levels of GDP per capita. Plots (b1)–(b2) show that in the 10-year data, the association is more straightforward: a negative slope is found with respect to *top1*, but this slope becomes less steep as the level of per capita GDP increases (see also note c to Table 3).

The findings indicate that top-end inequality and growth are related despite adding various control variables. The results on the level of top 1% share are qualitatively in line with the findings of Tuominen (2015). Moreover, the results in Table 3 show that government consumption and openness are positively related to future growth. Secondary education is also significant in most models.[22]

In summary, the results support a distribution–growth relationship that is found with respect to the *level* of (not *change* in) top-end inequality, and this association may evolve during the development process. In the 10-year data, the main results on top-end inequality are similar to those in subsection 4.1. In comparison, in the 5-year data, the results appear to be affected by the inclusion of additional covariates, and a U shape appears in plots (a1)–(a2) of

[21] In models (1) and (4) of Table 3, both terms $f(top1_t)$ and $f(top1_t - top1_{t-1})$ are linear. However, negative coefficients are obtained for $top1_t$ and $top1_{t-1}$ if the linear terms are written out in these two models. For example, model (1) gives $-0.065top1_t + 0.048(top1_t - top1_{t-1}) = -0.017top1_t - 0.048top1_{t-1}$. Thus, these specifications do not indicate a positive association between the level of *top1* and subsequent growth.

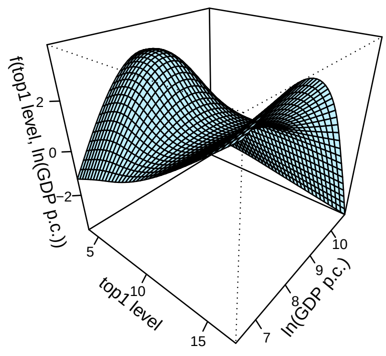
[22] Figure E.10 in Appendix E reveals that secondary schooling correlates positively with future growth in countries where the level of education is very low.

Table 3: Expanded models for 25 countries (data from the 1950s onward; GDP data from PWT 7.0): the effective degrees of freedom for each smooth and the coefficients for the linear terms. The dependent variable is the average annual log growth in the next period, where one period is 5 or 10 years. See Figure 3 for illustrations of the bivariate smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (5) and Figure E.10 in Appendix E for illustrations of the univariate smooths with $edf > 1$.

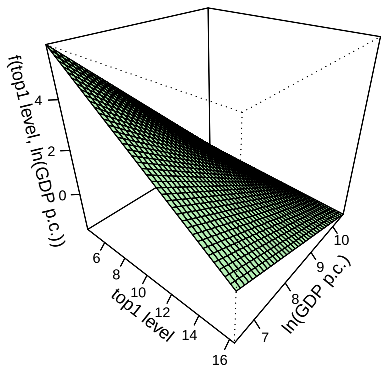
	5-year average data (N=210)			10-year average data (N=95)		
	(1)	(2)	(3)	(4)	(5)	(6)
$f(top1_t)$	[linear[a]] -0.065	-	-	[linear[a]] -0.161**	-	-
$f(top1_t - top1_{t-1})$	[linear[a]] 0.048	[linear[a]] 0.076	[linear[a]] 0.017	[linear[a]] 0.133	[linear[a]] 0.130	[linear[a]] 0.048
$f(ln(GDP\ p.c.)_t)$	[$edf \approx 2.6$[a]]***	-	[$edf \approx 2.6$[a]]***	[linear[a]] -1.274***	-	[$edf \approx 2.0$[a]]***
$f(top1_t, ln(GDP\ p.c.)_t)$	-	[$edf \approx 8.8$[b]]***	-	-	[$edf \approx 3.0$[b,c]]***	-
$f(government\ consumption_t)$	[linear[a]] 0.180***	[linear[a]] 0.187***	[linear[a]] 0.193***	[linear[a]] 0.256***	[linear[a]] 0.234***	[linear[a]] 0.335***
$f(price\ level\ of\ investment_t)$	[linear[a]] -0.006	[linear[a]] -0.012*	[linear[a]] -0.007	[linear[a]] 0.000	[linear[a]] 0.003	[$edf \approx 1.3$[a]]
$f(openness_t)$	[linear[a]] 0.008**	[linear[a]] 0.008**	[linear[a]] 0.008**	[linear[a]] 0.005'	[linear[a]] 0.006*	[linear[a]] 0.008*
$f(investment_t)$	[linear[a]] -0.004	[linear[a]] -0.011	[linear[a]] -0.008	[linear[a]] 0.085***	[linear[a]] 0.079***	[linear[a]] 0.050
$f(secondary\ schooling_t)$	[$edf \approx 3.0$[a]]**	[$edf \approx 2.7$[a]]	[$edf \approx 3.0$[a]]**	[$edf \approx 3.4$[a]]**	[$edf \approx 3.3$[a]]*	[$edf \approx 3.1$[a]]**
$f(tertiary\ schooling_t)$	[linear[a]] 0.930	[linear[a]] 1.105	[linear[a]] 0.810	[linear[a]] 1.353	[linear[a]] 1.232	[linear[a]] 1.386
adjusted r^2	0.46	0.48	0.47	0.53	0.53	0.64
AIC	784	778	781	325	325	306

***, **, *, ' indicate significance at the 1, 5, 10, and 15% levels, respectively.
The p-values for the parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are provided. The smooth terms' significance levels are based on approximate p-values.
All specifications include decade dummies and random country-specific effects.

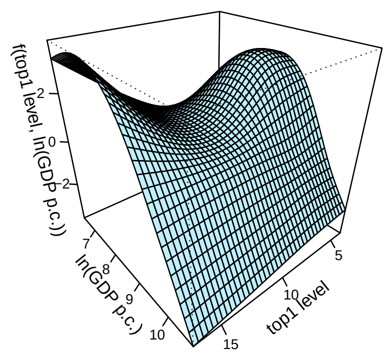
[a] The basis dimension k for the smooth before imposing identifiability constraints is $k = 5$.
[b] The basis dimension k for the smooth before imposing identifiability constraints is $k = 5^2 = 25$ (tensor product smooth using rank 5 marginals).
[c] With just 3 degrees of freedom, the tensor product smooth refers to $\theta_1 top1_t + \theta_2 ln(GDP\ p.c.)_t + \theta_3 top1_t ln(GDP\ p.c.)_t$, where $\theta_\bullet$ are coefficients. When model (5) is estimated using this form in place of $f(top1_t, ln(GDP\ p.c.)_t)$, the coefficients are $\hat{\theta}_1 = -1.062$*, $\hat{\theta}_2 = -2.134$***, and $\hat{\theta}_3 = 0.096$'. For example, if $GDP\ p.c.$ is 8100 (2005 I\$), then $ln(GDP\ p.c.) \approx 9$, and the slope with respect to $top1$ is approximately -0.20. Correspondingly, if $GDP\ p.c.$ is 22000 (2005 I\$), then $ln(GDP\ p.c.) \approx 10$, and the slope is approximately -0.10. Plots (b1)–(b2) of Figure 3 illustrate this change in the slope.



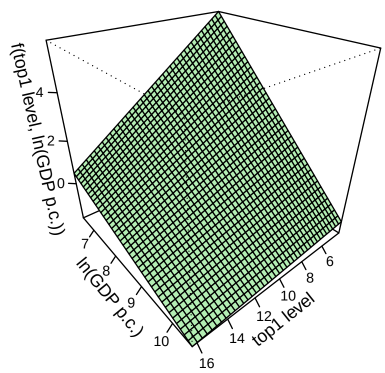
(a1) 5−year average data, model (2),
view 1



(b1) 10−year average data, model (5),
view 1



(a2) 5−year average data, model (2),
view 2



(b2) 10−year average data, model (5),
view 2

Figure 3: Visualization of the expanded models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (5) of Table 3 (data from the 1950s onward; GDP data from PWT 7.0). Both smooths are illustrated from two views. The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth f. For additional illustrations, see Figure E.9 in Appendix E.

Figure 3 at "medium" levels of economic development (see also footnote 20). The next subsection investigates the data further by taking into account that the sample is composed of different types of countries.

4.2.2. Sample composition: different types of countries

This subsection focuses on the 5-year average data because the corresponding subsets of the 10-year average data would be very small. To be more specific, data from the 1950s onward were exploited in specifications similar to models (1) and (2) of Table 3 for different groups of countries.[23] Although the results were not statistically significant at the 10% level for all groups of countries, the findings help in understanding the whole-sample patterns.

The Continental and Southern European countries showed a negative link between the level of top-end inequality and growth, but this association was not statistically significant; a negative association was discovered between the change in top-end inequality and growth. For the Nordic countries, neither the level of *top1* nor the change in *top1* were statistically significantly related to growth. For the English-speaking countries, a negative (or slightly inverse U-shaped) association between the level of *top1* and growth was discovered; the relationship between the change in *top1* and growth was not statistically significant. In comparison, data on the small and very diverse group of "less-advanced" countries showed a positive relationship between the level of top-end inequality and subsequent growth; the association between the change in top-end inequality and growth was inverse U-shaped, but it was not statistically significant.[24]

These results help explain the shape of the smooth $f(top1_t, ln(GDP\ p.c.)_t)$ in plots (a1) and (a2) of Figure 3. The U shape at "medium" levels of economic development appears to reflect a combination of different types of

[23]English-speaking: Australia, Canada, Ireland, New Zealand, the United Kingdom, and the United States (N=60). Continental and Southern European: Germany, France, Italy, the Netherlands, Portugal, Spain, and Switzerland (N=52). Nordic: Denmark, Finland, Norway, and Sweden (N=37). "Less-advanced:" Argentina, China, India, Indonesia, Mauritius, and South Africa (N=41). Note that Japan (N=11) and Singapore (N=9) are difficult to fit into these categories.

[24]Furthermore, results for the "less-advanced" countries indicated that secondary schooling and government consumption are positively (and statistically significantly) related to subsequent growth. These countries appear to have the greatest influence on the results with respect to schooling and government consumption at the whole-sample level.

countries: the relationship between the level of *top1* and growth may be different in "less-advanced" and "advanced" countries (at least when 5-year periods are studied). This finding is in accordance with Tuominen (2015), but a larger sample would be required to be able to discuss this further. In conclusion, the result of a positive association of top incomes to growth in "less-advanced" countries should be taken very cautiously due to sparse data. Thus, the main conclusions are drawn for currently "advanced" countries.

Finally, the group of 18 "advanced" countries was studied separately. These countries demonstrated that the negative relationship between the *level* of top-end inequality and growth is weak (or no longer significant) at "high" levels of economic development.[25] The "fading association" may explain why Andrews et al. (2011) do not find significant results on top 1% shares in 12 wealthy countries. Andrews et al. also report that their results on changes in top incomes are not in line with the inverse U result of Banerjee and Duflo (2003). The currently studied group of 18 "advanced" countries did not show a statistically significant pattern between the *change* in top 1% share and future growth. However, this "non-result" for changes in top-end inequality may be a consequence of many things. For example, the current sample may be too focused on wealthy countries (compared to the sample used by Banerjee and Duflo, 2003), or the top-income measure may miss something that Gini coefficients capture. This reasoning motivated an additional investigation that is discussed in the next subsection.

4.2.3. Example: fewer countries, shorter series, and Gini coefficients

Different parts of the distribution may be differently related to growth (see, e.g., Voitchovsky, 2005). For this reason, this subsection provides an example of expanding the estimated models with the Gini coefficients used by Forbes (2000) and Banerjee and Duflo (2003). They use observations from the "high quality" sample of the Deininger and Squire (1996) data on approximately 5-year intervals, and their sample includes 45 countries, of which 21 appear also in the current study.[26] However, different timing of the

[25] This group included Japan and the English-speaking, Continental and Southern European, and Nordic countries. This group of countries was also checked with the 10-year data, and the results for top-end inequality were qualitatively similar to those with the 5-year data.

[26] Because the results by Banerjee and Duflo (2003) motivate the current study, the same Gini source is of interest. Data quality issues are beyond the scope of the current study.

available observations in the data limits the countries to 18, of which almost all are "advanced" economies. The data span approximately 30 years but are not balanced. Appendix B provides details.

Table 4: Models with Gini coefficients for 18 countries (GDP data from PWT 7.0): the effective degrees of freedom for each smooth and the coefficients for the linear terms. The dependent variable is the average annual log growth in the subsequent period, where one period is 5 years. See Appendix B for more information on the Gini data and period definitions. Figure 4 provides illustrations of the bivariate smooth $f(top1_t, ln(GDP\ p.c.)_t)$ in models (3) and (4).

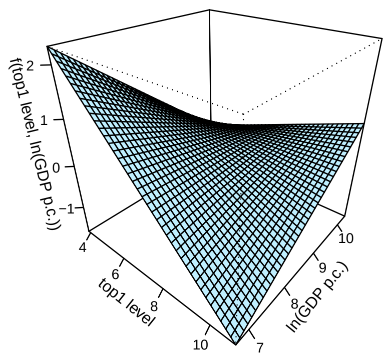
	5-year average data (N=62)			
	(1)	(2)	(3)	(4)
$f(top1_t)$	[linear[a]] 0.005	-	-	-
$f(top1_t - top1_{t-1})$	[linear[a]] -0.183	[$edf \approx 1.1$[a]]	[$edf \approx 1.3$[a]]	[linear[a]] -0.133
$f(ln(GDP\ p.c.)_t)$	[linear[a]] 0.446*	[$edf \approx 1.3$[a]]	-	-
$f(top1_t, ln(GDP\ p.c.)_t)$	-	-	[$edf \approx 3.0$[b,c]]*	[$edf \approx 3.0$[b,d]]*
$f(Gini_t)$	[linear[a]] 0.080**	-	-	[linear[a]] 0.067*
$f(Gini_t - Gini_{t-1})$	[linear[a]] 0.067	[linear[a]] 0.116**	[linear[a]] 0.124**	[linear[a]] 0.075'

***, **, *, ' indicate significance at the 1, 5, 10, and 15% levels, respectively.
The p-values for parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are reported.
The smooth terms' significance levels are based on approximate p-values.

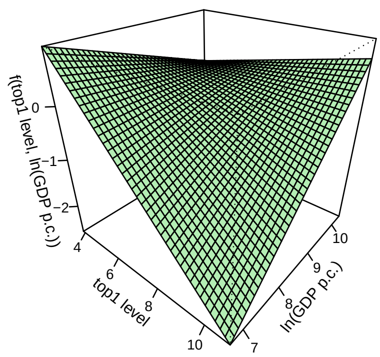
Note: All models include decade dummies, random country effects, and controls for government consumption, price level of investment, openness, investment, average years of secondary schooling, and average years of tertiary schooling (almost all controls enter the models linearly).

[a] The basis dimension k for the smooth before imposing identifiability constraints is $k = 3$.
[b] The basis dimension k for the smooth before imposing identifiability constraints is $k = 3^2 = 9$ (tensor product smooth using rank 3 marginals).
[c] With only 3 degrees of freedom, the tensor product smooth refers to $\theta_1 top1_t + \theta_2 ln(GDP\ p.c.)_t + \theta_3 top1_t ln(GDP\ p.c.)_t$, where $\theta_\bullet$ are coefficients. When model (3) is estimated using this form in place of $f(top1_t, ln(GDP\ p.c.)_t)$, the coefficients are $\hat{\theta}_1 = -1.928$**, $\hat{\theta}_2 = -1.609$**, and $\hat{\theta}_3 = 0.205$**. For example, if $GDP\ p.c.$ is 8100 (2005 \$I), then $ln(GDP\ p.c.) \approx 9$, and the slope with respect to $top1$ is approximately -0.08; if $GDP\ p.c.$ is 22000 (2005 \$I), then $ln(GDP\ p.c.) \approx 10$, and the slope is approximately 0.12. Plots (a1)–(a2) in Figure 4 illustrate this change in the slope.
[d] With just 3 degrees of freedom, the tensor product smooth refers to $\theta_1 top1_t + \theta_2 ln(GDP\ p.c.)_t + \theta_3 top1_t ln(GDP\ p.c.)_t$, where $\theta_\bullet$ are coefficients. When model (4) is estimated using this form in place of $f(top1_t, ln(GDP\ p.c.)_t)$, the coefficients are $\hat{\theta}_1 = -1.621$**, $\hat{\theta}_2 = -0.944$, and $\hat{\theta}_3 = 0.167$**. For example, if $ln(GDP\ p.c.) = 9$, the slope with respect to $top1$ is approximately -0.12; if $ln(GDP\ p.c.) = 10$, the slope is approximately 0.05. This change in the slope is illustrated in plots (b1)–(b2) of Figure 4.

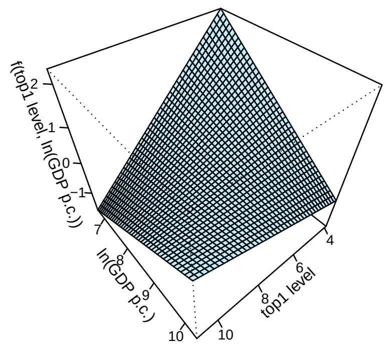
Table 4 provides the results of models with Gini coefficients for 18 countries. Linear terms were suggested for most covariates. In accordance with earlier findings, the *change* in top 1% share is not statistically significantly related to future growth. Moreover, Figure 4 illustrates the smooth func-



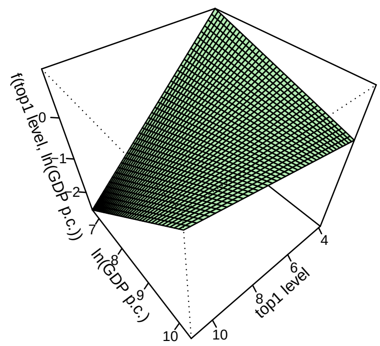
(a1) model (3), view 1



(b1) model (4), view 1



(a2) model (3), view 2



(b2) model (4), view 2

Figure 4: Visualizations of the smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (3) and (4) of Table 4. Both smooths are illustrated from two views. The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth f.

tions $f(top1_t, ln(GDP\ p.c.)_t)$ of models (3) and (4) in plots (a1)–(a2) and (b1)–(b2), respectively. These plots show a negative association between the *level* of top-end inequality and subsequent growth, and this relation fades as the level of GDP per capita increases; thus, the overall shape of the smooths appears to be in line with the previous results. However, a more detailed investigation reveals that India and Indonesia cause the negative association between the level of top 1% share and growth at "low" levels of economic development ($ln(GDP\ p.c.) < 8$, in this case). The other 16 countries in this subset have higher per capita GDP, and, in keeping with previous findings, the relationship is not very clear at these levels of per capita GDP. The link between the level of top 1% share and growth is close to zero (maybe even starting to turn positive) at "high" levels of development; see also notes c and d to Table 4.[27]

The small sample size provides a good reason for being cautious about the results in Table 4, but the findings suggest that the Gini coefficients and the top 1% income shares may be differently related to growth. The results also indicate that more data are needed to establish the inverted U result with respect to *changes* in the Gini coefficient.[28] However, these findings should be checked in later studies when more data are available. The current study does not speculate further on the results in Table 4 for this reason. Using alternative Gini data sets with the top income shares would also be interesting, but this is left for future studies. However, these findings, combined with the previous subsection's checks, illustrate why it is reasonable to investigate different subsets of the data that may represent different types of countries.

5. Conclusions

Banerjee and Duflo (2003) suggest that changes in the Gini coefficient, in any direction, are related to lower future growth. The current study

[27]As a further check, India and Indonesia (six observations in total) were excluded from the analysis: the remaining 16 wealthy countries (all had $ln(GDP\ p.c.) > 9$) showed that the *top1*–growth association is not significant at the 10% level, and this is in line with previous findings related to "high" levels of economic development. The results on the Gini coefficients were qualitatively similar to those reported in Table 4.

[28]In the sample used by Banerjee and Duflo, the largest changes in the Gini coefficients took place in countries that are not in the currently studied subset of the data. See Table 2 in Banerjee and Duflo (2003, p. 282).

investigates the association between the change in inequality and growth, but a different inequality measure is used. However, due to data unavailability, the current study is more focused on "advanced" countries, although some "less-advanced" countries are included. This study finds that future growth is more closely related to the *level* of top 1% income share than to the *change* in this measure. This finding is robust to various specifications.

Furthermore, it appears that the relationship between top-end inequality and growth is not constant during the development process. The main results focus on currently "advanced" countries, and various specifications in this study demonstrate that the level of top-end inequality does not correlate positively with subsequent growth in these countries in the medium or long run; this study discovers a negative association that is likely to fade as the level of per capita GDP increases. The main results related to the level of top-end inequality and subsequent growth are in accordance with the findings in a preceding study by Tuominen (2015). Although the current study abstains from causal inference, the results coincide with the growing literature suggesting that high inequality does not stimulate growth in the long term.

Finally, this study provides evidence that the sample composition matters. For example, the study provides tentative results on the association between top 1% income shares and subsequent growth in "less-advanced" countries. These findings indicate that the relationship may be different from what was discovered for "advanced" countries. "Less-advanced" economies need to be studied further when more data become available. Moreover, it will be interesting to investigate how the economic downturn after 2008 will affect the results of future studies.

Appendix A. Information on the top 1% income share series

Table A.5: Sources for the top 1% income share series used in this study. Series excluding capital gains have been used whenever possible. For more information on the series, see the original source and also Atkinson and Piketty (2007, 2010). The *top1* series in the 5-year average data are plotted in Figure A.5.

Country	**Source**
Argentina	Alvaredo et al. (2012)
Australia	Alvaredo et al. (2012)
Canada	Alvaredo et al. (2012)[a]
China	Alvaredo et al. (2012)
Denmark	Alvaredo et al. (2012)
Finland	Alvaredo et al. (2012) and Marja Riihelä (2011)[b]
France	Alvaredo et al. (2012)
Germany	Alvaredo et al. (2012)
India	Alvaredo et al. (2012)
Indonesia	Alvaredo et al. (2012)
Ireland	Alvaredo et al. (2012)
Italy	Alvaredo et al. (2012)
Japan	Alvaredo et al. (2012)
Mauritius	Alvaredo et al. (2012)
Netherlands	Alvaredo et al. (2012)
New Zealand	Alvaredo et al. (2012)
Norway	Alvaredo et al. (2012)
Portugal	Alvaredo et al. (2012)
Singapore	Alvaredo et al. (2012)
South Africa	Alvaredo et al. (2012)
Spain	Alvaredo et al. (2012)
Sweden	Alvaredo et al. (2012)
Switzerland	Alvaredo et al. (2012)[c]
United Kingdom	Alvaredo et al. (2012)
United States	Alvaredo et al. (2012)

Additional notes:

[a]Figures for the years 1982–2000 (in the annual series) are averages of the two alternative series provided in Alvaredo et al. (2012).

[b]Updated Finnish data covering years from 1993 onward. Received directly from Marja Riihelä by email (Feb 11, 2011).

[c]For all years except 1933, the annual estimates relate to income averaged over the year shown and the following year in the database (Alvaredo et al., 2012). Thus, a repeated value for two consecutive years is used as a basis for calculations in this study.

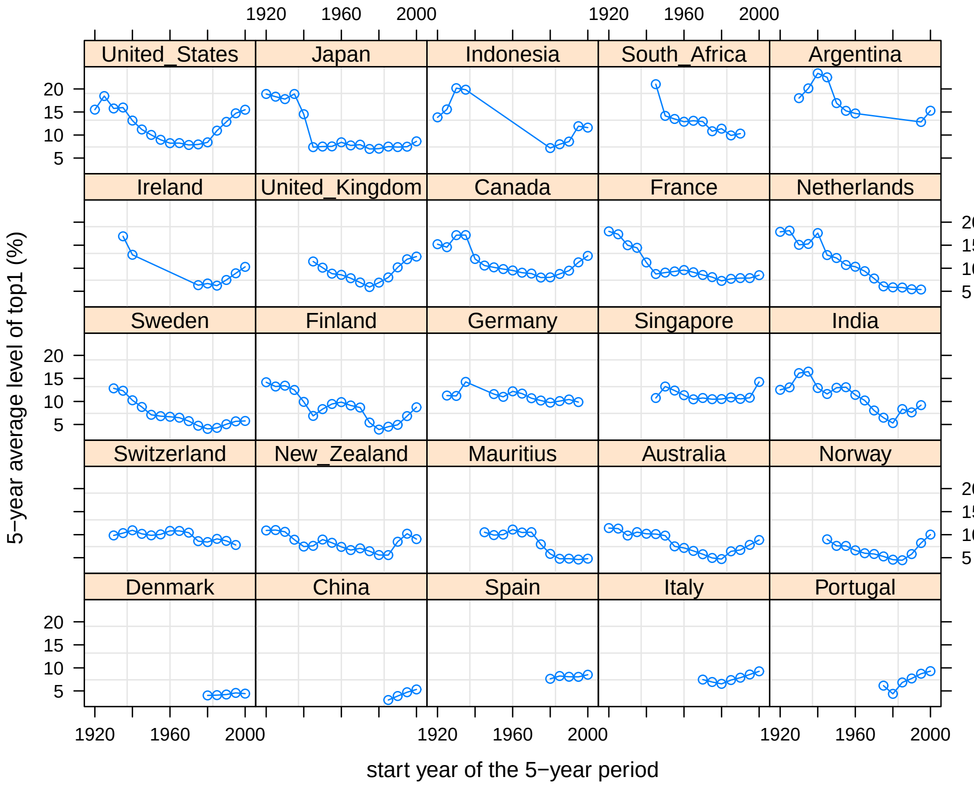


Figure A.5: Top 1% income shares for each country (5-year average data used in the models of Table 2; the time periods t are 1925–29, 1930–34, ..., and 2000–04; values from period 1920–24 are also plotted if they have been used in the construction of the "change in top 1% share" variable). Data source: see Table A.5.

Appendix B. Sources and definitions of other variables

Long series, simplified models (annual observations span 1920–2008):

– GDP per capita, 1990 international GK$; Maddison (2010). See Figure B.6.

Expanded models (annual observations span 1950–2009):

– GDP per capita: PPP converted GDP per capita (Laspeyres), derived from growth rates of domestic absorption, at 2005 constant prices (2005 I$/person); PWT 7.0 by Heston et al. (2011) (*"rgdpl2"*)
– Government consumption share of PPP converted GDP per capita at current prices (%); PWT 7.0 by Heston et al. (2011) (*"cg"*)
– Investment share of PPP converted GDP per capita at current prices (%); PWT 7.0 by Heston et al. (2011) (*"ci"*)
– Openness at current prices (%); PWT 7.0 by Heston et al. (2011) (*"openc"*)
– Price level of investment (PPP over investment/XRAT, where XRAT is national currency units per US dollar); PWT 7.0 by Heston et al. (2011) (*"pi"*)
– Average years of secondary schooling for total population (population aged 25 and over); Barro and Lee (2010); available every 5 years from 1950
– Average years of tertiary schooling for total population (population aged 25 and over); Barro and Lee (2010); available every 5 years from 1950
– Note: "China Version 2" data from PWT 7.0 (Heston et al., 2011) is used.

Gini data by Deininger and Squire (1996), "high quality" sample:
This sample is also used by Forbes (2000) and Banerjee and Duflo (2003, denoted by B&D in this appendix).

– Models of Table 4 include the following 18 countries: Australia, Canada, Denmark, Finland, France, Germany, India, Indonesia, Italy, Japan, the Netherlands, New Zealand, Norway, Singapore, Spain, Sweden, the United Kingdom, and the United States.
– Note that Argentina, Mauritius, South Africa, and Switzerland are not included in the sample used by Forbes and B&D. Moreover, China, Ireland, and Portugal are not studied in Table 4 because the observations on *top1* and Gini variables are not available for the same periods.
– The Gini series are constructed as in Forbes and B&D: the Gini measure every 5 years is picked for each country. If Gini is not available, then the closest measure in the 5 years preceding the date is used. Forbes and B&D create their Gini data using the following 5-year periods: 1961–65, 1966–70, 1971–75, 1976–80, 1981–85, and 1986–90; and they refer to these periods as 1965, 1970, 1975, 1980, 1985, and 1990, respectively.
– In this study, the closest corresponding period is used. This means that the period 1961–65 (1965 in Forbes and B&D) corresponds to the period 1960–1964 in this study's period structure, 1966–70 (1970 in Forbes and B&D) corresponds to 1965–69 in this study, ..., and 1986–90 (1990 in Forbes and B&D) corresponds to 1985–89 here.
– Thus, in the models of Table 4, the periods t are 1965–69, 1970–74, ..., and 1985–89. The descriptive statistics for the Gini coefficient variables are as follows:
$Gini_t$ N=62 ; min 23.3 ; mean 33.7 ; max 44.0, and
$Gini_t - Gini_{t-1}$ N=62 ; min -8.2 ; mean -0.2 ; max 5.2.

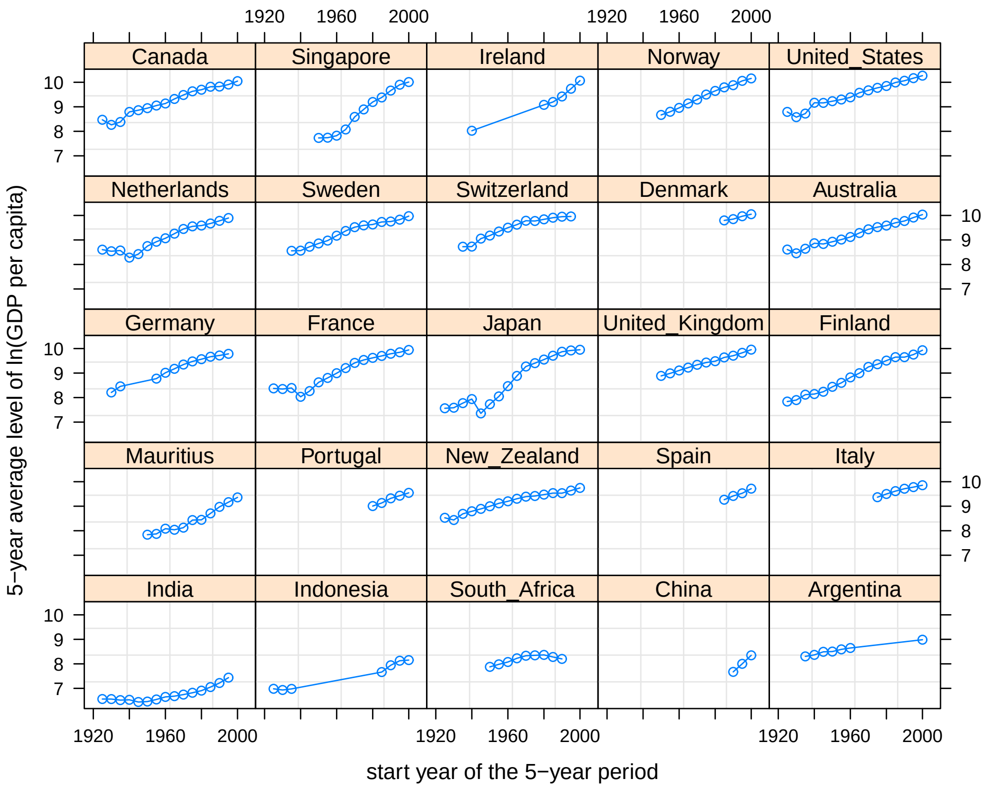


Figure B.6: Level of economic development for each country (5-year average data used in the models of Table 2; the time periods t are 1925–29, 1930–34, ..., and 2000–04). Data source: Maddison (2010).

Appendix C. Tensor product smooths

This appendix provides additional information to section 3. Tensor product smooths are recommended if one uses a smooth that contains more than one variable, but the scales of those variables are fundamentally different (i.e., measured in different units). Smooths of several variables are constructed from marginal smooths using the tensor product construction. The basic idea of a smooth function of two covariates is provided as an example.

Consider a smooth comprised of two covariates, x and z. Assume that we have low-rank bases to represent smooth functions f_x and f_z of the covariates. We can then write:

$$f_x(x) = \sum_{i=1}^{I} \alpha_i a_i(x) \quad \text{and} \quad f_z(z) = \sum_{l=1}^{L} \delta_l d_l(z),$$

where α_i and δ_l are parameters, and the $a_i(x)$ and $d_l(z)$ are known (chosen) basis functions such as those in the cubic regression spline basis.

Consider then the smooth function f_x. We want to convert it to a smooth function of both x and z. This can be done by allowing the parameters α_i to vary smoothly with z. We can write:

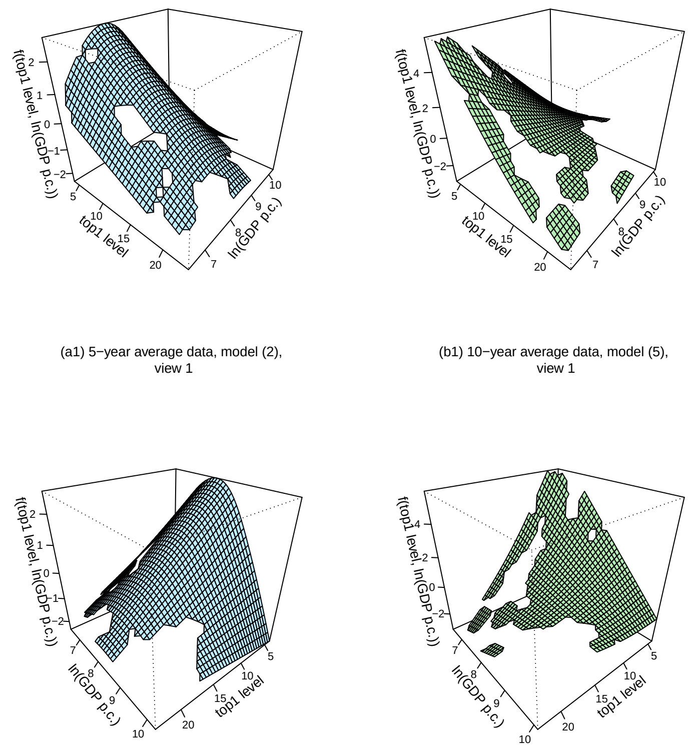
$$\alpha_i(z) = \sum_{l=1}^{L} \delta_{il} d_l(z),$$

and the tensor product basis construction gives:

$$f_{xz}(x, z) = \sum_{i=1}^{I} \sum_{l=1}^{L} \delta_{il} d_l(z) a_i(x).$$

The tensor product smooth has a penalty for each marginal basis. For further technical details, see Wood (2006).

Appendix D. Additional plots: long series from the 1920s



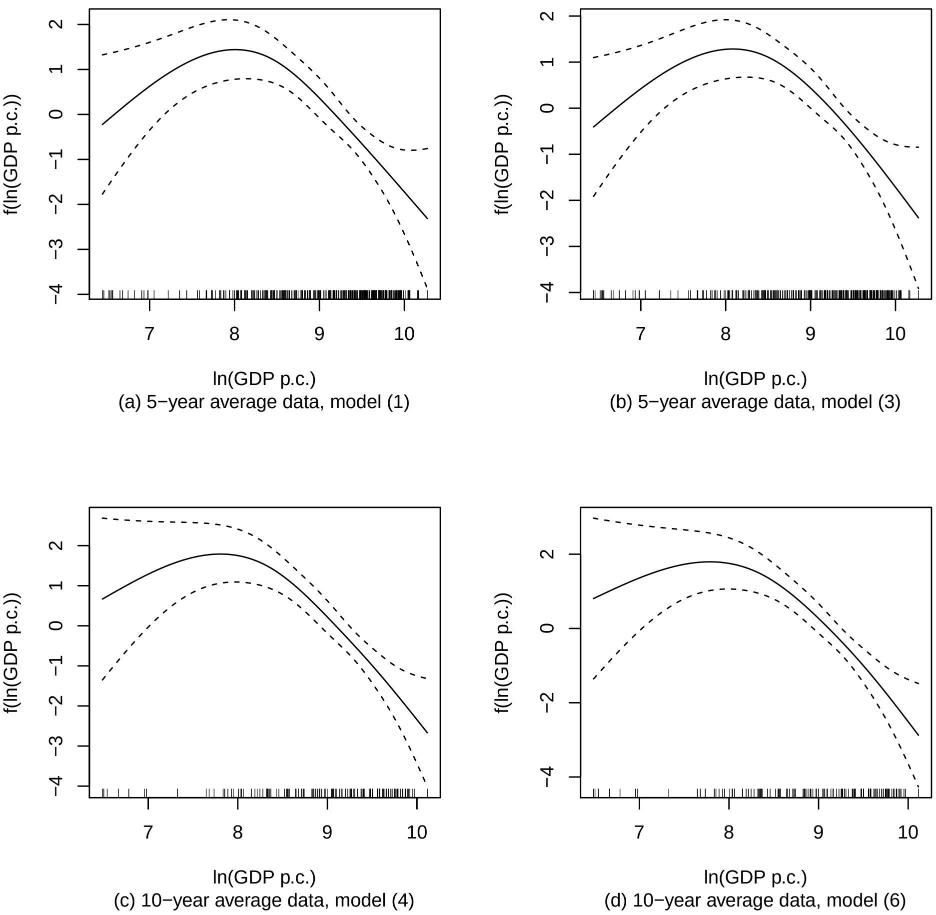
(a1) 5−year average data, model (2), view 1

(b1) 10−year average data, model (5), view 1

(a2) 5−year average data, model (2), view 2

(b2) 10−year average data, model (5), view 2

Figure D.7: Visualization of the simplified models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (5) of Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth function f. The smooths are illustrated from two views. In all plots, plot grid nodes that are too far from the true data points of the top 1% share and ln(GDP per capita) are excluded: the grid has been scaled into the unit square along with *top1* and GDP variables; grid nodes more than 0.1 from the predictor variables are excluded. Compare to Figure 2.



(a) 5−year average data, model (1)

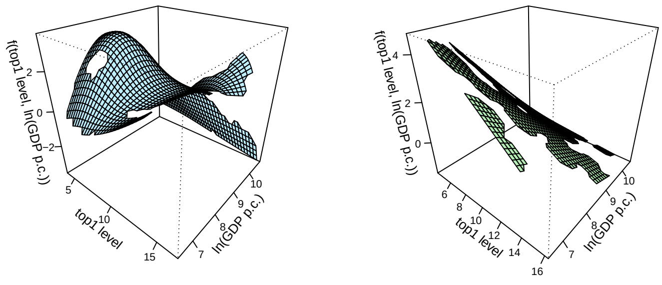
(b) 5−year average data, model (3)

(c) 10−year average data, model (4)

(d) 10−year average data, model (6)

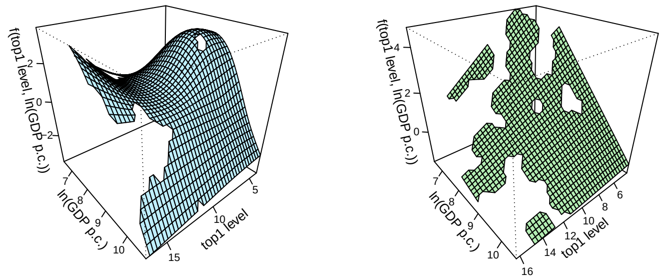
Figure D.8: Visualization of the simplified models' smooths $f(ln(GDP\ p.c.)_t)$ provided in Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as dashed lines and the covariate values as a rug plot along the horizontal axis.

Appendix E. Additional plots: series from the 1950s



(a1) 5−year average data, model (2), view 1

(b1) 10−year average data, model (5), view 1



(a2) 5−year average data, model (2), view 2

(b2) 10−year average data, model (5), view 2

Figure E.9: Visualization of the expanded models: smooths $f(top1_t, ln(GDP\ p.c.)_t)$ in models (2) and (5) of Table 3 (data from the 1950s onward; GDP data from PWT 7.0). The horizontal axes have the top 1% income share and ln(GDP per capita); the vertical axis has the smooth function f. The smooths are illustrated from two views. In all plots, plot grid nodes that are too far from the true data points of the top 1% share and ln(GDP per capita) are excluded: the grid has been scaled into the unit square along with *top1* and GDP variables; grid nodes more than 0.1 from the predictor variables are excluded. Compare to Figure 3.

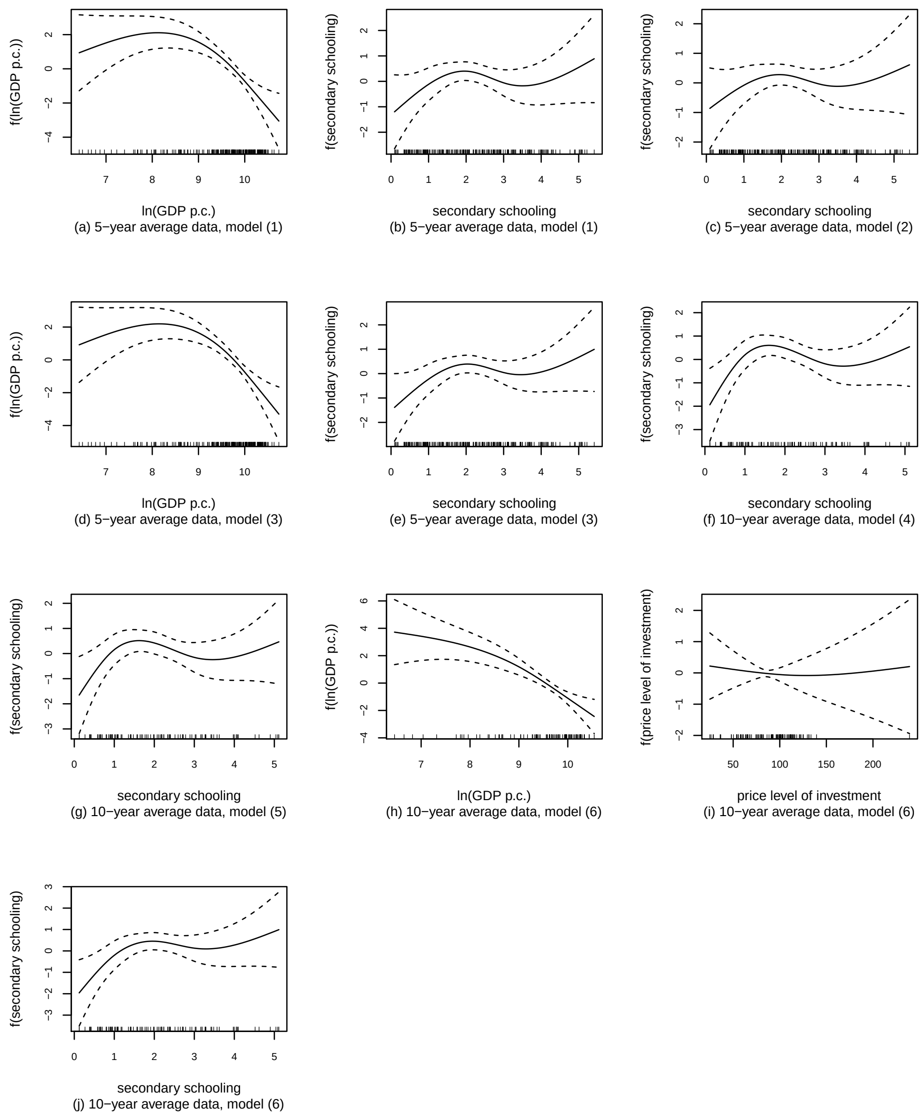


Figure E.10: Visualization of the expanded models' univariate smooths provided in Table 3 (data from the 1950s onward; GDP data from PWT 7.0). Each plot presents the smooth function f as a solid line. The plots also show the 95% Bayesian credible intervals as dashed lines and the covariate values as a rug plot along the horizontal axis.

References

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2012. The World Top Incomes Database. Data downloaded from website: `http://g-mond.parisschoolofeconomics.eu/topincomes` (January 4, 2012).

Andrews, D., Jencks, C., Leigh, A., 2011. Do Rising Top Incomes Lift All Boats? B.E. Journal of Economic Analysis & Policy 11(1), Article 6.

Atkinson, A.B., 2007. Measuring Top Incomes: Methodological Issues, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 18–42.

Atkinson, A.B., Brandolini, A., 2001. Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study. Journal of Economic Literature 39(3), 771–799.

Atkinson, A.B., Piketty, T. (Eds.), 2007. Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T. (Eds.), 2010. Top Incomes: A Global Perspective. Oxford University Press, Oxford.

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

Banerjee, A.V., Duflo, E., 2003. Inequality and Growth: What Can the Data Say? Journal of Economic Growth 8(3), 267–299.

Barro, R.J., 2000. Inequality and Growth in a Panel of Countries. Journal of Economic Growth 5(1), 5–32.

Barro, R., Lee, J.–W., 2010. A New Data Set of Educational Attainment in the World, 1950-2010. NBER Working Paper No. 15902. Data downloaded from website: `http://www.barrolee.com/`, Version 2.0, 07/10 (February 15, 2011).

Bénabou, R., 1996. Inequality and Growth, in: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual. The MIT Press, Cambridge. pp. 11–74.

Bénabou, R., 2000. Unequal Societies: Income Distribution and the Social Contract. American Economic Review 90(1), 96–129.

Chambers, D., Krause, A., 2010. Is the Relationship between Inequality and Growth Affected by Physical and Human Capital Accumulation? Journal of Economic Inequality 8(2), 153–172.

Deininger, K., Squire, L., 1996. A New Data Set Measuring Income Inequality. World Bank Economic Review 10(3), 565–591. Data downloaded from website: http://go.worldbank.org/UVPO9KSJJ0 (March 25, 2013).

Forbes, K.J., 2000. A Reassessment of the Relationship between Inequality and Growth. American Economic Review 90(4), 869–887.

Galor, O., Moav, O., 2004. From Physical to Human Capital Accumulation: Inequality and the Process of Development. Review of Economic Studies 71(4), 1001–1026.

Galor, O., Moav, O., Vollrath, D., 2009. Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. Review of Economic Studies 76(1), 143–179.

Galor, O., Zeira, J., 1993. Income Distribution and Macroeconomics. Review of Economic Studies 60(1), 35–52.

Halter, D., Oechslin, M., Zweimüller, J., 2014. Inequality and growth: the neglected time dimension. Journal of Economic Growth 19(1), 81–104.

Hastie, T., Tibshirani, R., 1986. Generalized additive models (with discussion). Statistical Science 1(3), 297–318.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, New York.

Herzer, D., Vollmer, S., 2013. Rising top incomes do not raise the tide. Journal of Policy Modeling 35(4), 504–519.

Heston, A., Summers, R., Aten, B., 2011. Penn World Table Version 7.0. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. Version: June 3, 2011. Data downloaded from website: http://pwt.econ.upenn.edu/php_site/pwt_index.php (January 4, 2012).

Kuznets, S., 1953. Shares of Upper Income Groups in Income and Saving. NBER Publication No. 55, New York.

Leigh, A., 2007. How Closely Do Top Income Shares Track Other Measures of Inequality? Economic Journal 117(524), F589–F603.

Leigh, A., 2009. Top Incomes, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 150–174.

Li, H., Zou, H.–f., 1998. Income Inequality is not Harmful for Growth: Theory and Evidence. Review of Development Economics 2(3), 318–334.

Maddison, A., 2010. Statistics on World Population, GDP and Per Capita GDP, 1–2008 AD. Data on GDP per capita. Data downloaded from website: http://www.ggdc.net/MADDISON/Historical_Statistics/vertical-file_02-2010.xls (November 12, 2010).

Perotti, R., 1993. Political Equilibrium, Income Distribution, and Growth. Review of Economic Studies 60(4), 755–776.

Perotti, R., 1996. Growth, Income Distribution, and Democracy: What the Data Say. Journal of Economic Growth 1(2), 149–187.

Piketty, T., 2001. Les Hauts revenus en France au 20e siècle: inégalités et redistribution, 1901–1998. B. Grasset, Paris.

Piketty, T., 2003. Income Inequality in France 1901–1998. Journal of Political Economy 111(5), 1004–1042.

Piketty, T., Saez, E., 2006. The Evolution of Top Incomes: A Historical and International Perspective. American Economic Review 96(2), 200–205.

Roine, J., Vlachos, J., Waldenström, D., 2009. The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data? Journal of Public Economics 93(7–8), 974–988.

Roine, J., Waldenström, D., 2015. Long-Run Trends in the Distribution of Income and Wealth, in: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution Vol. 2A. North-Holland, Amsterdam. pp. 469–592.

Tuominen, E., 2015. Top-end inequality and growth: Empirical evidence. University of Tampere, mimeo.

Voitchovsky, S., 2005. Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution. Journal of Economic Growth 10(3), 273–296.

Voitchovsky, S., 2009. Inequality and Economic Growth, in: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford. pp. 549–574.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton FL.

Essay III.

Reversal of the Kuznets curve: Study on the inequality–development relation using top income shares data[1]

Elina Tuominen

Abstract

In this study, recently published top 1% income share series are exploited in studying the inequality–development association in 26 countries from 1900 to 2010. The top income shares data are of high quality and provide interesting possibilities for studying slow development processes. Because many empirical inequality–development studies have challenged the use of quadratic specifications, this study addresses the issue of functional form by applying penalized spline methods. The relationship between the top 1% income share and development is found to experience a reversal at the highest levels of development and, thus, a positive association is now observed in many "advanced" economies. In an additional analysis covering a shorter time period, the discovered positive relationship holds at the highest levels of development when controls for two sectoral measures are included.

Keywords: inequality, top incomes, development, nonlinearity, longitudinal data
JEL classification: N30, O11, O15

Acknowledgments
This study has received funding from the Finnish Cultural Foundation and the Academy of Finland (project 268863). The author thanks Ravi Kanbur, Matti Tuomala, and Arto Luoma for the discussions. The study has also benefited from comments at the Finnish Doctoral Programme in Economics (FDPE) Workshops, the Annual Meeting of the Finnish Economic Association 2014, the Annual Meeting of the Austrian Economic Association 2014, the IIPF Congress 2014, and the UNU-WIDER Conference 2014 on "Inequality—measurement, trends, impacts, and policies." Remaining errors are the author's own.

[1]An earlier version of this study was published by UNU-WIDER, Helsinki (UNU-WIDER Working Paper 2015/036).

1. Introduction

In his seminal paper, Kuznets (1955) presented the famous "inverted-U hypothesis" between inequality and economic development; inequality first increases and then decreases as the country develops.[2] He suggested that during this process, the focus of the economy shifts from agriculture to modern sectors.[3] In addition to this famous idea of a sectoral shift, Kuznets discussed various other factors that affect the income distribution during the development process. For example, he noted that the concentration of savings at the top of the distribution induces inequality in the distribution before taxes and transfers, and he discussed equalizing forces such as political pressure for redistribution. Subsequently, various theoretical models have generated a Kuznets-type curve (e.g., Robinson, 1976; Greenwood & Jovanovic, 1990; Galor & Tsiddon, 1996; Aghion & Bolton, 1997; Dahan & Tsiddon, 1998). Empirical studies have presented mixed evidence on the shape of the inequality–development association, and the debate has focused on whether the results support the inverse-U hypothesis. A short and selective introduction to the empirical literature is provided next.[4]

In empirical applications, the chosen functional form plays an important role. For example, a cross-sectional study by Ahluwalia (1976) supports the inverted-U link, but Anand and Kanbur (1993) challenge the data quality and chosen functional forms. In comparison, Huang (2004), Lin et al. (2006), and Huang and Lin (2007) apply nonparametric methods to cross-sectional data and find evidence for the Kuznets hypothesis. However, it is possible that cross-sectional data cannot capture the complexity of the process. Panel studies have become more common after the construction of new inequality data sets. Possibly the most famous panel data set is by Deininger and Squire (1996). Although these data have been exploited in several studies, parametric analyses have shown differing results (e.g., Deininger & Squire, 1998; Barro, 2000). Further, Atkinson and Brandolini (2001) demonstrate that also this inequality data set has its shortcomings.

[2] Using data from the United States, Germany, and the United Kingdom, Kuznets (1955) got an impression of constancy in inequality around the turn of the twentieth century, followed by a secular decline in inequality at least since the 1920s.

[3] Kuznets (1955) provided numerical illustrations where (under certain assumptions) a mere population shift from the rural to urban sector can affect the overall income distribution: inequality first increases, and then declines.

[4] Further, Fields (2001) and Frazer (2006) provide overviews of the literature.

Recent studies suggest that using flexible methods is well-founded in inequality–development investigations. Frazer (2006) applies nonparametric regression in his study that spans approximately 50 years. In his pooled models, he discovers a nonlinear Gini–development association that is more complex than a second-degree polynomial. Specifically, he finds that the curve may be flat before it experiences a negative slope. His illustrations also show that the association may turn positive at the highest levels of development, but the confidence interval becomes wide at these development levels. Moreover, Zhou and Li (2011) conduct a nonparametric investigation on the inequality–development association using unbalanced panel data for the period 1962–2003. They find an inverse-U relation between Gini coefficients and economic development, but only after a certain level of development is reached. Further, Desbordes and Verardi (2012) use semiparametric methods with Gini data for the 1960–2000 period and provide empirical evidence for the latter stages of the Kuznets-type relation. Desbordes and Verardi also show that misspecified functional forms can lead to differing results on the inequality–development association.

Various inequality indices—including top income shares—have shown an upward trend in many countries over the last 20–30 years, and the inverse-U association has been challenged. In addition, List and Gallet (1999) study an unbalanced panel from 1961 to 1992 and find that, at the highest levels of economic development, there is a positive correlation between inequality and development. Although List and Gallet admit that the positive association may be a result of various factors, they suggest that one explanation is a new shift from manufacturing toward services in advanced economies.

To bring new insights into the inequality–development literature, the current study applies penalized regression spline methods to top 1% income share data. The World Top Incomes Database provides unprecedentedly long inequality series that cover almost a century for many countries (Alvaredo et al., 2013b). During this period, some countries have faced not only urbanization but also more advanced stages of development. Due to data unavailability, the focus of the study is on "advanced" countries; however some "less-advanced" countries are also included in the total sample of 26 countries. The data are of high quality compared to many other inequality data. Moreover, Leigh (2007) and Roine and Waldenström (2015) provide evidence that these series reflect changes in other inequality indices over time. Thus, it is interesting to exploit top income shares in inequality–development studies, particularly when other alternatives for long series are not available.[5]

This study finds that the inequality–development association is U-shaped after a certain development level when inequality is measured in terms of the top 1% income share. In an additional investigation encompassing the years 1980–2009, the positive association (at the highest levels of economic development) is robust to including controls for urbanization and the service sector. Moreover, there are similarities in the overall shape of the inequality–development relationship when one compares the results of this paper to the pooled results in Frazer (2006), although the studies use different distributional measures.

The remainder of this study is organized in the following manner: Section 2 introduces the data used in the empirical analysis, and section 3 describes the estimation method. Section 4 provides empirical results including sensitivity analysis. Finally, section 5 presents the conclusions.

2. Data

2.1. Top 1% income shares

Many of the available Gini series have suffered from comparability problems, both in time and between countries, and the series have not covered long time intervals. Using tax and population statistics, it is possible to compose long and fairly consistent series on top income shares. Kuznets (1953) was the first to use this kind of data to produce top income share estimates, and Piketty (2001, 2003) generalized Kuznets's approach. Following Piketty, different researchers have constructed top income share series using similar methods.[6] According to Leigh (2007), the evolution of top income shares is similar to that of various other inequality indices over time. In addition, Roine and Waldenström (2015) conclude that top income shares are useful in describing inequality.

Top income data can be easily accessed using the World Top Incomes

[5]To the best of the author's knowledge, there are no previous studies that exploit the new top income share series in this context.

[6]For more information on the methodology see, for example, Atkinson (2007). In addition, the advantages and limitations of the top income share series are discussed by Piketty and Saez (2006), Leigh (2007), and Roine and Waldenström (2015). Furthermore, Atkinson et al. (2011) provide a thorough overview of the top income literature.

Database by Alvaredo et al. (2013b).[7] The top 1% income shares in 26 countries from 1900 to 2010 are exploited, but the longitudinal data are not balanced (note that this is pre-tax income). Most of the data are from the English-speaking, Continental European, Southern European, and Nordic countries; however Japan, Singapore, and some "less-advanced" countries are also included.[8] The top 1% income share (*top1*) series are presented graphically in Appendix A. Table 1 provides summary statistics.

On the basis of the existing top income literature, an inverse U-shaped association between *top1* and economic development is not expected. For example, in the English-speaking countries, the evolution of the top 1% income shares resembles U over the twentieth century because there has been a significant increase since the 1980s; whereas the top 1% shares in Continental Europe and Japan have remained fairly stable during the past three decades. Further, Atkinson et al. (2011) and Roine and Waldenström (2015) discuss the problems of fitting the evolution of top income shares into the approach where the inequality–development relation is described by sectoral shifts. Other factors—also indicated by Kuznets (1955)—seem relevant, particularly taxation and the concentration of savings at the top.[9] Moreover, "superstar" theories and the possibility of changing norms are examples of suggested explanations for the recent increase in top incomes in many countries. For more discussion, see, for example, Piketty and Saez (2006) and Alvaredo et al. (2013a).

2.2. Economic development and sectoral variables

The level of economic development is measured in a traditional manner using GDP per capita. The GDP per capita data (1990 international GK$) are available annually until 2010 in the Maddison Project update (Bolt & van Zanden, 2013). Data from 1900 are used whenever available. In an additional analysis encompassing the years 1980–2009, the models include controls for

[7] The first book on these series, edited by Atkinson and Piketty (2007), contrasted the evidence from the Continental Europe and English-speaking countries. The second volume, also edited by Atkinson and Piketty, was published in 2010. The database builds on these volumes, and the project is ongoing.

[8] Argentina, China, Colombia, India, Indonesia, Mauritius, and South Africa.

[9] Roine et al. (2009) provide empirical evidence for the negative association between tax progressivity and top income shares. Moreover, Kanbur (2000) notes that inequality–development studies tend to minimize the role of policy.

Table 1: Descriptive statistics.

Annual data (1900–2010)	N	min	mean	max
top1	1609	2.7	10.6	28.0
ln(GDP p.c.)	1609	6.4	8.9	10.4
Data averaged over 5-year periods (1980–2009)[a]	N	min	mean	max
top1	129	3.0	8.8	20.5
ln(GDP p.c.)	129	7.2	9.5	10.3
urbanization	129	22.1	71.1	100.0
service sector	129	17.7	62.5	78.6

[a]The 5-year periods are defined as 1980–84, 1985–89, ..., and 2005–09.

two sectors, namely, urban and service sectors. It should be interesting to see whether the inclusion of sectoral variables affects the relationship between top-end inequality and economic development. Urbanization data describe the *population residing in urban areas (%)* (United Nations, 2012). These data are available every five years. The service sector is measured with *employment in service sector (% of total employment)* (World Bank, 2014a), and these data are available from 1980 onward. See Table 1 for descriptive statistics.

Although the investigated time span becomes considerably shorter with the two sectoral variables, this approach can be considered an extension to previous studies. For example, Frazer (2006) reports controlling for urbanization but does not provide detailed results on the inequality–urbanization relationship. Desbordes and Verardi (2012) do not include sectoral variables in their empirical models.[10]

3. Estimation method

Additive models provide a flexible framework for investigating the association between inequality and development.[11] This study follows the approach

[10]Kanbur and Zhuang (2013) is a recent example of focusing on the inequality–urbanization relationship in four Asian countries in the spirit of Kuznets (1955).

[11]Additive models are a special case of generalized additive models (GAMs). GAMs were introduced by Hastie and Tibshirani (1986, 1990). They present a GAM as a generalized linear model with a linear predictor that involves a sum of smooth functions of covariates. Some of the covariates can enter the model in linear form. Note here the analogy to "generalized linear models" and "linear models." This paper is restricted to a special case: it uses an identity link and assumes normality in errors, which leads to additive models.

presented in Wood (2006). The basic idea is that the model's predictor is a sum of linear and smooth functions of covariates:

$$\mathbb{E}(Y_i) = \boldsymbol{X}_i^* \boldsymbol{\theta} + f_1(x_{1i}) + f_2(x_{2i}) + ...$$

In the above presentation, Y_i is the response variable (here: *top1*), $\boldsymbol{X}_i^*$ is a row of the model matrix for any strictly parametric model components, $\boldsymbol{\theta}$ is the corresponding parameter vector, and the $f_\bullet$ are smooth functions of the covariates, $x_\bullet$.

The flexibility of these models comes at the cost of two problems. First, one needs to represent the smooth functions $f_\bullet$ in some manner. One way to represent these smooths is to use cubic regression splines, which is the approach adopted in this study. A cubic regression spline is a curve constructed from sections of cubic polynomials that are joined together so that the resulting curve is continuous up to the second derivative. The points at which sections are joined (and the end points) are the knots of the spline, and these locations must be chosen. The spline can be represented in terms of its values at the knots.[12] Second, the amount of smoothness that functions $f_\bullet$ will have needs to be chosen. Overfit is to be avoided and, thus, departure from smoothness is penalized. The appropriate degree of smoothness for $f_\bullet$ can be estimated from the data by, for example, maximum likelihood.

Illustration

Consider a model containing only one smooth function of one covariate: $y_i = f(x_i) + \epsilon_i$, where ϵ_i are i.i.d. $N(0, \sigma^2)$ random variables. To estimate function f here, f is represented so that the model becomes a linear model. This is possible by choosing a basis, defining the space of functions of which f (or a close approximation to it) is an element. In practice, one chooses basis functions, which are treated as known.

Assume that the function f has a representation $f(x) = \sum_{j=1}^{k} \beta_j b_j(x)$, where β_j are unknown parameters and $b_j(x)$ are known basis functions. Using a chosen basis for f implies that we have a linear model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where the model matrix $\boldsymbol{X}$ can be represented using basis functions such as those in the cubic regression spline basis. The departure from smoothness can be penalized with $\int f''(x)^2 dx$. The penalty $\int f''(x)^2 dx$ can be expressed as

[12] There are usually two extra conditions that specify that the second derivative of the curve should be zero at the two end knots.

$\boldsymbol{\beta}^T \boldsymbol{S} \boldsymbol{\beta}$, where $\boldsymbol{S}$ is a coefficient matrix that can be expressed in terms of the known basis functions.

Accordingly, the penalized regression spline fitting problem is to minimize $\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + \lambda \boldsymbol{\beta}^T \boldsymbol{S} \boldsymbol{\beta}$, with respect to $\boldsymbol{\beta}$. The problem of estimating the degree of smoothness is a problem of estimating the smoothing parameter λ.[13] The penalized least squares estimator of $\boldsymbol{\beta}$, given λ, is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T \boldsymbol{X} + \lambda \boldsymbol{S})^{-1} \boldsymbol{X}^T \boldsymbol{y}$. Thus, the expected value vector is estimated as $\widehat{\mathbf{E}(\boldsymbol{y})} = \hat{\boldsymbol{\mu}} = \boldsymbol{A}\boldsymbol{y}$, where $\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{X} + \lambda \boldsymbol{S})^{-1} \boldsymbol{X}^T$ is called an influence matrix.

This setting can be augmented to include several covariates and smooths. Given a basis, an additive model is simply a linear model with one or more associated penalties.

Practical notes

The size of basis dimension for each smooth is usually not critical in estimation, because it only sets an upper limit on the flexibility of a term. Smoothing parameters control the effective degrees of freedom (*edf*). Effective degrees of freedom are defined as trace($\boldsymbol{A}$), where $\boldsymbol{A}$ is the influence matrix. The effective degrees of freedom can be used to measure the flexibility of a model. It is also possible to divide the effective degrees of freedom into degrees of freedom for each smooth. For example, a simple linear term would have one degree of freedom, and *edf*=2.3 can be thought of as a function that is slightly more complex than a second-degree polynomial.

Confidence (credible) intervals for the model terms can be derived using Bayesian methods, and approximate *p*-values for model terms can be calculated. Models can be compared using information criteria such as the Akaike information criterion (AIC). When using the AIC for penalized models (models including smooth terms), the degrees of freedom are the effective degrees of freedom, not the number of parameters. Moreover, random effects can be included in these models. For further details, see Wood (2006).[14]

---

[13] In the estimation, one faces a bias–variance tradeoff: on the one hand, the bias should be small, but on the other hand, the fit should be smooth. One needs to compromise between the two extremes. $\lambda \longrightarrow \infty$ results in a straight line estimate for f, and $\lambda = 0$ leads to an unpenalized regression spline estimate.

[14] The results presented in this study are obtained using the R software package "mgcv" (version 1.7-21), which includes a function "gam." Basis construction for cubic regression splines is used (the knots are placed evenly through the range of covariate values by default). The maximum likelihood method is used in the selection of the smoothing

4. Results

In the baseline models, the estimation is implemented with annual data from 1900 to 2010. The results are also checked by studying different subsets of the sample and changing the data structure from annual to 5-year average data. Finally, in an additional analysis, urbanization and service sector variables are included in models with 5-year average data spanning the years from 1980 to 2009.

4.1. Baseline models

The baseline results are for annual data spanning 1900–2010. The models are of the form

$$top1_{it} = \alpha + f(ln(GDP\ p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it},$$

where i refers to country and t to year, α is a constant, f is a smooth function that is described using a penalized cubic regression spline, δ_{decade} is a time dummy (one decade is the reference category), u_i is a country effect, and $\epsilon_{it} \sim N(0, \sigma^2)$ is the error term. The country effects can be omitted, fixed (i.e., dummy variables), or random ($u_i \sim N(0, \sigma_u^2)$). Different strategies in modeling country effects are reported because the literature does not follow a unified approach. Thus, the reader can also see when and how the chosen specification affects the results.

Details of the model without country effects are provided in column (1) of Table 2. Models (2) and (3) of Table 2 include country effects, and the table shows that including these effects improves the model fit. Figure 1 illustrates the smooth functions f in these three models. The fixed-effect (FE) and random-effect (RE) specifications give practically identical fits. In all three specifications, there is a possibility of a flat curve at lower levels of development (ln(GDP per capita) < 8, approximately). Further, after a certain level of development (ln(GDP per capita) > 8.5, approximately), all smooths show U shape (or J shape).[15]

---

parameters. The identifiability constraints (due to, for example, the model's additive constant term) are taken into account by default. The function "gam" also allows for simple random effects: it represents the conventional random effects in a GAM as penalized regression terms. More details can be found in Wood (2006) and the R project's web pages (`http://cran.r-project.org/`).

[15]Note: $\exp(8) \approx 2980$ and $\exp(8.5) \approx 4910$ (1990 international GK\$).

Table 2: Baseline models, using annual data (years 1900–2010): effective degrees of freedom for each smooth. Intercepts, country effects, and time effects[a] are not reported. For graphical illustration of smooth functions f, see Figure 1.

	dependent variable: $top1_t$ (N=1609)		
	(1)	(2)	(3)
$f(ln(GDP\ p.c.)_t)$	$[edf \approx 9.2^{b}]$***	$[edf \approx 10.4^{b}]$***	$[edf \approx 10.4^{b}]$***
country effects	no	fixed	random
adjusted r^2	0.59	0.82	0.82
AIC	7950	6642	6642

*** indicates significance at the 1% level.
The smooth terms' significance levels are based on approximate p-values.

[a] All models (1)–(3) include time effects. Time effects are dummy variables for different decades. However, all observations for 2000–2010 are considered in the "last" decade.
[b] The basis dimension of the smooth before imposing identifiability constraints is $k = 15$.

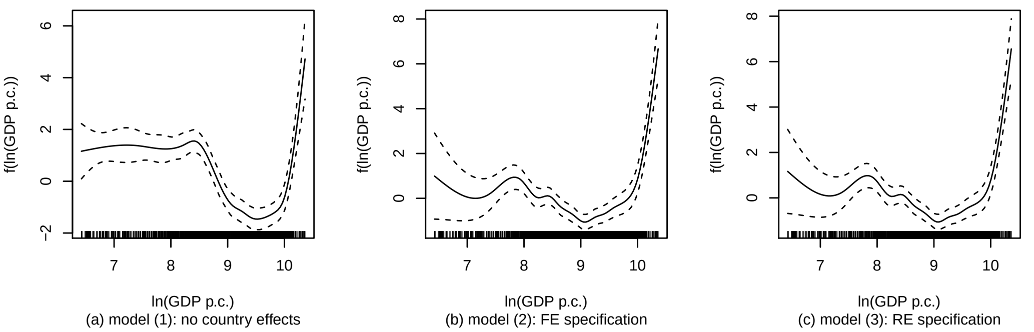


Figure 1: Illustration of the *top1*–development relation (annual data 1900–2010). See Table 2 for details. The solid line represents the smooth function $f(ln(GDP\ p.c))$. The plots also show the 95% Bayesian credible intervals (dashed) and the covariate values as a rug plot along the horizontal axis.

The overall shape of $f(ln(GDP\ p.c.))$ resembles the shape that Frazer (2006) shows for the Gini–development relationship (except for the steep positive slope at the highest levels of development). This similarity supports the notion that top income shares reflect the same characteristics as the traditional Gini coefficients. Even the downward peak close to ln(GDP per capita) ≈ 9.5 in plot (a) of Figure 1 appears to be reasonable compared to Frazer's pooled models.[16]

4.2. Sensitivity of the baseline models' results

In the first check, the English-speaking, Nordic, Continental and Southern European, and "less-advanced" countries were studied separately.[17] More detailed information on the models with random country effects is reported in Table B.5 in Appendix B. The illustrations of the smooths $f(ln(GDP\ p.c.))$ in these specifications are provided in Figure 2. Plots in Figure 2 illustrate that the association is not uniform at lower levels of development (ln(GDP per capita) < 8.5, approximately). However, there seems to be a pattern that holds as countries reach a higher level of economic development: there is a negative relationship between *top1* and the level of development when $8.5 <$ ln(GDP per capita) < 9.5 (approximately). In general, the shape of the association between top-end inequality and development is fairly uniform when ln(GDP per capita) > 8.5. The results in Figure 2 are also in line with plot (c) of Figure 1, which illustrates the corresponding random-effect specification with the entire sample. Moreover, the main results of the fixed-effect specifications for separate groups accorded with those in Figure 2.[18]

The second check was concerned with the sensitivity of excluding groups of countries from the entire sample. The previously discovered U shape (or J shape) emerges again at development levels ln(GDP per capita) > 8.5 (approximately), and the downward peak of the U is located between $9 <$ ln(GDP per capita) < 10. More detailed information on the models with random country effects is provided in Figure B.5 in Appendix B. Further-

[16]Note: $\exp(9.5) \approx 13360$ (1990 international GK$ in the current study).

[17]Singapore and Japan do not fit into these categories and were, thus, not included in these group-wise investigations.

[18]Only in the group of Continental and Southern European countries the curve may be flat at the highest levels of economic development. Detailed results on the fixed-effect specifications are not provided for the sake of brevity.

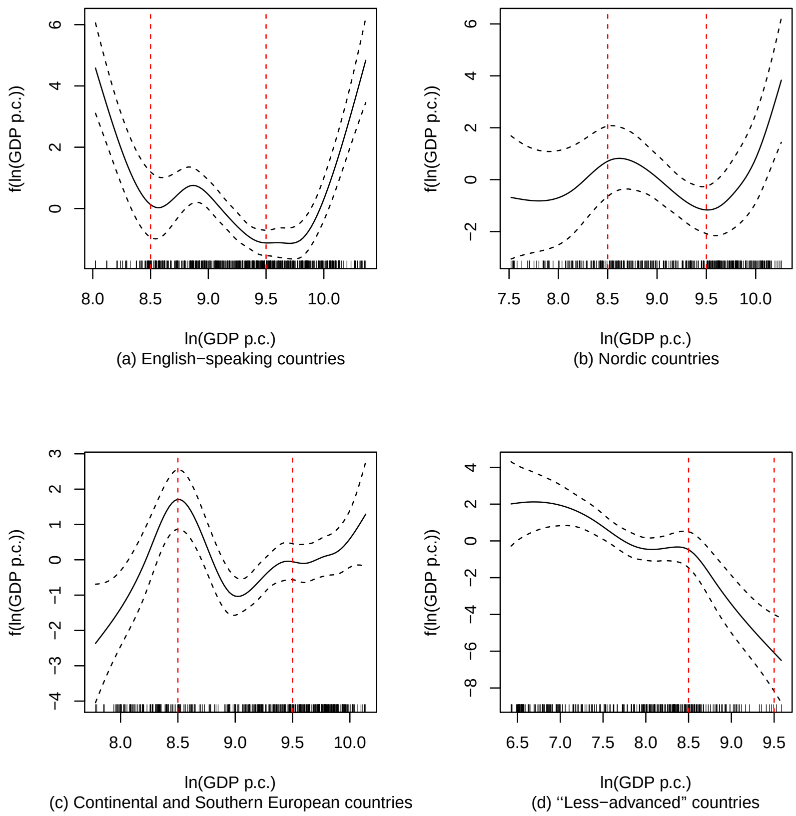


Figure 2: Illustration of the *top1*–development relation with four different subsets of the sample (annual data 1900–2010). The models include decade dummies and random country effects (Table B.5 in Appendix B provides details). The solid line represents the smooth function $f(ln(GDP\ p.c.))$. The plots also show the 95% Bayesian credible intervals (dashed) and the covariate values as a rug plot along the horizontal axis. Vertical dashed lines have been added to highlight the idea of a negative slope between 8.5 < ln(GDP per capita) < 9.5 (approximately).

more, when the corresponding fixed-effect specifications were studied, the results were similar to those reported in Figure B.5.[19]

Finally, the annual data results were checked against the corresponding results with the 5-year average data. The main findings with the 5-year data spanning 1900–2009 do not contradict the results presented in subsection 4.1. Appendix C provides graphical illustrations. Thus, the overall results do not seem to depend on the choice between annual and 5-year average data. The next investigations are conducted with 5-year averages, but the models are augmented with sectoral variables.

4.3. Additional analysis: controlling for two sectors

This subsection provides an additional analysis where models include controls for urban and service sectors. The analysis is implemented using 5-year averages, where the periods are 1980–1984, 1985–1989, ..., and 2005–2009.[20] The studied specifications are as given below:

$$\begin{aligned} \mathit{top1}_{it} = \alpha &+ f_1(\mathit{ln}(\mathit{GDP\ p.c.})_{it}) + f_2(\mathit{urbanization}_{it}) \\ &+ f_3(\mathit{service\ sector}_{it}) + \delta_{decade} + u_i + \epsilon_{it}, \end{aligned}$$

where i refers to country and t to 5-year period, α is a constant, smooth functions f_j $(j = 1, 2, 3)$ are approximated using penalized cubic regression splines, δ_{decade} is a fixed time effect (one decade is the reference category), u_i is a country effect, and $\epsilon_{it} \sim N(0, \sigma^2)$ is the error term; the values of the top 1% share, ln(GDP per capita) and sectoral variables are now period averages. As before, the country effects can be omitted, fixed, or random depending on the specification. Initially, all smooths f_j were allowed to enter in a flexible form, but a linear term was suggested for the service sector variable in some models. The models in question were then re-estimated with this linearity restriction.

Table 3 provides details on models with two sectors. Models (2) and (3) have linear terms for the service sector, and the coefficients are provided

[19]The detailed results on the fixed-effect specifications are not reported for the sake of brevity. In addition, the effect of excluding Japan and Singapore from the sample was tested because these two countries do not fit into the discussed categorization. The main results that relate to "medium" and "high" levels of development are not sensitive to including or excluding these countries.

[20]Taking period averages should reduce potential short-run disturbances. Moreover, the urbanization variable is available every five years.

Table 3: Models with two sectors, using 5-year average data (years 1980–2009): effective degrees of freedom for each smooth $f_\bullet$ and coefficients for linear terms. Intercepts, country effects, and time effects[a] are not reported. The smooths with *edf* > 1 are illustrated in Figure 3.

	dependent variable: $top1_t$ (N=129)		
	(1)	(2)	(3)
$f_1(ln(GDP\ p.c.)_t)$	[$edf \approx 4.7^{b}$]***	[$edf \approx 5.3^{b}$]***	[$edf \approx 5.4^{b}$]***
$f_2(urbanization_t)$	[$edf \approx 5.8^{b}$]***	[$edf \approx 3.6^{b}$]*	[$edf \approx 4.1^{b}$]**
$f_3(service\ sector_t)$	[$edf \approx 2.9^{b}$]***	[linear[b]] 0.096**	[linear[b]] 0.120***
country effects	no	fixed	random
adjusted r^2	0.72	0.94	0.93
AIC	542	361	371

***, **, * indicate significance at the 1, 5, and 10% levels, respectively.
The p-values for parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are reported. The smooth terms' significance levels are based on approximate p-values.

[a]All models (1)–(3) include time effects. Time effects are dummy variables for different decades.

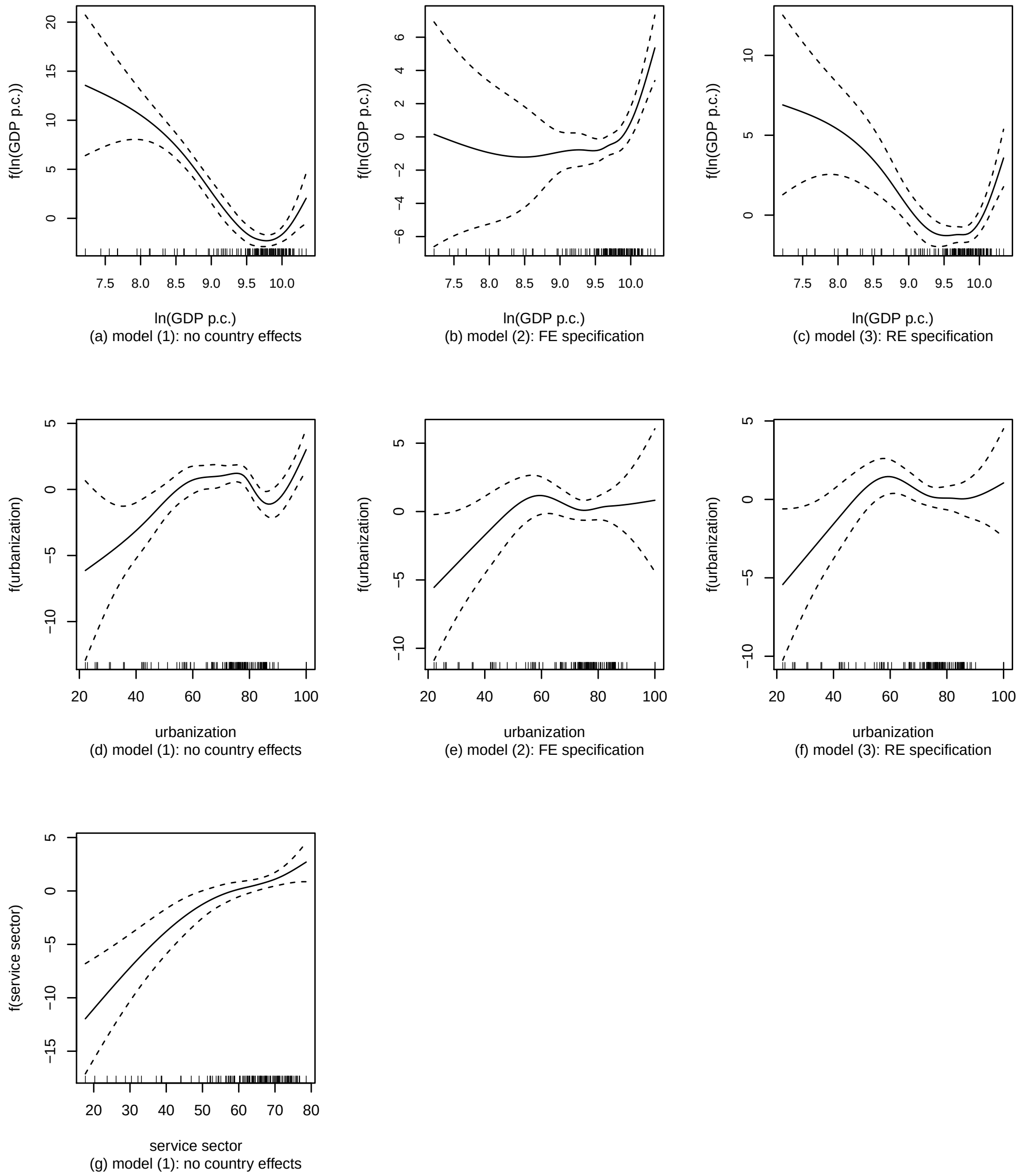
[b]The basis dimension of the smooth before imposing identifiability constraints is $k = 10$.

in the table. Figure 3 provides plots of the other smooth functions. The results on sectoral variables are fairly uniform, irrespective of the country-effect specification. The urbanization smooth resembles an inverted-U curve (particularly in plots (e) and (f) of Figure 3). The association between the top 1% share and employment in services is positive, which leads to speculation regarding whether this illustrates a new structural shift.

Let us then focus on the GDP per capita variable. In plots (a) and (c) of Figure 3, the model without country effects and the model with random country effects show very similar shapes for the smooth $f(ln(GDP\ p.c.))$, and the overall shape does not contradict previously reported results.[21] In contrast, the fixed-effect specification in plot (b) does not confirm the U-shaped relationship at "medium-to-high" levels of development.[22] However, the positive relationship at the highest levels of GDP per capita is discovered in all three specifications, and the "turning point" is located close to

[21]Moreover, Frazer (2006) controls for urbanization in the sensitivity checks of his pooled model and finds that the overall shape of the Gini–development relationship holds.

[22]This conclusion regarding the smooth $f(ln(GDP\ p.c.))$ does not change if the sectoral variables are excluded from the model with fixed country effects (when period 1980–2009 is studied).



(a) model (1): no country effects
(b) model (2): FE specification
(c) model (3): RE specification
(d) model (1): no country effects
(e) model (2): FE specification
(f) model (3): RE specification
(g) model (1): no country effects

Figure 3: Illustrations of the smooths, using 5-year average data (years 1980–2009). See Table 3 for the details of the models. The solid line represents the smooth function $f_{\bullet}$. The plots also show the 95% Bayesian credible intervals (dashed) and covariate values as a rug plot along the horizontal axis.

ln(GDP per capita) ≈ 9.5. Thus, the discovered positive association holds at the highest levels of development when two sectors are controlled for.

These results were also checked against leaving country groups out of the sample, one group at a time. The categorization was the same as that in the previous subsection (and in the lower panel of Table B.5). The random-effect specifications were intuitive when compared to the model (3) of Table 3. In comparison, the fixed-effect specifications were slightly more sensitive to the exclusion of country groups, but also these findings were reasonable when compared to the whole-sample results of model (2) in Table 3.[23] For brevity, the details of these checks are not reported.

Finally, an alternative measure for the service sector was tested. Data on *services, etc., value added (% of GDP)* (World Bank, 2014b) begin from the 1960s for some countries, but Swiss data are not available. Results related to ln(GDP per capita) and urbanization did not change. The alternative service sector measure correlated positively with *top1*, but it was not statistically significant at the 10% level in specifications with country effects. However, these results were not in conflict with the models of Table 3. Thus, the details of these checks are not reported.

5. Discussion

A vast number of empirical studies have explored the relationship between inequality and development, but the results have been mixed. The current paper addresses the issue by applying flexible methods to new data. The results of the current study are based on an unbalanced longitudinal data from 26 countries over the years 1900–2010. Various specifications in this paper suggest a negative association between the top 1% income share and ln(GDP per capita) after a certain point in the development process. Furthermore, the current study finds that this relationship turns positive at even

[23] Main findings with the FE specifications: When the Continental and Southern European countries were excluded from the sample, GDP per capita variable was not statistically significantly related to the top 1% share at the 10% level; both sectoral variables correlated positively with the top 1% share. In comparison, when the "less-advanced" countries were excluded, the sectoral variables were not significantly related to the top 1% share at the 10% level, but—as expected—there was a statistically significant, positive relationship between per capita GDP and the top income share. Further, excluding either the English-speaking or the Nordic countries from the sample barely affected the main conclusions.

higher levels of economic development. Thus, the data suggest a reversal of the Kuznets curve after a certain development level is reached. However, the current sample includes only some "less-advanced" countries, and more research is needed when new data become available.

In an additional analysis encompassing the period 1980–2009, this study assumes a broad interpretation of Kuznets's idea of sectoral shifts. The analysis is descriptive, but the results favor that something more than sectoral shifts are needed to explain changes in top-end inequality in the course of economic development. Specifically, the discovered positive association between the top 1% share and economic development (at the highest levels of development) holds when measures for urbanization and service sector are included. This accords with the existing literature on top incomes, which has highlighted other explanations for the evolution of top income shares.

Appendix A. Top 1% income share series

Table A.4: Top 1% income share series (years 1900–2010). For better comparability, series excluding capital gains have been selected whenever possible. For more information, see the source and Atkinson and Piketty (2007, 2010). The series are plotted in Figure A.4.

Country	***N***	**Source**
Argentina	39	Alvaredo et al. (2013b)
Australia	90	Alvaredo et al. (2013b)
Canada	91	Alvaredo et al. (2013b)[b]
China	18	Alvaredo et al. (2013b)
Colombia	18	Alvaredo et al. (2013b)
Denmark	95	Alvaredo et al. (2013b)
Finland	90	Alvaredo et al. (2013b)[c]
France	96	Alvaredo et al. (2013b)[d]
Germany	47	Alvaredo et al. (2013b)
India	71	Alvaredo et al. (2013b)
Indonesia	28	Alvaredo et al. (2013b)
Ireland	37	Alvaredo et al. (2013b)
Italy	34	Alvaredo et al. (2013b)
Japan	110	Alvaredo et al. (2013b)
Mauritius	52[a]	Alvaredo et al. (2013b)
Netherlands	55	Alvaredo et al. (2013b)
New Zealand	83	Alvaredo et al. (2013b)
Norway	69	Alvaredo et al. (2013b)
Portugal	24	Alvaredo et al. (2013b)
Singapore	59[a]	Alvaredo et al. (2013b)
South Africa	62[a]	Alvaredo et al. (2013b)
Spain	30	Alvaredo et al. (2013b)
Sweden	79	Alvaredo et al. (2013b)
Switzerland	74	Alvaredo et al. (2013b)[e]
United Kingdom	60	Alvaredo et al. (2013b)
United States	98	Alvaredo et al. (2013b)
	total: 1609	

[a]There would be more top 1% income share observations, but GDP per capita data are not available: Mauritius (+4), Singapore (+3), and South Africa (+9).
[b]Two partially overlapping series are available. Here; series up to 1981 is based on tax data, and series from 1982 is based on Longitudinal Administrative Database.
[c]Two partially overlapping series are available. Here; series up to 1989 is based on tax data, and the series from 1990 is based on the Income Distribution Survey.
[d]In the original source, the figure for 1905 is averaged for 1900–1910.
[e]For all years except 1933, the estimates relate to income averaged over the year shown and the following year. Thus, repeated value for two consecutive years is used in this study.

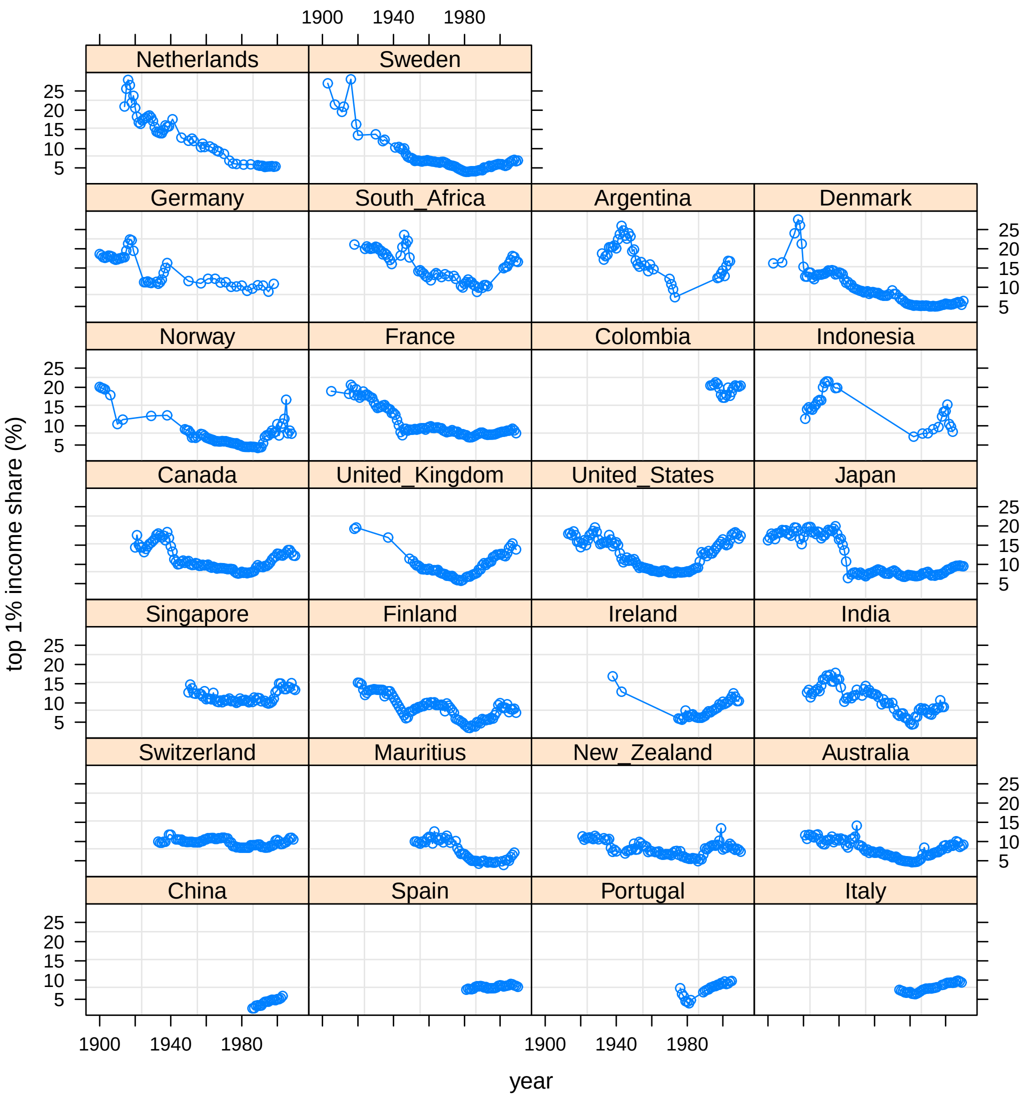


Figure A.4: Top 1% income share series for each country (years 1900–2010). See Table A.4 for details. Data source: Alvaredo et al. (2013b).

Appendix B. Model details: subsets of the sample

Table B.5: Subsets of the sample. Results of models with fixed time effects[a] and random country effects, using annual data (years 1900–2010): effective degrees of freedom for each smooth.

$top1_{it} = \alpha + f(ln(GDP\ p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it}$	N	smooth f
In Figure 2:		
English-speaking [b]	459	$[edf \approx 7.4^{\mathrm{f}}]$***
Nordic [c]	333	$[edf \approx 5.9^{\mathrm{f}}]$***
Continental and Southern Europe [d]	360	$[edf \approx 6.9^{\mathrm{f}}]$***
"Less-advanced" [e]	288	$[edf \approx 5.4^{\mathrm{f}}]$***
In Figure B.5:		
Without English-speaking [b]	1150	$[edf \approx 10.0^{\mathrm{g}}]$***
Without Nordic [c]	1276	$[edf \approx 10.0^{\mathrm{g}}]$***
Without Continental/Southern Europe [d]	1249	$[edf \approx 9.8^{\mathrm{g}}]$***
Without "less-advanced" [e]	1321	$[edf \approx 9.5^{\mathrm{g}}]$***

*** indicates significance at the 1% level.
The smooth terms' significance levels are based on approximate p-values.

[a] Time effects are dummy variables for different decades. However, all observations for 2000–2010 are considered in the "last" decade. One decade is the reference category.
[b] Australia, Canada, Ireland, New Zealand, the United Kingdom, and the United States.
[c] Denmark, Finland, Norway, and Sweden.
[d] France, Germany, Italy, the Netherlands, Portugal, Spain, and Switzerland.
[e] Argentina, China, Colombia, India, Indonesia, Mauritius, and South Africa.
[f] The basis dimension of the smooth before imposing identifiability constraints is $k = 10$.
[g] The basis dimension of the smooth before imposing identifiability constraints is $k = 15$.

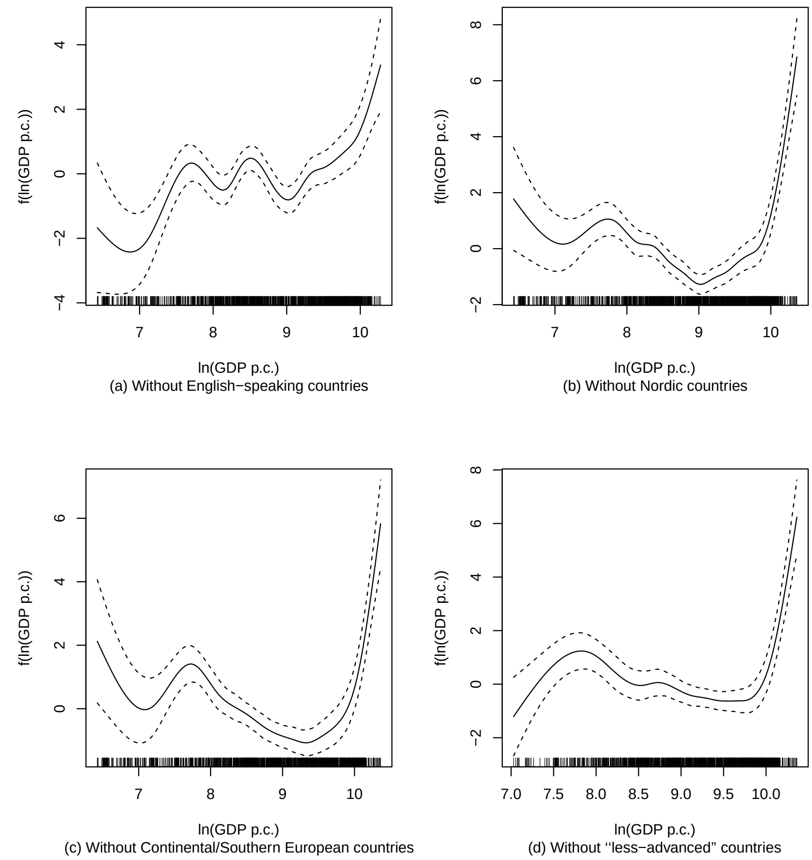


Figure B.5: The effect of leaving country groups out of the sample (annual data 1900–2010). The models include decade dummies and random country effects. See Table B.5 for model details. The solid line represents the smooth function $f(ln(GDP\ p.c.))$. The plots also show the 95% Bayesian credible intervals (dashed) and covariate values as a rug plot along the horizontal axis. The shapes of these smooths can be compared to plot (c) of Figure 1, which illustrates the corresponding random-effect specification with the entire sample.

Appendix C. 5-year average data: results using the long series

The baseline models with the 5-year average data (discussed at the end of subsection 4.2) are of the form $top1_{it} = \alpha + f(ln(GDP\ p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it}$, where i refers to country and t to 5-year period,[24] α is a constant, f is a smooth function that is described using a penalized cubic regression spline, δ_{decade} is a fixed time effect (one decade is the reference category), u_i is a country effect (omitted, fixed, or random), and ϵ_{it} is the conventional error term; the values for top 1% share and ln(GDP per capita) refer to period averages. Figure C.6 below describes the smooths f.[25] The obtained shapes of $f(ln(GDP\ p.c.))$ are close to the corresponding ones in Figure 1. Thus, changing the modeling strategy from annual to 5-year average data does not influence the overall shapes of the corresponding smooths.

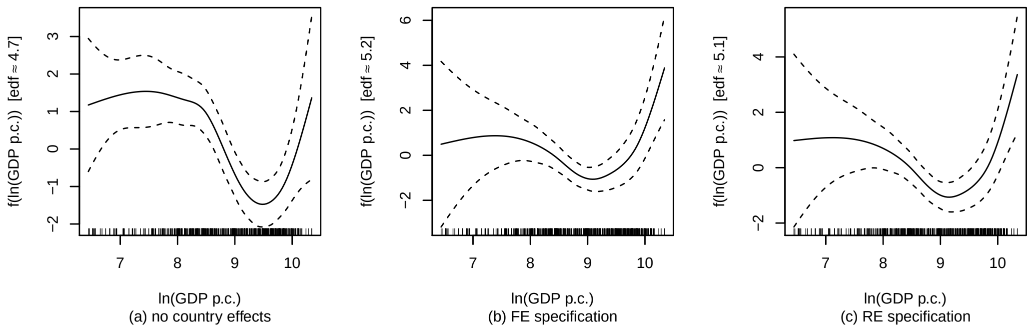


Figure C.6: Illustration of the *top1*–development relation, using 5-year average data (years 1900–2009, here N=376). The solid line represents the smooth function $f(ln(GDP\ p.c.))$. The figure also shows the 95% Bayesian credible intervals (dashed), and covariate values as a rug plot along the horizontal axis. Plot (a) represents a model without country effects, plot (b) illustrates a model with country-specific fixed effects, and plot (c) represents a model with country-specific random effects. All models include decade dummies.

[24]These periods are 1900–04, 1905–09, ..., and 2005–09.

[25]The basis dimension of the smooth before imposing identifiability constraints is $k = 10$.

References

Aghion, P., Bolton, P., 1997. A Theory of Trickle-Down Growth and Development. Review of Economic Studies 64(2), 151–172.

Ahluwalia, M.S., 1976. Inequality, Poverty and Development. Journal of Development Economics 3(4), 307–342.

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2013a. The Top 1 Percent in International and Historical Perspective. Journal of Economic Perspectives 27(3), 3–20.

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2013b. The World Top Incomes Database. Data downloaded from website: `http://g-mond.parisschoolofeconomics.eu/topincomes` (September 12, 2013).

Anand, S., Kanbur, S.M.R., 1993. Inequality and development: A critique. Journal of Development Economics 41(1), 19–43.

Atkinson, A.B., 2007. Measuring Top Incomes: Methodological Issues, in: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford. pp. 18–42.

Atkinson, A.B., Brandolini, A., 2001. Promise and Pitfalls in the Use of “Secondary” Data-Sets: Income Inequality in OECD Countries as a Case Study. Journal of Economic Literature 39(3), 771–799.

Atkinson, A.B., Piketty, T. (Eds.), 2007. Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T. (Eds.), 2010. Top Incomes: A Global Perspective. Oxford University Press, Oxford.

Atkinson, A.B., Piketty, T., Saez, E., 2011. Top incomes in the Long Run of History. Journal of Economic Literature 49(1), 3–71.

Barro, R.J., 2000. Inequality and Growth in a Panel of Countries. Journal of Economic Growth 5(1), 5–32.

Bolt, J., van Zanden, J.L., 2013. The First Update of the Maddison Project; Re-Estimating Growth Before 1820. Maddison Project Working Paper 4. Data downloaded from website: `http://www.ggdc.net/maddison/maddison-project/data/mpd_2013-01.xlsx` (September 12, 2013).

Dahan, M., Tsiddon, D., 1998. Demographic Transition, Income Distribution, and Economic Growth. Journal of Economic Growth 3(1), 29–52.

Deininger, K., Squire, L., 1996. A New Data Set Measuring Income Inequality. World Bank Economic Review 10(3), 565–591.

Deininger, K., Squire, L., 1998. New ways of looking at old issues: Inequality and growth. Journal of Development Economics 57(2), 259–287.

Desbordes, R., Verardi, V., 2012. Refitting the Kuznets curve. Economics Letters 116(2), 258–261.

Fields, G.S., 2001. Distribution and Development: A New Look at the Developing World. The MIT Press, Cambridge, Massachusetts.

Frazer, G., 2006. Inequality and Development Across and Within Countries. World Development 34(9), 1459–1481.

Galor, O., Tsiddon, D., 1996. Income Distribution and Growth: the Kuznets Hypothesis Revisited. Economica 63(250), S103–S117.

Greenwood, J., Jovanovic, B., 1990. Financial Development, Growth and the Distribution of Income. Journal of Political Economy 98(5), 1076–1107.

Hastie, T., Tibshirani, R., 1986. Generalized additive models (with discussion). Statistical Science 1(3), 297–318.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, New York.

Huang, H.–C.R., 2004. A flexible nonlinear inference to the Kuznets hypothesis. Economics Letters 84(2), 289–296.

Huang, H.–C.R., Lin, S.–C., 2007. Semiparametric Bayesian inference of the Kuznets hypothesis. Journal of Development Economics 83(2), 491–505.

Kanbur, R., 2000. Income Distribution and Development, in: Atkinson, A.B., Bourguignon F. (Eds.), Handbook of Income Distribution Vol. 1. North-Holland, Amsterdam. pp. 791–841.

Kanbur, R., Zhuang, J., 2013. Urbanization and Inequality in Asia. Asian Development Review 30(1), 131–147.

Kuznets, S., 1953. Shares of Upper Income Groups in Income and Saving. NBER Publication No. 55, New York.

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

Leigh, A., 2007. How Closely Do Top Income Shares Track Other Measures of Inequality? Economic Journal 117(524), F589–F603.

Lin, S.–C., Huang, H.–C.R., Weng, H.-W., 2006. A semi-parametric partially linear investigation of the Kuznets' hypothesis. Journal of Comparative Economics 34(3), 634–647.

List, J.A., Gallet, C.A., 1999. The Kuznets Curve: What Happens After the Inverted-U? Review of Development Economics 3(2), 200–206.

Piketty, T., 2001. Les Hauts revenus en France au 20e siècle: inégalités et redistribution, 1901–1998. B. Grasset, Paris.

Piketty, T., 2003. Income Inequality in France 1901–1998. Journal of Political Economy, 111(5), 1004–1042.

Piketty, T., Saez, E., 2006. The Evolution of Top Incomes: A Historical and International Perspective. American Economic Review 96(2), 200–205.

Robinson, S., 1976. A Note on the U Hypothesis Relating Income Inequality and Economic Development. American Economic Review 66(3), 437–440.

Roine, J., Vlachos, J., Waldenström, D., 2009. The long-run determinants of inequality: What can we learn from top income data? Journal of Public Economics 93(7–8), 974–988.

Roine, J., Waldenström, D., 2015. Long-Run Trends in the Distribution of Income and Wealth, in: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution Vol. 2A. North-Holland, Amsterdam. pp. 469–592.

United Nations, 2012. Department of Economic and Social Affairs, Population Division. World Urbanization Prospects: The 2011 Revision (POP/DB/WUP/Rev.2011/1/F2). Data downloaded from website: `http://esa.un.org/unpd/wup/CD-ROM/WUP2011-F02-Proportion_Urban.xls` (November 28, 2013).

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton FL.

World Bank, 2014a. World Development Indicators, Employment in services (% of total employment). Data downloaded from website: `http://data.worldbank.org/indicator/SL.SRV.EMPL.ZS/` (February 25, 2014).

World Bank, 2014b. World Development Indicators, Services, etc., value added (% of GDP). Data downloaded from website: `http://data.worldbank.org/indicator/NV.SRV.TETC.ZS/` (February 25, 2014).

Zhou, X., Li, K.–W., 2011. Inequality and development: Evidence from semiparametric estimation with panel data. Economics Letters 113(3), 203–207.