



# TAMPERE ECONOMIC WORKING PAPERS

CHANGES OR LEVELS?  
REASSESSMENT OF THE RELATIONSHIP  
BETWEEN TOP-END INEQUALITY AND GROWTH

Elina Tuominen

Working Paper 109  
September 2016

SCHOOL OF MANAGEMENT  
FI-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191  
ISBN 978-952-03-0251-1

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

Elina Tuominen

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

September 2016

---

## 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 Markus Jääntti, Jukka Pirttilä, Olli Ropponen, Hannu Tanninen, Matti Tuomala, and Jari Vainiomäki, 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.<sup>1</sup> 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.<sup>2</sup> 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-

---

<sup>1</sup>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)

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

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-

---

<sup>3</sup>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.

<sup>4</sup>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.

<sup>5</sup>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.<sup>6</sup> 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.<sup>7</sup> 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 (2016). The current study augments the preceding investigation by exploring the *change* in this

---

<sup>6</sup>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.”

<sup>7</sup>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 may become weaker 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.<sup>8</sup> This study focuses on the top 1% (note that this is pre-tax income). The

---

<sup>8</sup>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). The updated data used to be available in the World Top Incomes Database by Alvaredo et al. (2012). The top income project is ongoing, and the updated data are now available in the World Wealth and Income Database by Alvaredo et al. (2016).

top 1% income shares (*top1*) in 25 countries from the 1920s to the 2000s 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)	<i>N</i>	min	mean	max
<i>top1<sub>t</sub></i>	275	3.9	9.6	23.4
<i>top1<sub>t</sub> - top1<sub>t-1</sub></i>	275	-7.2	-0.2	4.6
<i>ln(GDP p.c.)<sub>t</sub></i>	275	6.4	8.9	10.3
<i>growth<sub>t+1</sub></i>	275	-15.2	2.4	16.1
Expanded models (data from the 1950s onward)	<i>N</i>	min	mean	max
<i>top1<sub>t</sub></i>	210	3.9	8.5	16.9
<i>top1<sub>t</sub> - top1<sub>t-1</sub></i>	210	-6.9	0.0	3.4
<i>ln(GDP p.c.)<sub>t</sub></i>	210	6.4	9.5	10.7
<i>government consumption<sub>t</sub></i>	210	4.0	9.4	18.3
<i>investment<sub>t</sub></i>	210	10.6	24.0	54.4
<i>price level of investment<sub>t</sub></i>	210	18.9	87.0	294.6
<i>openness<sub>t</sub></i>	210	8.0	64.7	386.3
<i>secondary schooling<sub>t</sub></i>	210	0.1	2.2	5.4
<i>tertiary schooling<sub>t</sub></i>	210	0.0	0.3	1.7
<i>growth<sub>t+1</sub></i>	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.<sup>9</sup> 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.<sup>10,11</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 log growth in the subsequent period),  $\mathbf{X}_i^*$  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,  $x_\bullet$ .

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

---

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

<sup>10</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>11</sup>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.<sup>12</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  $f(x) = \sum_{j=1}^k \beta_j b_j(x)$ , 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.

Accordingly, 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>13</sup> The penalized least squares estimator of  $\boldsymbol{\beta}$ , given  $\lambda$ , is  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T \mathbf{y}$ . 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.

---

<sup>12</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>13</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.

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  $\text{trace}(\mathbf{A})$ , 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  $\text{edf}=2.1$  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>14</sup>

## **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.

---

<sup>14</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/>).

#### 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  $\ln(\text{GDP per capita})$  as covariates, and the dependent variable is the future log growth of GDP per capita; the GDP per capita data of Maddison (2010) are 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:

$$\begin{aligned} \text{growth}_{i,t+1} &= \alpha + f_1(\text{top1}_{it}) + f_2(\text{top1}_{it} - \text{top1}_{i,t-1}) + f_3(\ln(\text{GDP p.c.})_{it}) \\ &\quad + \delta_{\text{decade}} + u_i + \epsilon_{it}, \\ \text{growth}_{i,t+1} &= \alpha + f_{13}(\text{top1}_{it}, \ln(\text{GDP p.c.})_{it}) + f_2(\text{top1}_{it} - \text{top1}_{i,t-1}) \\ &\quad + \delta_{\text{decade}} + u_i + \epsilon_{it}, \quad \text{and} \\ \text{growth}_{i,t+1} &= \alpha + f_2(\text{top1}_{it} - \text{top1}_{i,t-1}) + f_3(\ln(\text{GDP p.c.})_{it}) \\ &\quad + \delta_{\text{decade}} + u_i + \epsilon_{it}, \end{aligned}$$

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),  $u_i$  refers to a 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 per capita variables are used as period averages.<sup>15</sup> The random-effect spec-

---

<sup>15</sup>In annual data, growth would refer to the difference of  $\ln(\text{GDP p.c.})$  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  $\ln(\text{GDP p.c.})$  values in 1930–35), and the change in  $\text{top1}$  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  $\ln(\text{GDP p.c.})$  values in 2005–08 (i.e.,  $\text{growth}_{t+1}$  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  $\ln(\text{GDP p.c.})$  values in 2000–08 (i.e.,  $\text{growth}_{t+1}$  is not an average of ten annual growth rates but eight). Thus, 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.<sup>16</sup> The second specification with a bivariate smooth  $f(\text{top}1_t, \ln(\text{GDP } p.c.)_t)$  allows for a very flexible interaction between the level of top-end inequality and the level of economic development—the specification stems from Tuominen (2016). 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.

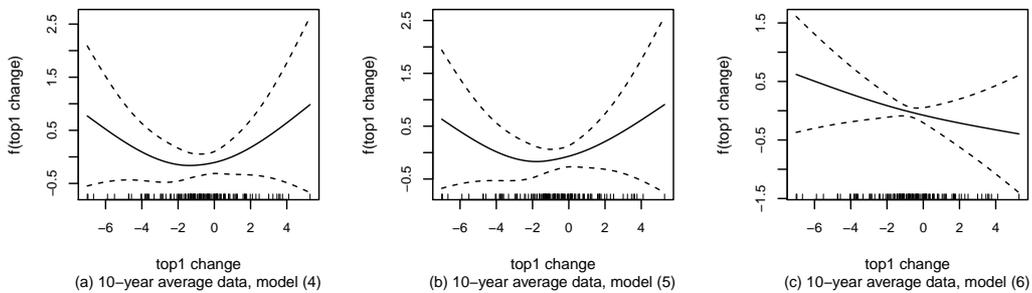


Figure 1: Visualization of the simplified models: smooths  $f(\text{top}1_t - \text{top}1_{t-1})$  provided in Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as dashed 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.,  $f(\text{top}1_t - \text{top}1_{t-1})$ ) 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

<sup>16</sup>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 <sup>a</sup> ] -0.146**	-	-	[linear <sup>a</sup> ] -0.197**	-	-
$f(top1_t - top1_{t-1})$	[linear <sup>a</sup> ] 0.145	[linear <sup>a</sup> ] 0.135	[linear <sup>a</sup> ] 0.070	[ $edf \approx 2.0^a$ ] See Fig. 1 (a)	[ $edf \approx 1.9^a$ ] See Fig. 1 (b)	[ $edf \approx 1.2^a$ ] See Fig. 1 (c)
$f(\ln(GDP p.c.)_t)$	[ $edf \approx 2.5^a$ ]*** See Fig. D.8 (a)	-	[ $edf \approx 2.5^a$ ]*** See Fig. D.8 (b)	[ $edf \approx 2.6^a$ ]*** See Fig. D.8 (c)	-	[ $edf \approx 2.6^a$ ]*** See Fig. D.8 (d)
$f(top1_t, \ln(GDP p.c.)_t)$	-	[ $edf \approx 5.1^b$ ]*** See Fig. 2 (a1)–(a2)	-	-	[ $edf \approx 3.8^b$ ]*** See Fig. 2 (b1)–(b2)	-
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.

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

<sup>b</sup>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.<sup>17</sup> 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.

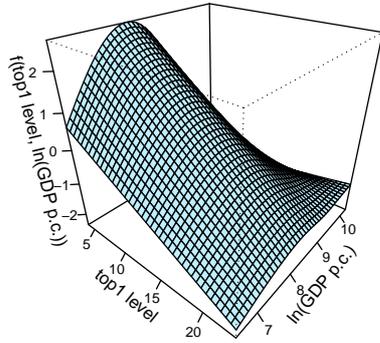
In the current sample, 18 out of the 25 countries are “advanced,” and the other countries comprise a heterogeneous group. As a small check, these “advanced” countries were studied separately to see whether the other seven countries affected the main results above. Specifications similar to models (1)–(2) and (4)–(5) of Table 2 were fitted for this subset of the data.<sup>18</sup> 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.

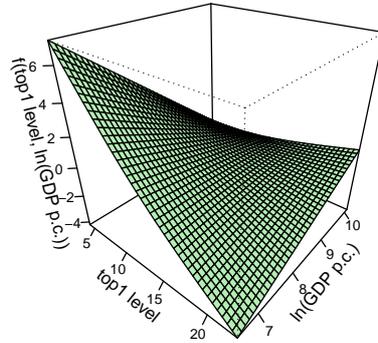
---

<sup>17</sup>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  $-0.146top1_t + 0.145(top1_t - top1_{t-1}) \approx -0.145top1_{t-1}$ . 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.

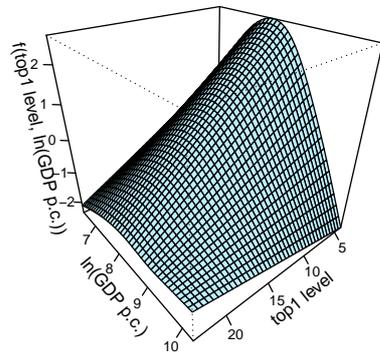
<sup>18</sup>Australia, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States ( $N=212$  in the 5-year data;  $N=96$  in the 10-year data).



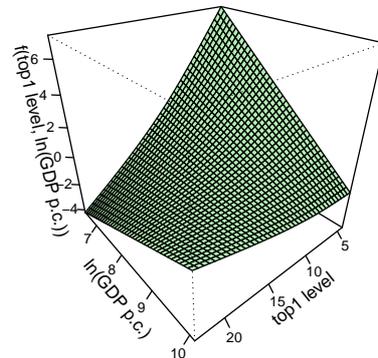
(a1) 5-year average data, model (2),  
view 1



(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  $f(\text{top}1_t, \ln(\text{GDP } p.c.)_t)$  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  $\ln(\text{GDP } p.c.)$ ; 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).<sup>19</sup> 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.<sup>20</sup>

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

---

<sup>19</sup>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).

<sup>20</sup>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).<sup>21</sup> 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  $\ln(\text{GDP per capita})$ ; 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 (2016). 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.<sup>22</sup>

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

---

<sup>21</sup>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.

<sup>22</sup>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 <sup>a</sup> ] -0.065	-	-	[linear <sup>a</sup> ] -0.161**	-	-
$f(top1_t - top1_{t-1})$	[linear <sup>a</sup> ] 0.048	[linear <sup>a</sup> ] 0.076	[linear <sup>a</sup> ] 0.017	[linear <sup>a</sup> ] 0.133	[linear <sup>a</sup> ] 0.130	[linear <sup>a</sup> ] 0.048
$f(\ln(GDP p.c.)_t)$	[ $edf \approx 2.6^a$ ] <sup>***</sup> See Fig. E.10 (a)	-	[ $edf \approx 2.6^a$ ] <sup>***</sup> See Fig. E.10 (d)	[linear <sup>a</sup> ] -1.274 <sup>***</sup>	-	[ $edf \approx 2.0^a$ ] <sup>***</sup> See Fig. E.10 (h)
$f(top1_t, \ln(GDP p.c.)_t)$	-	[ $edf \approx 8.8^b$ ] <sup>***</sup> See Fig. 3 (a1)–(a2)	-	-	[ $edf \approx 3.0^{b,c}$ ] <sup>***</sup> See Fig. 3 (b1)–(b2)	-
$f(government\ consumption_t)$	[linear <sup>a</sup> ] 0.180 <sup>***</sup>	[linear <sup>a</sup> ] 0.187 <sup>***</sup>	[linear <sup>a</sup> ] 0.193 <sup>***</sup>	[linear <sup>a</sup> ] 0.256 <sup>***</sup>	[linear <sup>a</sup> ] 0.234 <sup>***</sup>	[linear <sup>a</sup> ] 0.335 <sup>***</sup>
$f(price\ level\ of\ investment_t)$	[linear <sup>a</sup> ] -0.006	[linear <sup>a</sup> ] -0.012*	[linear <sup>a</sup> ] -0.007	[linear <sup>a</sup> ] 0.000	[linear <sup>a</sup> ] 0.003	[ $edf \approx 1.3^a$ ] See Fig. E.10 (i)
$f(openness_t)$	[linear <sup>a</sup> ] 0.008**	[linear <sup>a</sup> ] 0.008**	[linear <sup>a</sup> ] 0.008**	[linear <sup>a</sup> ] 0.005'	[linear <sup>a</sup> ] 0.006*	[linear <sup>a</sup> ] 0.008*
$f(investment_t)$	[linear <sup>a</sup> ] -0.004	[linear <sup>a</sup> ] -0.011	[linear <sup>a</sup> ] -0.008	[linear <sup>a</sup> ] 0.085 <sup>***</sup>	[linear <sup>a</sup> ] 0.079 <sup>***</sup>	[linear <sup>a</sup> ] 0.050
$f(secondary\ schooling_t)$	[ $edf \approx 3.0^a$ ] <sup>**</sup> See Fig. E.10 (b)	[ $edf \approx 2.7^a$ ] See Fig. E.10 (c)	[ $edf \approx 3.0^a$ ] <sup>**</sup> See Fig. E.10 (e)	[ $edf \approx 3.4^a$ ] <sup>**</sup> See Fig. E.10 (f)	[ $edf \approx 3.3^a$ ] <sup>*</sup> See Fig. E.10 (g)	[ $edf \approx 3.1^a$ ] <sup>**</sup> See Fig. E.10 (j)
$f(tertiary\ schooling_t)$	[linear <sup>a</sup> ] 0.930	[linear <sup>a</sup> ] 1.105	[linear <sup>a</sup> ] 0.810	[linear <sup>a</sup> ] 1.353	[linear <sup>a</sup> ] 1.232	[linear <sup>a</sup> ] 1.386
AIC	784	778	781	325	325	306

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

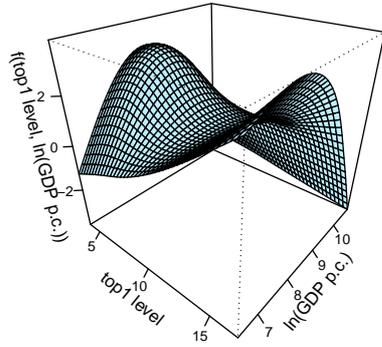
The  $p$ -values for the parametric terms are calculated using the Bayesian estimated covariance matrix of the parameter estimators; only the significance levels are provided. The smooth terms' significance levels are based on approximate  $p$ -values.

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

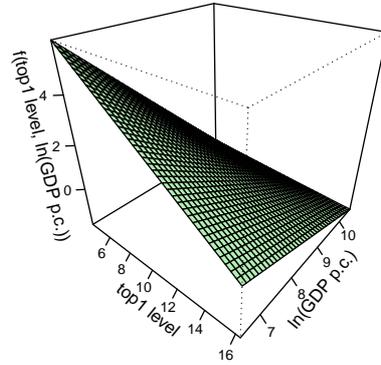
<sup>a</sup>The basis dimension  $k$  for the smooth before imposing identifiability constraints is  $k = 5$ .

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

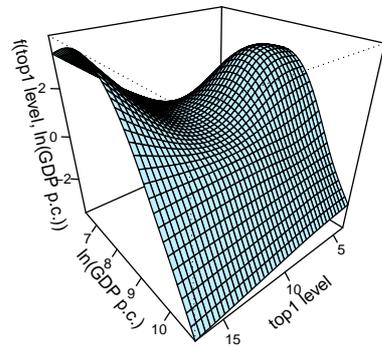
<sup>c</sup>With just 3 degrees of freedom, the tensor product smooth refers to  $\theta_1 top1_t + \theta_2 \ln(GDP p.c.)_t + \theta_3 top1_t \ln(GDP p.c.)_t$ , where  $\theta_\bullet$  are coefficients. When model (5) is estimated using this form in place of  $f(top1_t, \ln(GDP p.c.)_t)$ , the coefficients 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.



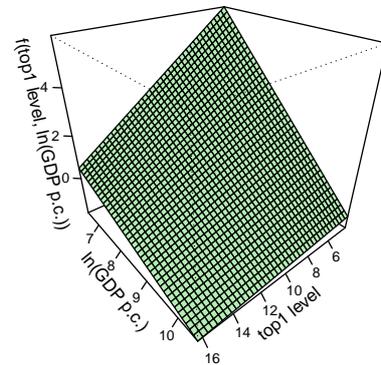
(a1) 5-year average data, model (2),  
view 1



(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  $f(\text{top}1_t, \ln(\text{GDP } p.c.)_t)$  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  $\ln(\text{GDP per capita})$ ; 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.<sup>23</sup> 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.<sup>24</sup>

These results help explain the shape of the smooth  $f(\text{top1}_t, \ln(\text{GDP } p.c.)_t)$  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

---

<sup>23</sup>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.

<sup>24</sup>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 (2016), 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.<sup>25</sup> 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.<sup>26</sup> However, different timing of the

---

<sup>25</sup>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.

<sup>26</sup>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 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 functions  $f(\text{top}1_t, \ln(\text{GDP } p.c.)_t)$  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 ( $\ln(\text{GDP } p.c.) < 8$ , 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.<sup>27</sup>

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.<sup>28</sup> 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

---

<sup>27</sup>As a further check, India and Indonesia (six observations in total) were excluded from the analysis: the remaining 16 wealthy countries (all had  $\ln(\text{GDP } p.c.) > 9$ ) 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.

<sup>28</sup>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).

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 <sup>a</sup> ] 0.005	-	-	-
$f(top1_t - top1_{t-1})$	[linear <sup>a</sup> ] -0.183	[ $edf \approx 1.1^a$ ]	[ $edf \approx 1.3^a$ ]	[linear <sup>a</sup> ] -0.133
$f(\ln(GDP\ p.c.)_t)$	[linear <sup>a</sup> ] 0.446*	[ $edf \approx 1.3^a$ ]	-	-
$f(top1_t, \ln(GDP\ p.c.)_t)$	-	-	[ $edf \approx 3.0^{b,c}$ ]* See Fig. 4 (a1)–(a2)	[ $edf \approx 3.0^{b,d}$ ]* See Fig. 4 (b1)–(b2)
$f(Gini_t)$	[linear <sup>a</sup> ] 0.080**	-	-	[linear <sup>a</sup> ] 0.067*
$f(Gini_t - Gini_{t-1})$	[linear <sup>a</sup> ] 0.067	[linear <sup>a</sup> ] 0.116**	[linear <sup>a</sup> ] 0.124**	[linear <sup>a</sup> ] 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.

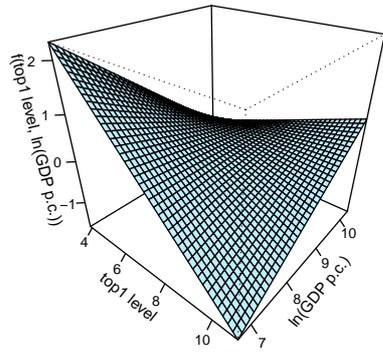
*Note:* All models include decade dummies, random country effects, and controls for government consumption, price level of investment, openness, investment, average years of secondary schooling, and average years of tertiary schooling (almost all controls enter the models linearly).

<sup>a</sup>The basis dimension  $k$  for the smooth before imposing identifiability constraints is  $k = 3$ .

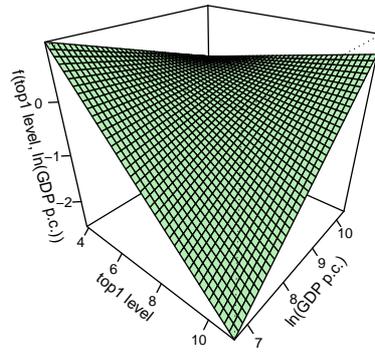
<sup>b</sup>The basis dimension  $k$  for the smooth before imposing identifiability constraints is  $k = 3^2 = 9$  (tensor product smooth using rank 3 marginals).

<sup>c</sup>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.

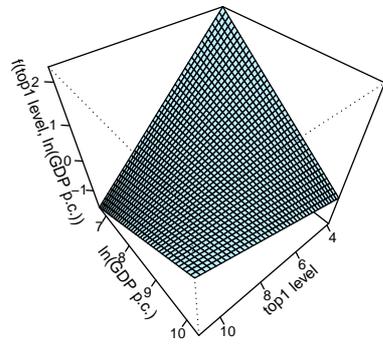
<sup>d</sup>With just 3 degrees of freedom, the tensor product smooth refers to  $\theta_1 top1_t + \theta_2 \ln(GDP\ p.c.)_t + \theta_3 top1_t \ln(GDP\ p.c.)_t$ , where  $\theta_\bullet$  are coefficients. When model (4) is estimated using this form in place of  $f(top1_t, \ln(GDP\ p.c.)_t)$ , the coefficients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



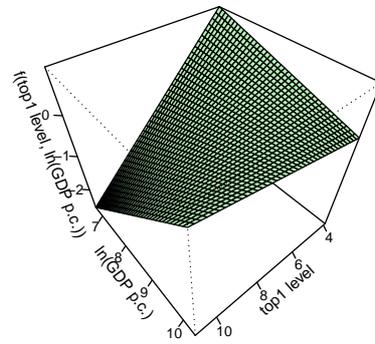
(a1) model (3), view 1



(b1) model (4), view 1



(a2) model (3), view 2



(b2) model (4), view 2

Figure 4: Visualizations of the smooths  $f(\text{top1}_t, \ln(\text{GDP p.c.})_t)$  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  $\ln(\text{GDP per capita})$ ; the vertical axis has the smooth  $f$ .

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 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 become weaker 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 (2016). 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. The *top1* series in the 5-year average data are plotted in Figure A.5. For more information on the series, see the new version of the database (Alvaredo et al., 2016) and also Atkinson and Piketty (2007, 2010).

Country	Source
Argentina	Alvaredo et al. (2012)
Australia	Alvaredo et al. (2012)
Canada	Alvaredo et al. (2012) <sup>a</sup>
China	Alvaredo et al. (2012)
Denmark	Alvaredo et al. (2012)
Finland	Alvaredo et al. (2012) and Marja Riihelä (2011) <sup>b</sup>
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) <sup>c</sup>
United Kingdom	Alvaredo et al. (2012)
United States	Alvaredo et al. (2012)

Additional notes:

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

<sup>b</sup>Updated Finnish data covering years from 1993 onward. Received directly from Marja Riihelä by email (Feb 11, 2011).

<sup>c</sup>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.

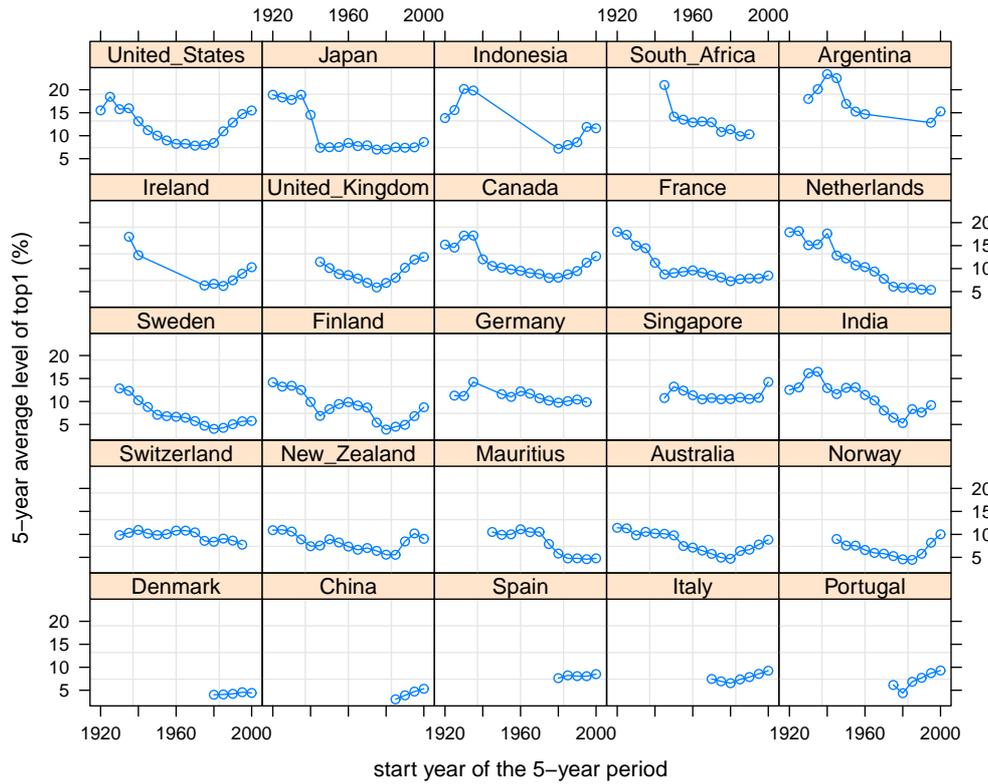


Figure A.5: Top 1% income shares for each country (5-year average data used in the models of Table 2; the time periods  $t$  are 1925–29, 1930–34, ..., and 2000–04; values from period 1920–24 are also plotted if they have been used in the construction of the “change in top 1% share” variable). Data source: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

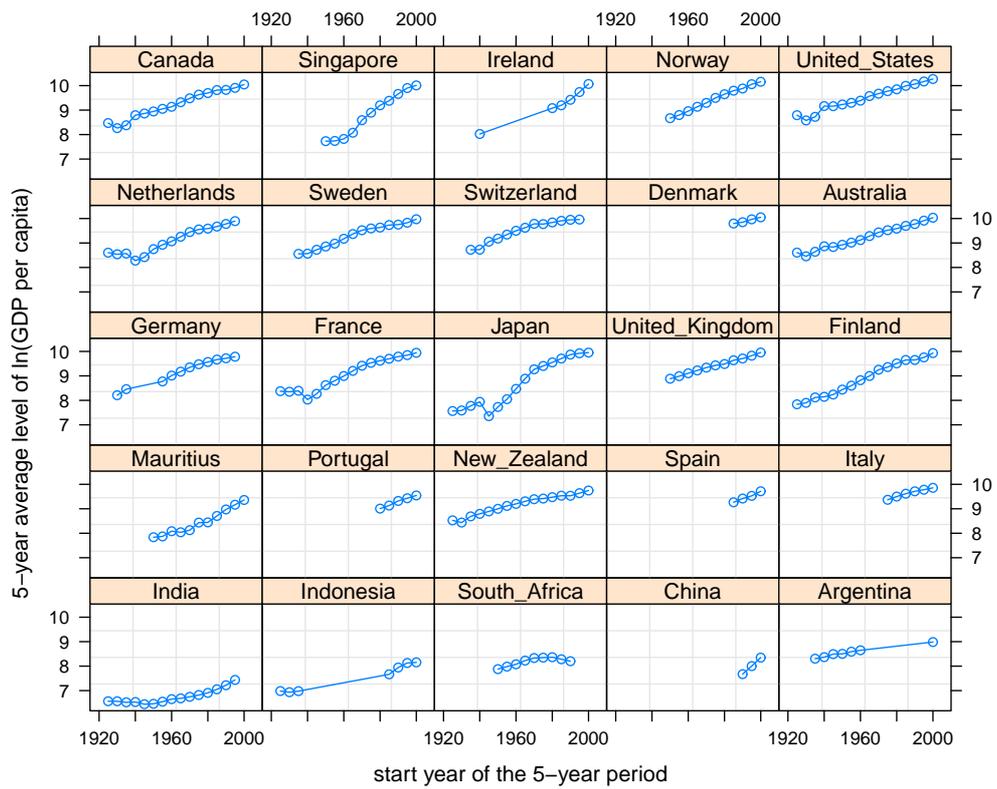


Figure B.6: Level of economic development for each country (5-year average data used in the models of Table 2; the time periods  $t$  are 1925–29, 1930–34, ..., and 2000–04). Data source: Maddison (2010).

## 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  $\alpha_i$  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:

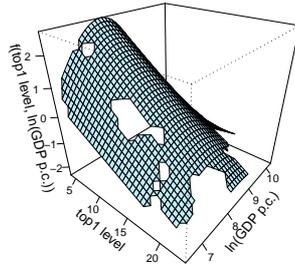
$$\alpha_i(z) = \sum_{l=1}^L \delta_{il} d_l(z),$$

and the tensor product basis construction gives:

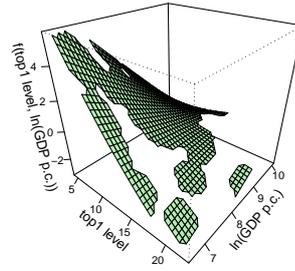
$$f_{xz}(x, z) = \sum_{i=1}^I \sum_{l=1}^L \delta_{il} d_l(z) a_i(x).$$

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

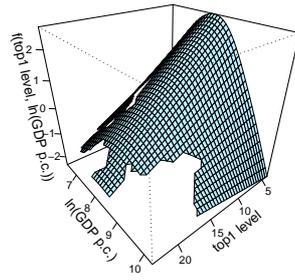
## Appendix D. Additional plots: long series from the 1920s



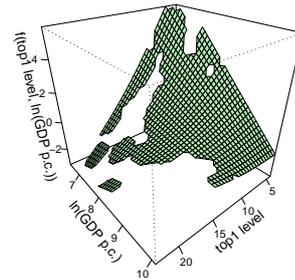
(a1) 5-year average data, model (2),  
view 1



(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  $f(\text{top1}_t, \ln(\text{GDP p.c.})_t)$  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  $\ln(\text{GDP per capita})$ ; 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  $\ln(\text{GDP per capita})$  are excluded: the grid has been scaled into the unit square along with  $\text{top1}$  and GDP variables; grid nodes more than 0.1 from the predictor variables are excluded. Compare to Figure 2.

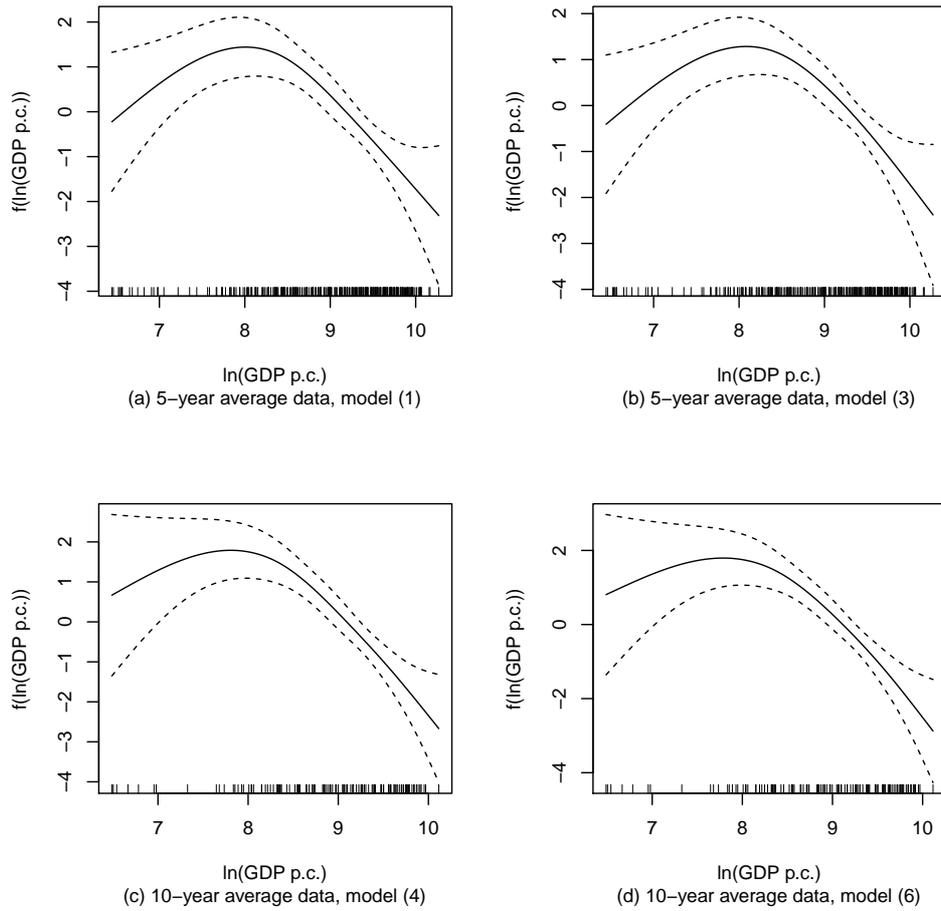
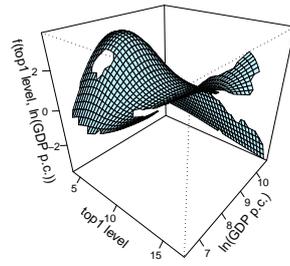
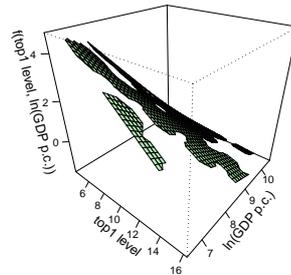


Figure D.8: Visualization of the simplified models' smooths  $f(\ln(\text{GDP p.c.}))_t$  provided in Table 2 (data from the 1920s onward; GDP data from Maddison, 2010). Each plot presents the smooth function as a solid line. The plots also show the 95% Bayesian credible intervals as dashed lines and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

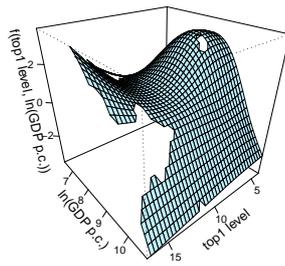
## Appendix E. Additional plots: series from the 1950s



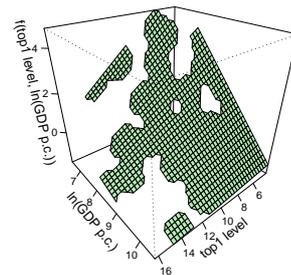
(a1) 5-year average data, model (2),  
view 1



(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  $f(\text{top}1_t, \ln(\text{GDP } p.c.)_t)$  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  $\ln(\text{GDP per capita})$ ; 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 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 3.

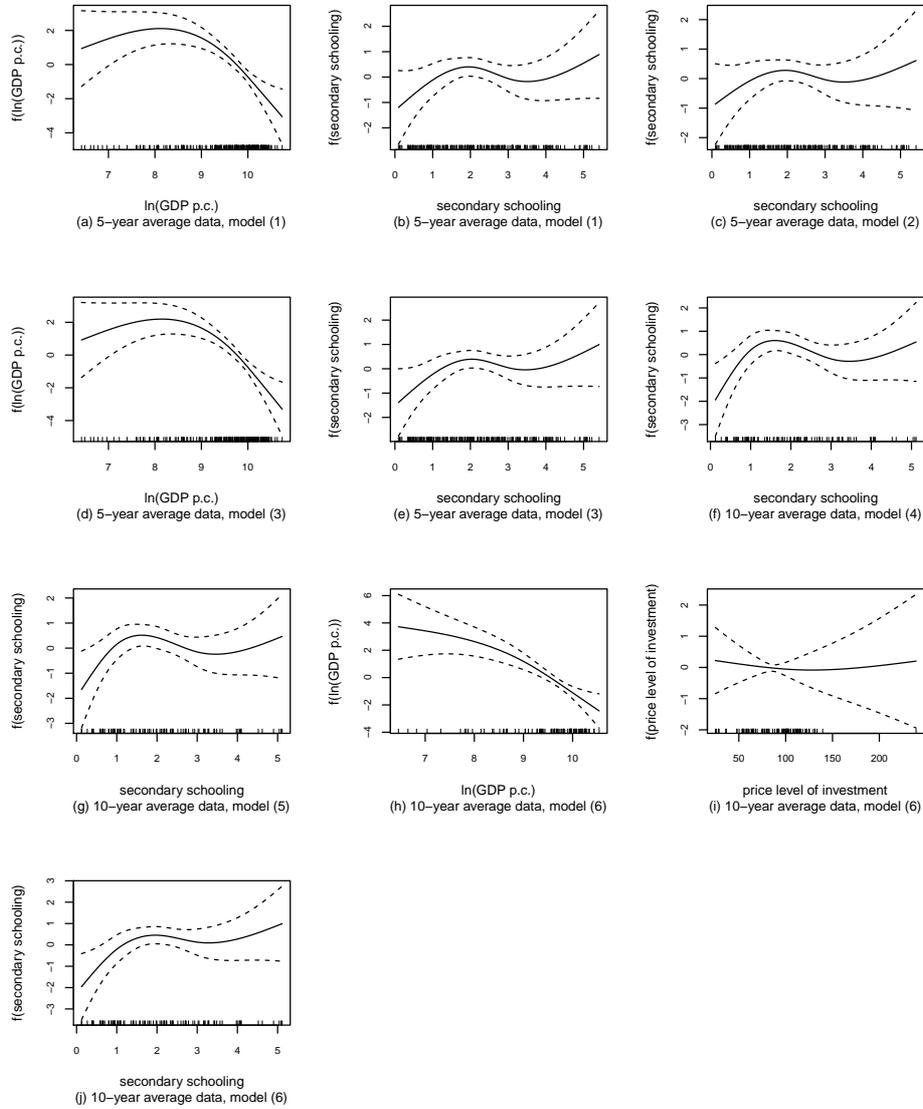


Figure E.10: Visualization of the expanded models' univariate smooths provided in Table 3 (data from the 1950s onward; GDP data from PWT 7.0). Each plot presents the smooth function  $f$  as a solid line. The plots also show the 95% Bayesian credible intervals as dashed lines and the covariate values 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).
- Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., Zucman, G., 2016. The World Wealth and Income Database. Website: <http://www.wid.world>.
- 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/UVP09KSJJ0> (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](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.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).
- 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., 2016. Top-end inequality and growth: Empirical evidence. Tampere Economic Working Papers No. 108 (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.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.