



# TAMPERE ECONOMIC WORKING PAPERS

## TOP-END INEQUALITY AND GROWTH: EMPIRICAL EVIDENCE

Elina Tuominen

Working Paper 108  
September 2016

SCHOOL OF MANAGEMENT  
FI-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191  
ISBN 978-952-03-0250-4

# Top-end inequality and growth: Empirical evidence

Elina Tuominen

*University of Tampere, Finland (elina.tuominen@uta.fi)*

September 2016

---

## 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, Markus Jäntti, 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).<sup>1</sup>

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

---

<sup>1</sup>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).<sup>2</sup> 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.<sup>3</sup> 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.<sup>4</sup> 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.<sup>5</sup> 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

---

<sup>2</sup>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.

<sup>3</sup>Most results are based on Gini coefficient data.

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

<sup>5</sup>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.<sup>6</sup>

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.<sup>7</sup> 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.

---

<sup>6</sup>However, the inequality measures used by Voitchovsky (2005) do not emphasize the very top of the distribution.

<sup>7</sup>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.<sup>8,9</sup>

This study finds a negative medium- to long-run relationship between top 1% income shares and subsequent growth, but this association may 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.<sup>10</sup> This study focuses on the top 1% income share series (note that this is pre-tax income). Most of the data are from “advanced”

---

<sup>8</sup>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 (Tuominen, 2016).

<sup>9</sup>Moreover, Roine et al. (2009) study top incomes and growth, but they discuss determinants of top-end inequality.

<sup>10</sup>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 the top of the 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.<sup>11</sup>

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.<sup>12</sup> 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 Wealth and Income Database (Alvaredo et al., 2015), and the data project is ongoing.

<sup>11</sup>Investment is an example of this kind of variable.

<sup>12</sup>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
$topI_t$	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
$topI_t$	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 top-end inequality and growth.<sup>13,14</sup> 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:

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

In the above presentation,  $Y_i$  is the response variable (here: average annual future growth),  $\mathbf{X}_i^*$  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,  $x_\bullet$ .

<sup>13</sup>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. This study uses an identity link and assumes normality in errors, which leads to additive models.

<sup>14</sup>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.<sup>15</sup> 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  $\epsilon_i$  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  $\beta_j$  are unknown parameters and  $b_j(x)$  are known basis functions. Using a chosen basis for  $f$  implies that we have a linear model  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ , where the model matrix  $\mathbf{X}$  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  $\boldsymbol{\beta}^T \mathbf{S}\boldsymbol{\beta}$ , where  $\mathbf{S}$  is a coefficient matrix that can be expressed in terms of the known basis functions.

The penalized regression spline fitting problem is to minimize  $\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\boldsymbol{\beta}^T \mathbf{S}\boldsymbol{\beta}$ , with respect to  $\boldsymbol{\beta}$ . The problem of estimating the degree of smoothness is a problem of estimating the smoothing parameter  $\lambda$ .<sup>16</sup> The pen-

---

<sup>15</sup>There are usually two extra conditions that specify that the second derivative of the curve should be zero at the two end knots.

<sup>16</sup>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.  $\lambda \rightarrow \infty$  results in a straight line estimate for  $f$ , and  $\lambda = 0$  leads to an unpenalized regression spline estimate.

alized least squares estimator of  $\beta$ , given  $\lambda$ , is  $\hat{\beta} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T \mathbf{y}$ . Thus, the expected value vector is estimated as  $\widehat{\mathbf{E}(\mathbf{y})} = \hat{\boldsymbol{\mu}} = \mathbf{A} \mathbf{y}$ , where  $\mathbf{A} = \mathbf{X}(\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T$  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  $\text{trace}(\mathbf{A})$ , where  $\mathbf{A}$  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).<sup>17</sup>

---

<sup>17</sup>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, and 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  $\ln(\text{GDP per capita})$  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:

$$\begin{aligned} \text{growth}_{i,t+1} &= \alpha + f_1(\text{top1}_{it}) + f_2(\ln(\text{GDP } p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it}, \quad \text{and} \\ \text{growth}_{i,t+1} &= \alpha + f_{12}(\text{top1}_{it}, \ln(\text{GDP } p.c.)_{it}) + \delta_{decade} + u_i + \epsilon_{it}, \end{aligned}$$

where  $i$  refers to a country and  $t$  to a time period,  $\alpha$  is a constant, functions  $f_{\bullet}$  refer to smooth functions,  $\delta_{decade}$  refers to a fixed decade effect (one decade is the reference category),  $u_i$  refers to a simple country-specific random effect ( $u_i \sim N(0, \sigma_u^2)$ ), and  $\epsilon_{it} \sim N(0, \sigma^2)$  is the error term; inequality and GDP variables are used as period averages (except for annual data).<sup>18</sup> The random-effect specification allows for correlation over time within countries, and the

---

<sup>18</sup>In the annual data ( $t$  refers to 1920, 1921, ..., 2007), future growth corresponds to the difference of  $\ln(\text{GDP } p.c.)$  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  $\ln(\text{GDP } p.c.)$  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  $\ln(\text{GDP } p.c.)$  values in 1930–35), and so on. The only exception is the future growth for the last 5-year period (2000–04):  $\text{growth}_{t+1}$  is calculated using  $\ln(\text{GDP } p.c.)$  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  $\ln(\text{GDP } p.c.)$  values in 2000–08 (i.e.,  $\text{growth}_{t+1}$  is not an average of ten annual growth rates but eight). Thus, in the averaged

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 the inequality–growth relationship.<sup>19</sup>

Univariate smooth functions of the top 1% share and  $\ln(\text{GDP per capita})$  are studied in models (1), (3), and (5) of Table 2. Initially, the top 1% share and  $\ln(\text{GDP per capita})$  were allowed to enter in a flexible form, but  $f(\text{top1}_t)$  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  $\text{top1}$  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  $f(\text{top1}_t)$  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  $f(\ln(\text{GDP p.c.})_t)$ .

The bivariate smooths  $f(\text{top1}_t, \ln(\text{GDP p.c.})_t)$  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.<sup>20</sup> The U shape is no longer evident at “high” levels of  $\ln(\text{GDP per capita})$ . Plots (b1)–(b2) and (c1)–(c2) of Figure 2 show clear similarities in the 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  $\ln(\text{GDP per capita})$ , as can be seen in a comparison of the slope at different levels of  $\ln(\text{GDP per capita})$ . Furthermore, Figure D.6 in Appendix D provides additional plots

---

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.

<sup>19</sup>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).

<sup>20</sup>For example, in plot (a1), look at the shape of  $f$  at  $\ln(\text{GDP p.c.}) \approx 8$  ( $\text{GDP p.c.} \approx 3000$  in 1990 int. GK\$) or at  $\ln(\text{GDP p.c.}) \approx 8.5$  ( $\text{GDP p.c.} \approx 4900$  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)
$f(top1_t)$	[ $edf \approx 1.7^a$ ] See Fig. 1 (a)	-	[linear <sup>a</sup> ] -0.137**	-	[linear <sup>a</sup> ] -0.203***	-
$f(\ln(GDP p.c.)_t)$	[ $edf \approx 2.6^a$ ] <sup>***</sup> See Fig. 1 (b)	-	[ $edf \approx 2.6^a$ ] <sup>***</sup> See Fig. 1 (c)	-	[ $edf \approx 2.7^a$ ] <sup>***</sup> See Fig. 1 (d)	-
$f(top1_t, \ln(GDP p.c.)_t)$	-	[ $edf \approx 12.7^b$ ] <sup>***</sup> See Fig. 2 (a1)–(a2)	-	[ $edf \approx 5.1^b$ ] <sup>***</sup> See Fig. 2 (b1)–(b2)	-	[ $edf \approx 5.0^b$ ] <sup>***</sup> See Fig. 2 (c1)–(c2)
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.

<sup>a</sup>Basis dimension  $k$  for the smooth before imposing identifiability constraints is  $k = 5$ .

<sup>b</sup>Basis dimension  $k$  for the smooth before imposing identifiability constraints is  $k = 5^2 = 25$  (tensor product smooth using rank 5 marginals).

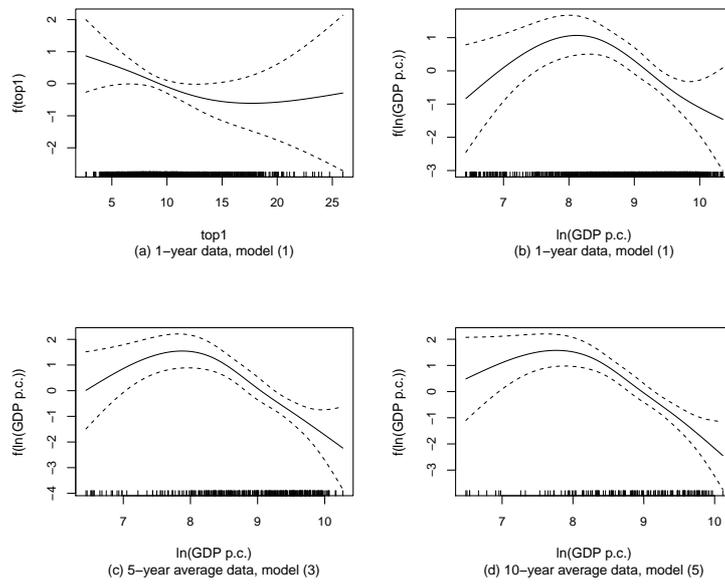
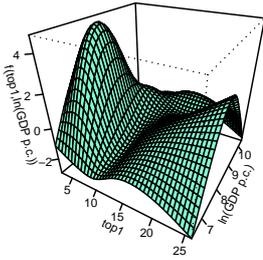
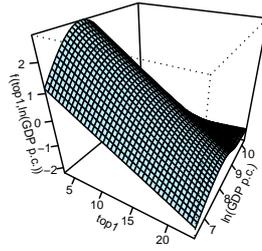


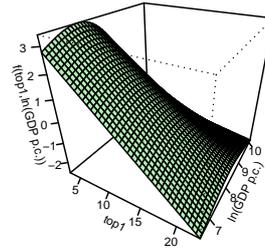
Figure 1: Visualization of the univariate smooths provided in Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as the dashed lines and the covariate values as a rug plot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



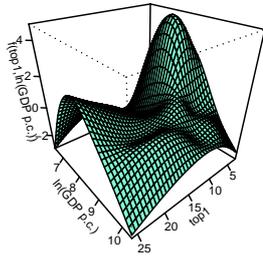
(a1) 1-year data,  
model (2), view 1



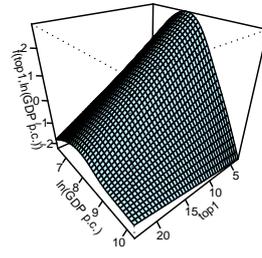
(b1) 5-year average data,  
model (4), view 1



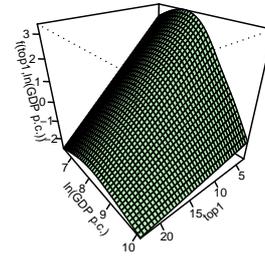
(c1) 10-year average data,  
model (6), view 1



(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  $f(\text{top1}_t, \ln(\text{GDP p.c.})_t)$  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  $\ln(\text{GDP per capita})$ ; 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.

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.<sup>21</sup> 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.<sup>22</sup> Before estimating expanded specifications with additional controls, it was checked that the results are not driven by the shorter time period (particularly excluding the war years) and the change of the GDP data source. However, details of these

---

<sup>21</sup>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.)

<sup>22</sup>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).

checks are omitted for the sake of brevity.<sup>23</sup>

### 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:

$$\begin{aligned} \text{growth}_{i,t+1} = & \alpha + f_1(\text{top1}_{it}) + f_2(\ln(\text{GDP p.c.})_{it}) + f_3(\text{schooling}_{it}) \\ & + f_4(\text{government consumption}_{it}) + f_5(\text{price level of investment}_{it}) \\ & + f_6(\text{openness}_{it}) + f_7(\text{investment}_{it}) + \delta_{\text{decade}} + u_i + \epsilon_{it}, \end{aligned}$$

where  $i$  refers to a country and  $t$  to a time period,  $\alpha$  is a constant, functions  $f_{\bullet}$  refer to smooth functions,  $\delta_{\text{decade}}$  refers to a fixed decade effect (one decade is the reference category),  $u_i$  is the country-specific random effect, and  $\epsilon_{it}$  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  $f_1(\text{top1}_{it}) + f_2(\ln(\text{GDP p.c.})_{it})$ , a bivariate smooth  $f_{12}(\text{top1}_{it}, \ln(\text{GDP p.c.})_{it})$  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  $f(\text{top1}_t)$  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  $f(\text{top1}_t, \ln(\text{GDP p.c.})_t)$  illustrate a more complex relationship.

Figure 3 provides illustrations of the smooths  $f(\text{top1}_t, \ln(\text{GDP p.c.})_t)$  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” levels of economic development, top-end inequality is in a slightly U-shaped

---

<sup>23</sup>In these checks, 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 association between *top1* and growth at “medium” levels of GDP per capita.

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)
$f(top1_t)$	[ $edf \approx 2.0^a$ ] See Fig. E.8 (a)	-	[linear <sup>a</sup> ] -0.220***	-
$f(\ln(GDP p.c.)_t)$	[ $edf \approx 2.3^a$ ]*** See Fig. E.8 (b)	-	[ $edf \approx 1.4^a$ ] See Fig. E.8 (d)	-
$f(top1_t, \ln(GDP p.c.)_t)$	-	[ $edf \approx 7.2^b$ ]*** See Fig. 3 (a1)–(a2)	-	[ $edf \approx 3.0^{b,c}$ ]*** See Fig. 3 (b1)–(b2)
$f(gov't\ consumption_t)$	[linear <sup>a</sup> ] 0.155***	[linear <sup>a</sup> ] 0.158***	[linear <sup>a</sup> ] 0.108**	[linear <sup>a</sup> ] 0.097*
$f(schooling_t)$	[linear <sup>a</sup> ] 0.093	[linear <sup>a</sup> ] 0.180	[ $edf \approx 2.8^a$ ] See Fig. E.8 (e)	[ $edf \approx 2.9^a$ ] See Fig. E.8 (h)
$f(price\ of\ investment_t)$	[linear <sup>a</sup> ] -0.015***	[linear <sup>a</sup> ] -0.013**	[ $edf \approx 2.9^a$ ]** See Fig. E.8 (f)	[ $edf \approx 2.7^a$ ] See Fig. E.8 (i)
$f(openness_t)$	[linear <sup>a</sup> ] 0.003	[linear <sup>a</sup> ] 0.005*	[ $edf \approx 1.7^a$ ] See Fig. E.8 (g)	[linear <sup>a</sup> ] 0.007**
$f(investment_t)$	[ $edf \approx 1.7^a$ ] See Fig. E.8 (c)	[linear <sup>a</sup> ] 0.031	[linear <sup>a</sup> ] 0.016	[linear <sup>a</sup> ] 0.020
AIC	749	752	319	321

\*\*\*, \*\*, \*, ' indicate significance at the 1, 5, 10, and 15% levels, respectively.

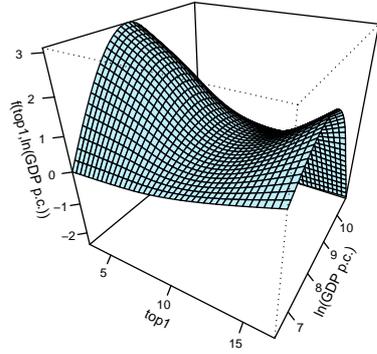
The  $p$ -values for parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are reported. The smooth terms' significance levels are based on approximate  $p$ -values.

All specifications include decade dummies and random country-specific effects.

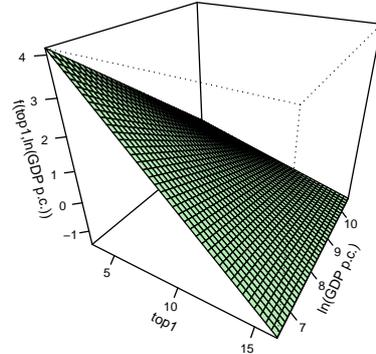
<sup>a</sup>Basis dimension  $k$  for the smooth before imposing identifiability constraints is  $k = 5$ .

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

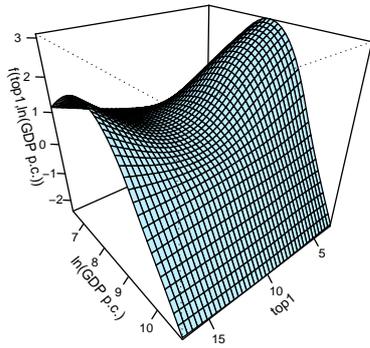
<sup>c</sup>With 3 degrees of freedom, the tensor product smooth refers to  $\theta_1 top1_t + \theta_2 \ln(GDP p.c.)_t + \theta_3 top1_t \ln(GDP p.c.)_t$ , where  $\theta_\bullet$  are coefficients. When model (4) is estimated using this specification in place of  $f(top1_t, \ln(GDP p.c.)_t)$ , the 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.



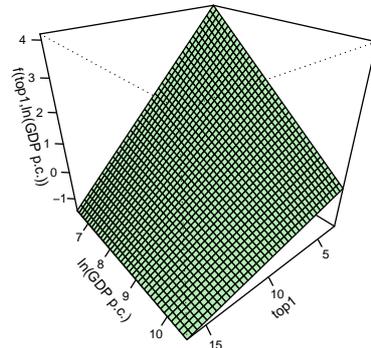
(a1) 5-year average data,  
model (2), view 1



(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.

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 (see also note c to Table 3). 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 (recall discussion at the beginning of section 3.2 and footnote 23). The next subsection provides additional 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.<sup>24</sup> Results for the Conti-

---

<sup>24</sup>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

mental 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  $\ln(\text{GDP per capita})$ . 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.<sup>25</sup> 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 1% 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 may become 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.

---

<sup>25</sup>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.<sup>26</sup> 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.<sup>27</sup>
5. **Finland:** Jäntti et al. (2010): Table 8A.2 (taxable income/population), years 1920–1992; Riihelä, M. (2011): updated figures, years 1993–2008.<sup>28</sup>
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.<sup>29</sup>
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.<sup>30</sup>
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.<sup>31</sup> (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.<sup>32</sup>
22. **United Kingdom:** Atkinson et al. (2010): Table 13A.2, years 1937, 1949–2005.<sup>33</sup>
23. **United States:** Saez (2010): Excel file, Table A1, years 1920–2008.<sup>34</sup>

---

<sup>26</sup>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 (Alvaredo et al., 2011). Currently, the updated data are available in the World Wealth and Income Database (Alvaredo et al., 2015).

<sup>27</sup>For more information and the original series, see Piketty and Qian (2010).

<sup>28</sup>Figures 1993–2008 received directly from Marja Riihelä by email (Feb 11, 2011).

<sup>29</sup>For more information and the original series, see Alvaredo and Pisano (2010).

<sup>30</sup>For more information and the original series, see Atkinson and Leigh (2007b).

<sup>31</sup>For more information and the original series, see Atkinson (2010).

<sup>32</sup>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.

<sup>33</sup>For more information and the original series, see Atkinson (2007b).

<sup>34</sup>For more information and the original series, see Piketty and Saez (2007).

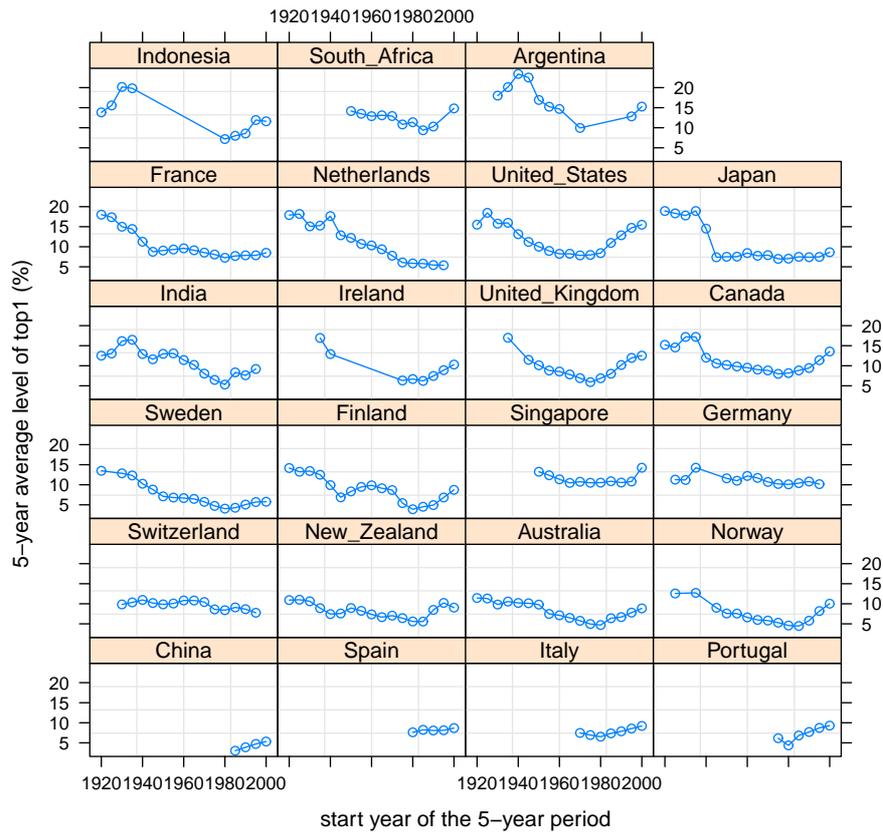


Figure A.4: Top 1% income shares for each country (5-year average data used in models of Table 2; the time periods are 1920–24, 1925–29, ..., and 2000–04). Data sources: see list of countries in this appendix.

## 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.

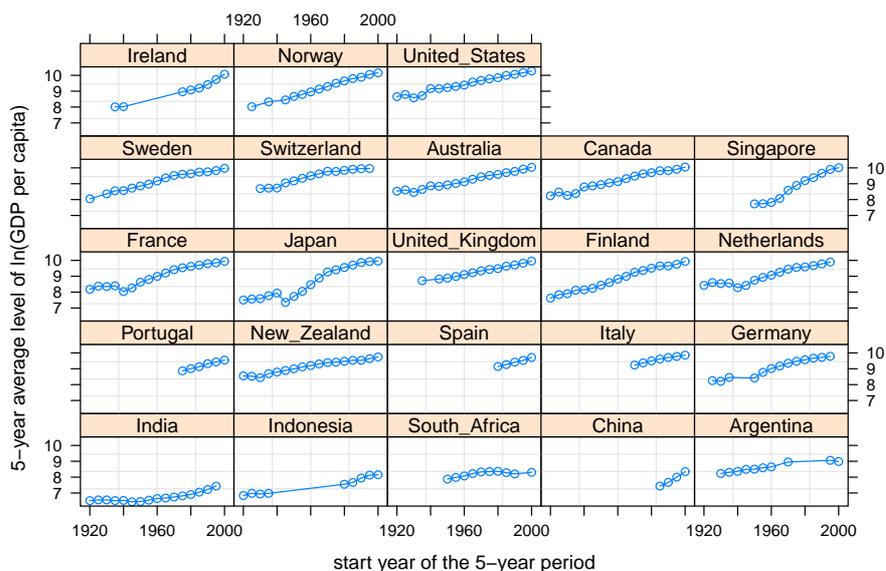


Figure B.5: Level of economic development for each country (5-year average data used in models of Table 2; the time periods are 1920–24, 1925–29, ..., and 2000–04). Data source: Maddison (2010).

## 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  $\alpha_i$  and  $\delta_l$  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  $\alpha_i$  to vary smoothly with  $z$ . We can write:

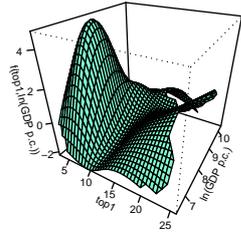
$$\alpha_i(z) = \sum_{l=1}^L \delta_{il} d_l(z),$$

and the tensor product basis construction gives:

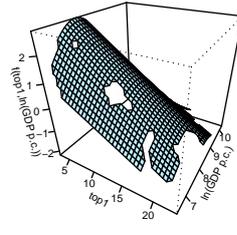
$$f_{xz}(x, z) = \sum_{i=1}^I \sum_{l=1}^L \delta_{il} d_l(z) a_i(x).$$

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

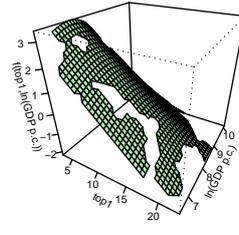
## Appendix D. Additional information, simplified models



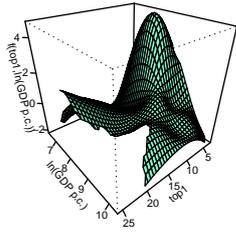
(a1) 1-year data,  
model (2), view 1



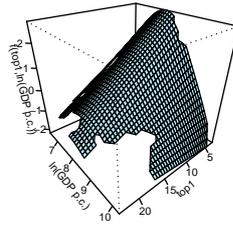
(b1) 5-year average data,  
model (4), view 1



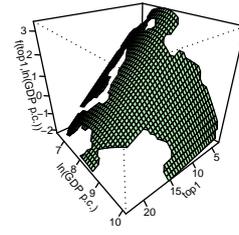
(c1) 10-year average data,  
model (6), view 1



(a2) 1-year data,  
model (2), view 2



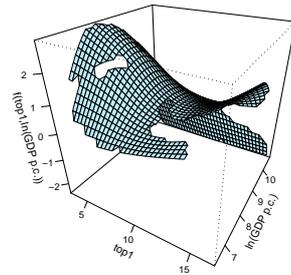
(b2) 5-year average data,  
model (4), view 2



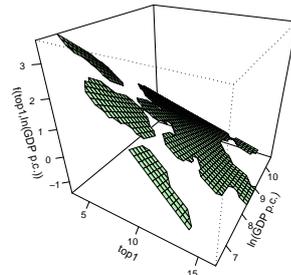
(c2) 10-year average data,  
model (6), view 2

Figure D.6: Visualization of the simplified models: smooths  $f(\text{top}1_t, \ln(\text{GDP } p.c.)_t)$  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  $\ln(\text{GDP per capita})$ ; 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  $\ln(\text{GDP per capita})$  are excluded: the grid has been scaled into the unit square along with  $\text{top}1$  and GDP variables; grid nodes more than 0.1 from the predictor variables are excluded. Compare to Figure 2.

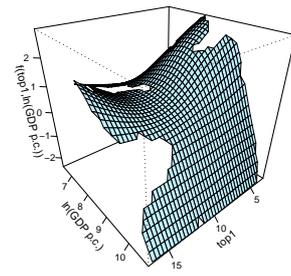
## Appendix E. Additional information, expanded models



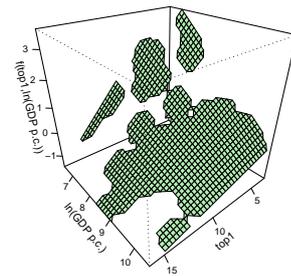
(a1) 5-year average data,  
model (2), view 1



(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  $\ln(\text{GDP per capita})$ ; 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  $\ln(\text{GDP per capita})$  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.

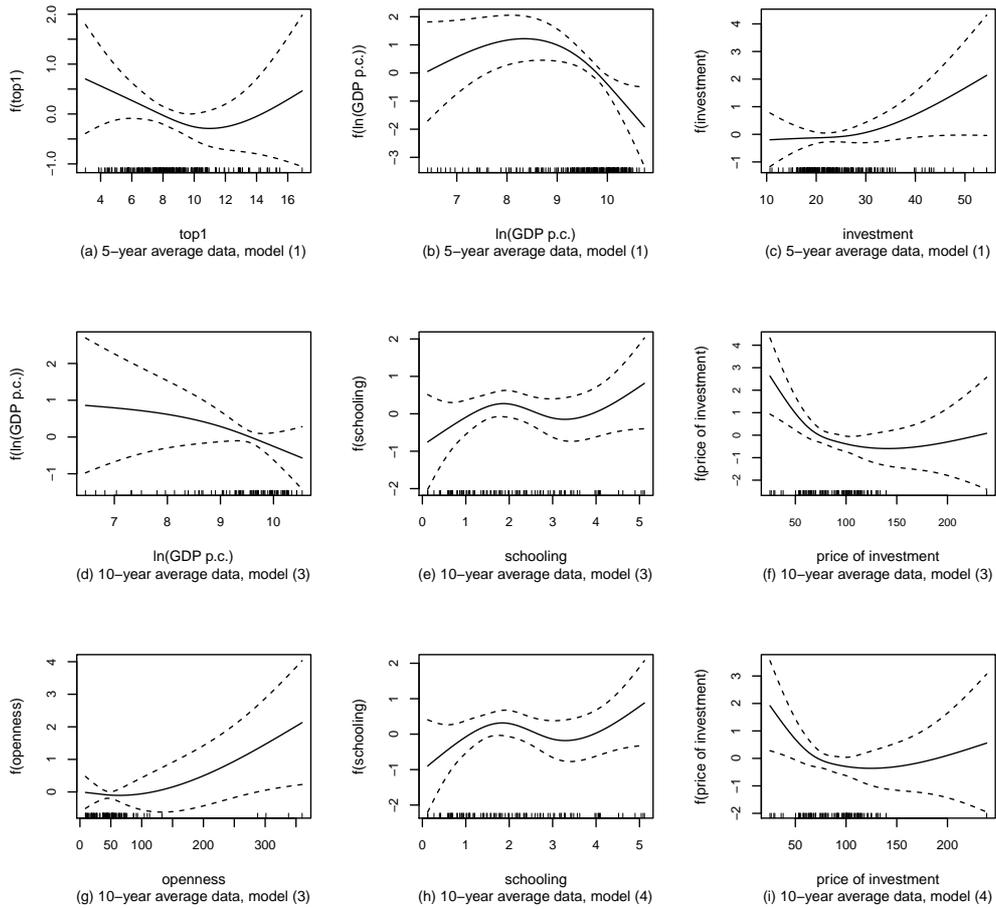


Figure E.8: Visualization of the expanded models: univariate smooths provided in Table 3 (data from the 1950s onward; GDP data from PWT 7.0). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as the dashed lines and the covariate values as 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., Atkinson, A.B., Piketty, T., Saez, E., & Zucman, G. (2015). The World Wealth and Income Database. Website: <http://www.wid.world>.
- 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](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.ggd.net/MADDISON/Historical\\_Statistics/vertical-file\\_02-2010.xls](http://www.ggd.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.
- Tuominen, E. (2016). Changes or levels? Reassessment of the relationship between top-end inequality and growth. *Tampere Economic Working Papers* No. 109 (September 2016), University of Tampere.
- 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.