

UNU-WIDER

30 YEARS OF RESEARCH
FOR DEVELOPMENT

WIDER Working Paper 2015/036

Reversal of the Kuznets curve

Study on the inequality–development relation using top income shares data

Elina Tuominen*

March 2015

Abstract This paper uses recently published top 1% income share series in studying the inequality–development association. The top income shares data are of high quality and cover about a century for some countries and thus provide an interesting opportunity to study slow development processes. The empirical inequality–development studies have started to call into question the use of parametric (quadratic) specifications. To address the issue of functional form, this study exploits penalized spline methods. The association between top 1% share and development is found to experience a reversal at later stages of development and, thus, a positive link is observed in many advanced economies. Although this study is not taking a strong stand on causality, additional analysis in this paper advocates that more than sectoral shifts are needed to explain distributional changes.

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

JEL classification: N30, O11, O15

Acknowledgements. This study has received funding from the Finnish Cultural Foundation and the Academy of Finland (project 268863). The author wants to thank Ravi Kanbur, Matti Tuomala and Arto Luoma for discussions. The paper has also benefited from comments made by participants in the FDPE Econometrics Workshop II/2013, FDPE Labour and Public Economics Workshop I/2014, Annual Meeting of the Finnish Economic Association 2014, IIPF Congress 2014, and UNU-WIDER Conference 2014 on ‘Inequality – measurement, trends, impacts, and policies’. Remaining errors are the author’s own.

*University of Tampere, elina.tuominen@uta.fi

This paper has been presented at the UNU-WIDER ‘Inequality—Measurement, Trends, Impacts, and Policies’ conference, held 5–6 September 2014 in Helsinki, Finland.

Copyright © UNU-WIDER 2015

ISSN 1798-7237 ISBN 978-92-9230-921-3 <https://doi.org/10.35188/UNU-WIDER/2015/921-3>

Typescript prepared by the author.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy-making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

UNU-WIDER, Katajanokanlaituri 6 B, 00160 Helsinki, Finland, wider.unu.edu

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the Institute or the United Nations University, nor by the programme/project sponsors, of any of the views expressed.

1. Introduction

In his seminal paper, Kuznets (1955) presented the famous ‘inverted-U hypothesis’ between inequality and economic development. He suggested that the relationship between inequality and economic development resembles an inverse-U curve as the focus of the economy shifts from agriculture to other sectors.¹ Many theoretical papers have studied the Kuznets curve and found support for it (e.g., Robinson, 1976; Galor & Tsiddon, 1996; Aghion & Bolton, 1997; Dahan & Tsiddon, 1998).

Empirical studies have presented mixed evidence on the shape of the inequality–development association, and the debate has focused on whether the results support the inverse-U hypothesis or not. In empirical applications, the chosen functional form of the GDP per capita variable plays an important role. For example, a cross-sectional study by Ahluwalia (1976) supports the inverted-U link but Anand and Kanbur (1993) challenge the results with respect to chosen functional forms and data quality. Also Deininger and Squire (1998) and Barro (2000) find differing results in their panel-data studies. Huang (2004), Lin et al. (2006) and Huang and Lin (2007) address the problem of functional form using non- or semiparametric methods and find evidence for the Kuznets hypothesis. However, the latter three studies rely on cross-sectional data, and it is possible that this type of data cannot capture the complexity of the process. Panel studies have become more common with the improvement in the data-sets.

Frazer (2006) relies on nonparametric methods in his study that spans approximately 50 years, and he uses the World Income Inequality Database (WIID). In his pooled models he finds a nonlinear Gini–development association that is more complex than a second degree polynomial. Specifically,

¹In his numerical illustrations, Kuznets (1955) assumed that incomes per capita are lower in the rural than in the urban sector, and that inequality within the rural sector is lower than or equal to inequality within the urban sector. He showed that if the ratio of the sectoral per capita incomes and the two intrasectoral distributions are held constant, the mere population shift from the agricultural sector to other (modern) sectors can affect the overall distribution: inequality first increases, and then declines. In addition—what is quite rarely mentioned—, Kuznets discussed various other conflicting forces behind the evolution of income distribution. For example, he noted that the cumulative effect of concentrated savings at the top of the distribution induce inequality in the distribution before taxes and transfers, and he discussed equalizing forces such as political pressure for redistribution.

he finds that the curve may be flat before it experiences a negative slope. His illustrations also show that there may be a positive association at the highest levels of development but this reversal is not statistically significant. Moreover, Zhou and Li (2011) conduct a nonparametric investigation on the inequality–development association using unbalanced panel data and they find an inverse-U relation between Gini coefficient and development, but only after a certain development level is reached. Also Desbordes and Verardi (2012) use semiparametric methods with Gini data and provide empirical evidence for the latter stages of the Kuznets relation. Desbordes and Verardi also conclude that mis-specified functional forms can lead to opposite results related to the inequality–development association. Thus, recent studies suggest that using flexible methods in the inequality–development studies is well-founded. The current study utilizes penalized regression splines.

The inequality data-set by Deininger and Squire (1996) was constructed because the reliability of the previously used data had been called into question. Since then, the Deininger–Squire data (or its subsets) have been widely used although also these data have been criticized. For example, Atkinson and Brandolini (2001) highlight that consistent inequality series have not been available. To bring new insights, this study exploits new inequality data on top 1% income shares (Alvaredo et al., 2013b) to study the inequality–development association. The top income share series are unprecedentedly long and cover almost the whole twentieth century for some countries. During this period some countries have faced not only urbanization but also more advanced stages of development. The focus of the paper is mainly in (currently) ‘advanced’ countries but also some (currently) ‘less-advanced’ countries are included. The data are of high quality compared to many other inequality data. Moreover, top income shares can be considered useful as a general measure of inequality, especially when other alternatives for long series are not available (see, e.g., Leigh, 2007; Roine & Waldenström, 2015).

The inverse-U association has been challenged but, e.g., Kanbur (2011) reminds that Kuznets (1955) discussed various forces beside the famous sectoral shift. For example, Atkinson (1995, pp. 25–26) suspects that Kuznets would not have been surprised if the inverse-U shape no longer holds. The shift from agricultural sector to urban sector has already taken place in the ‘advanced’ countries, and various inequality indices have shown an upward trend in many countries during the last 20–30 years. However, to address the issue of sectoral shifts, this study takes a broader interpretation of Kuznets’ ideas and investigates empirically also the possibility of a new structural shift

since the 1980s. For example, List and Gallet (1999) discuss the possibility of a shift from manufacturing toward services in the ‘advanced’ economies, and the current study extends this discussion.

This paper finds that the inequality–development association is U-shaped during the later phases of the development process when inequality is measured by the top 1% income share. In an additional investigation covering the years 1980–2009, the discovered positive slope (at high levels of development) is robust to including controls for urbanization and service sector. The results also show that the relationship between urbanization and top-end inequality can resemble an inverse U, and that employment in services is positively related to top 1% income shares. The current study refrains from taking a strong stand on causality. However, the results rather advocate that –in addition to sectoral changes– there are also other factors behind the evolution of top income shares. Moreover, there are clear similarities in the overall shape of the inequality–development relationship when one compares the results of this paper to those in Frazer (2006) even though the studies use different distributional measures.

The paper is organized as follows: Section 2 introduces the data used in the empirical analysis, and Section 3 describes the estimation method. Section 4 provides empirical results including sensitivity analysis. Finally, Section 5 concludes.

2. Data

2.1. Top 1% income share data

Many of the available Gini series have suffered from comparability problems both in time and between countries, and the series have not covered long time intervals. Using tax and population statistics, it is possible to compose long and fairly consistent series on top income shares. Kuznets (1953) was the first to use this kind of data to produce top income share estimates, and Piketty (2001, 2003) generalized Kuznets’ approach. Following Piketty, top income share series have been constructed by different researchers.² Accord-

²Different countries’ series have been constructed using similar methods. For more information about the methodology see, e.g., Atkinson (2007). Especially Piketty and Saez (2006), Leigh (2007), and Roine and Waldenström (2015) discuss the advantages and limitations of the top income share series. Furthermore, Atkinson et al. (2011) provide a thorough overview of the top-income literature.

ing to Leigh (2007), the evolution of top income shares is similar to that of various other inequality indices over time. Also Roine and Waldenström (2015) conclude that top income shares are useful as a general measure of inequality over time.

Top income data can be easily accessed using the *World Top Incomes Database* by Alvaredo et al. (2013b).³ The top 1% income shares of 26 countries from 1900 to 2010 are exploited but the longitudinal data are not balanced (note that this is pre-tax income). The data include, e.g., the English-speaking countries, Continental and Southern European countries, Nordic countries, and some ‘less-advanced’ countries⁴. The top 1% income share (*top1*) series are presented graphically in Appendix A.

During the first half of the twentieth century, top incomes consisted mainly of capital income. In most countries, capital incomes fell dramatically during wartime and the Great Depression. Also the distribution of earned incomes became more equal in many countries after the wars. One suggested explanation for the extended fall in top income shares is progressive taxation. In the English-speaking countries, the evolution of top 1% income shares resembles U over the twentieth century because there has been a significant increase since the 1980s. In contrast, top 1% shares in the Continental Europe and Japan have remained fairly stable during the past three decades. In the USA and other English-speaking countries, the growth in top income shares has been explained by growth in top wages after the 1970s. As top wages have increased, top executives and capital owners cohabit the top of the income distribution. In contrast, in Finland and Sweden capital incomes continue being important. One suggested explanation for the recent rise in top incomes is the decrease in the highest tax rates.⁵ Furthermore, it has been suggested that skill-biased technological change is not sufficient in explaining the surge in top earnings. Literature on ‘college premium’ does not describe well why the top 1% share has increased in comparison to the top 10% share, since basically the whole group of top 10% are college-educated.

³The first book on these series, edited by Atkinson and Piketty (2007), contrasted the evidence from the Continental Europe and the English-speaking countries. The second volume, also edited by Atkinson and Piketty, was published in 2010. The database builds on these series and the project is still on-going.

⁴Here: Argentina, China, Colombia, India, Indonesia, Mauritius, and South Africa.

⁵Roine et al. (2009) provide empirical evidence on the negative association between tax progressivity and top income shares.

Alternative explanations, e.g., theories on executive remuneration in a hierarchical structure and superstar theory, have been suggested. (Piketty & Saez, 2003, 2006; Atkinson et al., 2011; Alvaredo et al., 2013a)

On the basis of these findings an inverse-U-shaped *top1*–development association is not expected. Atkinson et al. (2011) and Roine and Waldenström (2015) discuss the problems of fitting top income shares into the Kuznets (1955) approach where the inverse-U relation is described by the shift from traditional to modern sector. Roine and Waldenström suggest that other factors pointed out by Kuznets deserve more attention: especially taxation, and the cumulative effect of concentrated savings at the top. But if one is willing to accept top 1% income shares as a more general inequality measure it is of interest to study empirically the idea of structural shifts after the 1970s. We have observed the expansion of the information technology and financial sector (or more generally put: the expansion of the service sector). For this reason, this study provides an additional analysis on the question of sectoral changes.

2.2. *Development and economic sectors*

Level of economic development is measured in a traditional way using GDP per capita. The GDP data are available annually until 2010 in the Maddison Project update (Bolt & van Zanden, 2013). Data from 1900 are used whenever available, and the series are plotted in Appendix B.

In an additional analysis, the baseline results are checked by controlling for two sectors, namely urban and service sectors. Urbanization data describe the *population residing in urban areas (%)* (United Nations, 2012). These data are available every 5 years. The service sector is measured with *employment in service sector (% of total employment)* (World Bank, 2014a), and these data are available from 1980 onwards. The sectoral variable series are plotted in Appendix C. Although the time span becomes considerably shorter with these two controls, studying the two sectors can be seen as an extension to previous studies. For example, Frazer (2006) reports controlling for urbanization but he does not provide detailed results on the inequality–urbanization relationship. Whereas Desbordes and Verardi (2012) do not include sectoral variables in their empirical models.⁶

⁶Kanbur and Zhuang (2013) is a recent example of focusing on the inequality–urbanization relationship in four Asian countries in the spirit of Kuznets (1955).

3. Estimation method

Additive models provide a flexible framework to investigate the association between inequality and development.⁷ The approach presented in Wood (2006) is followed. The basic idea is that the model's predictor is a sum of linear and smooth functions of covariates:

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

where $Y_i \sim$ normal distribution.

In the above presentation Y_i is the response variable, \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. Firstly, one needs to represent the smooth functions f_{\bullet} in some way. One way to represent these smooths are cubic regression splines, which is the approach taken here. 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 are *knots* of the spline, and these locations must be chosen. The spline can be represented in terms of its values at the knots.⁸ Secondly, one needs to choose the amount of smoothness that functions f_{\bullet} have. One does not want to overfit, and thus the departure from smoothness is penalized. The appropriate degree of smoothness for the f_{\bullet} can be estimated from the data. Various selection criteria are available, e.g., (generalized) cross-validation or maximum likelihood.

Illustration

Consider a model containing only one smooth function of one covariate: $y_i = f(x_i) + \epsilon_i$, where ϵ_i are i.i.d. $N(0, \sigma^2)$ random variables. To estimate

⁷Additive models are a special case of generalized additive models (GAMs). GAMs were introduced by Hastie and Tibshirani (1986, 1990). They present a GAM as a generalized linear model with a linear predictor that involves a sum of smooth functions of covariates. Some of the covariates can enter in linear form. Note here the analogy to 'generalized linear models' and 'linear models'. This paper restricts to a special case: using an identity link and assuming normality in errors, which leads to additive models.

⁸Usually, there are two extra conditions specifying that the second derivative of the curve should be zero at the two end knots.

function f here, f can be 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 to be estimated is $f(x) = \sum_{j=1}^k \beta_j b_j(x)$, where β_j refers to coefficients that are estimated, and b_j to known basis functions. Using a chosen basis for f means 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 $\int f''(x)^2 dx$, where $f''(x) = \sum_{j=1}^k \beta_j b_j''(x) = \boldsymbol{\beta}^T \mathbf{b}''(x)$, and $\mathbf{b}''(x)$ is the vector of second derivatives of the basis functions evaluated at x . The penalty $\int f''(x)^2 dx$ can be expressed as $\boldsymbol{\beta}^T \mathbf{S}\boldsymbol{\beta}$, where \mathbf{S} is the *penalty matrix* that can be expressed in terms of the known basis functions.

The estimation procedure is based on the minimization of the sum of squared residuals subject to a penalizing term, i.e. $\min \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \boldsymbol{\beta}^T \mathbf{S}\boldsymbol{\beta}$, with respect to $\boldsymbol{\beta}$. The problem of estimating the degree of smoothness is a problem of estimating the smoothing parameter λ .⁹ The penalized least squares estimator of $\boldsymbol{\beta}$, given λ , 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 various covariates and smooths. Given a basis, an additive model is simply a linear model with one or more associated penalties.

Practical notes

Usually in estimation, the size of the basis dimension for each smooth is not critical because it only sets an upper limit on the flexibility of a term. Smoothing parameter controls the actual effective degrees of freedom (edf). Effective degrees of freedom are defined as $\text{trace}(\mathbf{A})$ where \mathbf{A} is the influence matrix. It is also possible to divide the effective degrees of freedom into degrees of freedom for each smooth. The effective degrees of freedom can be used to measure the flexibility of a model. For example, a simple linear term would have one degree of freedom, and $\text{edf}=2.3$ can be thought of as a

⁹In 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 un-penalized regression spline estimate.

function that is a bit more complex than a second degree polynomial.

Confidence (credible) intervals for the model terms or parameters can be derived using Bayesian methods. Also approximate p-values for model terms can be calculated. Models can be compared using information criteria, e.g., the Akaike information criterion (AIC). When using 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. In the next section, the effective degrees of freedom of each smooth term and model selection criteria are provided for each model. However, the focus will be on investigating graphical illustrations.¹⁰

4. Estimation results

Below, the baseline models refer to specifications without sectoral variables, and the estimation is implemented with annual data from 1900 to 2010. The results are also checked by studying different subsets of the data and changing the data structure from annual data to 5-year averages. Finally, to discuss issues related to sectoral shifts, urbanization and service sector variables are included in models with 5-year average data covering years from 1980 to 2009.

4.1. Baseline models

The baseline results are for annual data and cover the years 1900–2010 (whenever observations are available). The models are of the form:

$$top1_{it} = \alpha + f(\ln(GDP \text{ per capita})_{it}) + \delta_{decade} + u_i + \epsilon_{it},$$

where α is a constant, f is a smooth function that is described using a penalized cubic regression spline, i refers to country and t refers to year,

¹⁰The results presented in this paper are obtained using the R software package ‘mgcv’ (version 1.7-21), which includes a function ‘gam’. Marginal basis construction ‘cr’ for cubic regression splines is used. The knots are placed evenly through the range of covariate values (default). The maximum likelihood method ‘ML’ is used in the selection of the smoothing parameters. The identifiability constraints (due to e.g., 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/>).

δ_{decade} is a fixed time effect for each decade, u_i is a country effect and $\epsilon_{it} \sim N(0, \sigma^2)$ is the error term. The country effects can be fixed (dummy for each country) or random ($u_i \sim N(0, \sigma_u^2)$). Different strategies in modeling country effects are reported because the literature does not follow a unified approach.

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

	depending variable: $top1_t$ (N=1609)		
	(1)	(2)	(3)
$f(\ln(GDP \text{ per capita})_t)$	[edf=9.19 ^b] ^{***}	[edf=10.44 ^b] ^{***}	[edf=10.43 ^b] ^{***}
country effects	no	fixed	random
adjusted r^2	0.592	0.822	0.822
AIC	7949.7	6642.1	6642.0

Significance levels: (***) < 0.01 , (**) < 0.05 , (*) < 0.10.

These are calculated using the Bayesian estimated covariance matrix of the parameter estimators. The smooth terms' p-values are approximate and based on F statistics.

^aTime effects are simple dummy variables for each decade. The decades are considered as follows: 1900–1909 is the 1900s, 1910–1919 is the 1910s, etc. However, all observations 2000–2010 are considered in the 'last' decade.

^bThe basis dimension of the smooth before imposing identifiability constraints is $k = 15$.

Variables are described in Section 2 and Appendices A–B.

Source: author's calculations.

Results of the pooled model (i.e., without country effects) are provided in column (1) of Table 1. The models (2) and (3) in Table 1 include country effects. Figure 1 illustrates the smooth functions of these models. The fixed-effect (FE) and random-effect (RE) specifications give practically identical fits. In all three specifications, there is a possibility of a flat curve at lower levels of development (approximately $\ln(\text{GDP per capita}) < 8$). And after a certain level of development (approximately $\ln(\text{GDP per capita}) > 8$), all smooths show U shape. However, the downward peak between $9 < \ln(\text{GDP per capita}) < 10$ shifts when country effects are included. All specifications show a positive association at the highest levels of development.

The obtained overall shape of $f(\ln(\text{GDP per capita}))$ resembles the shape that Frazer (2006) shows for the Gini–development relationship (except for the significant positive slope at the highest levels of development). This similarity suggests that the top 1% income shares reflect same characteristics as the traditional Gini coefficients. Even the downward peak close to $\ln(\text{GDP per capita}) \approx 9.5$ in Figure 1(a) seems to be reasonable compared to Frazer's results.

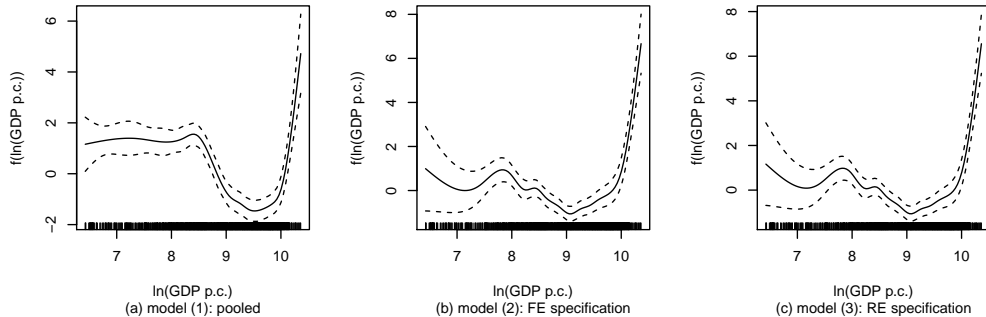


Figure 1: Illustration of the *top1*–development relation (annual data 1900–2010). See Table 1 for details. The plots present only the smooth function $f(\ln(\text{GDP per capita}))$ so the reader should focus on the shape instead of the level. The solid line represents the smooth function f . The plots also show the 95% Bayesian credible intervals (dashed) and the covariate values as a rug plot along the horizontal axis. Source: author’s calculations.

4.2. Sensitivity of the baseline models’ results

Different subsets of the data were studied first, and then the results were checked for the time-period structure. In summary, the main conclusions regarding the ‘medium’ or ‘higher’ levels of development hold.

First, the English-speaking, Nordic, Continental/Southern European, and ‘less-advanced’ countries were studied separately.¹¹ More detailed information on the models is reported in Appendix D (Table D.4). The illustrations of the smooths are provided in Figure 2 and discussed next. Results at lower levels of development (approximately $\ln(\text{GDP per capita}) < 8.5$) are not uniform. But there seems to be a pattern that holds as countries reach a higher level of economic development: there is a negative relationship between *top1* and the level of development when $8.5 < \ln(\text{GDP per capita}) < 9.5$ (approximately). This is highlighted in the group-wise plots. Only the Continental and Southern European countries’ plot (c) shows a downward peak close to $\ln(\text{GDP per capita}) \approx 9$. The English-speaking and the Nordic countries show a ‘turning point’ closer to $\ln(\text{GDP per capita}) \approx 9.5$ (plots (a) and (b), respectively). Plot (d) shows that it remains to be seen if this kind of a re-

¹¹Singapore and Japan do not fit into these categories and were, thus, not included in these group-wise investigations.

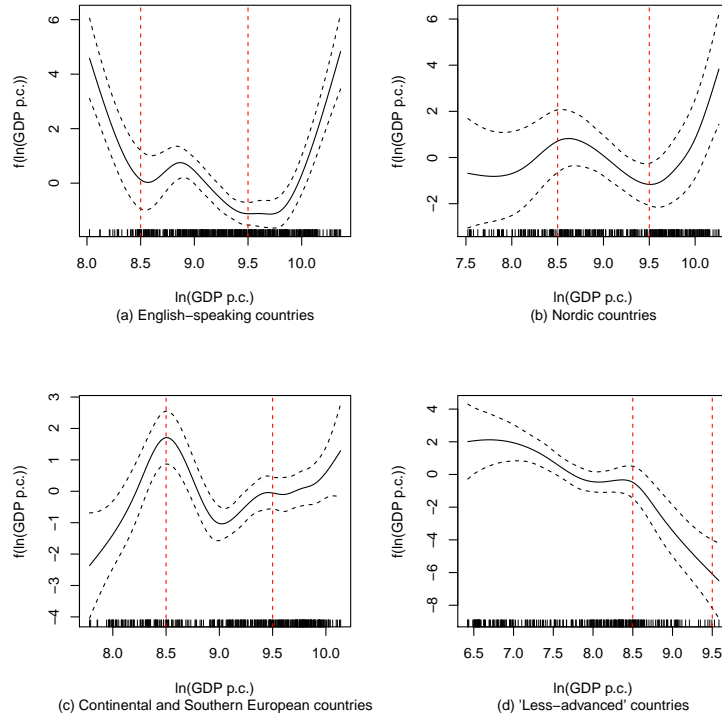


Figure 2: Illustration of the *top1*–development relation with four different subsets of the data (annual data 1900–2010). The models include decade dummies and random country effects (Table D.4 in Appendix D provides details). The solid line represents the smooth function $f(\ln(\text{GDP per capita}))$ so the reader should focus on the shape instead of the level. The plots also show the 95% Bayesian credible intervals (dashed), and the covariate values as a rug plot along the horizontal axis. Vertical, dashed lines have been added to highlight the idea of a negative slope between $8.5 < \ln(\text{GDP per capita}) < 9.5$ (approximately). Source: author’s calculations.

versal arises also for the ‘less-advanced’ economies. The smooths in Figure 2 can be compared to Figure 1(c) which illustrates the corresponding random-effect specification with the whole sample. In summary, the overall shape of the association between top-end inequality and development is fairly unified when $\ln(\text{GDP per capita}) > 8.5$.

The second check concerned the sensitivity to dropping groups of countries from the whole sample. The previously discovered U shape emerges again at development levels $\ln(\text{GDP per capita}) > 8$ (approximately). The downward peak of the U is again located between $9 < \ln(\text{GDP per capita}) < 10$. More detailed information is reported in Appendix D (Table D.4 and Figure D.8).¹²

Finally, the annual data results were checked against the corresponding results with the 5-year average data, where the consecutive periods are 1900–1904, 1905–1909, ..., 2000–2004, and 2005–2009. The results with the 5-year data do not differ from the results in Subsection 4.1. To avoid repetition, these results are not discussed here in detail. Appendix E provides graphical illustrations. Thus, the overall results do not seem to depend on the choice between annual or 5-year average data. This sensitivity check relates to the following investigation, where the analysis is conducted with the 5-year averages, but the models are augmented with sectoral variables.

4.3. *Controlling for two sectors*

This section extends the analysis by controlling for both urban and service sectors (as described in Subsection 2.2 and Appendix C). The service sector variable is available from 1980. Because the urbanization variable is available every 5 years, the analysis is implemented using 5-year average data. The averaged data are constructed using consecutive periods (1980–1984, 1985–1989, ..., 2005–2009). As the results below will show, limiting the investigated time period to mere 30 years does not alter the main findings.

The studied specifications are as follows:

¹²In addition, the effect of dropping Japan or Singapore from the sample was tested because these two countries do not fit into the discussed categorization. Dropping Singapore from the sample does not have an effect on the shape of $f(\ln(\text{GDP per capita}))$. Whereas leaving Japan out has some effect at the lowest levels of development, but the 95% credible interval is still quite wide at these levels of development ($\ln(\text{GDP per capita}) < 7.5$). Thus, the main results that relate to ‘medium’ or ‘high’ levels of development are not sensitive to these countries.

$$top1_{it} = \alpha + f_1(\ln(GDP \text{ per capita})_{it}) + f_2(urbanization_{it}) + f_3(employment \text{ in services}_{it}) + \delta_{decade} + u_i + \epsilon_{it},$$

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

Table 2: Results on models with two sectors, using 5-year average data (years 1980–2009): effective degrees of freedom for each smooth f_\bullet and coefficients for linear terms. Intercepts, country effects and time effects^a are not reported. The smooths with $edf > 1$ are illustrated in Figure 3.

	depending variable: $top1_t$ (N=129)		
	(1)	(2)	(3)
$f_1(\ln(GDP \text{ per capita})_t)$	[edf=4.67 ^b] ^{***}	[edf=5.29 ^b] ^{***}	[edf=5.38 ^b] ^{***}
$f_2(urbanization_t)$	[edf=5.81 ^b] ^{***}	[edf=3.59 ^b] [*]	[edf=4.11 ^b] ^{**}
$f_3(empl. \text{ in services}_t)$	[edf=2.90 ^b] ^{***}	[linear] 0.096 ^{**}	[linear] 0.120 ^{***}
country effects	no	fixed	random
adjusted r^2	0.724	0.940	0.935
AIC	542.3	361.5	370.7

Significance levels: (***) < 0.01 , (**) < 0.05 , (*) < 0.10 .

These are calculated using the Bayesian estimated covariance matrix of the parameter estimators. The smooth terms' p-values are approximate and based on F statistics.

^aTime effects are simple dummy variables for each decade. The decades are considered as follows: 1980–1989 is the 1980s, 1990–1999 is the 1990s, and 2000–2009 is the 2000s.

^bThe basis dimension of the smooth before imposing identifiability constraints is $k = 10$.

Variables are described in Section 2 and Appendices A–C.

Source: author's calculations.

Table 2 provides details on models with two sectors. There is a linear term for the service sector in models (2) and (3), and the coefficients are provided in the table. Figure 3 provides plots of the other smooth functions. In plots (a) and (c), the pooled and random-effect specifications give very similar shapes for the smooth $f(\ln(\text{GDP per capita}))$, and the overall shape does not contradict previously reported results. In the fixed-effect specification (plot (b)), the 95% credible interval is very wide at lower levels of development.

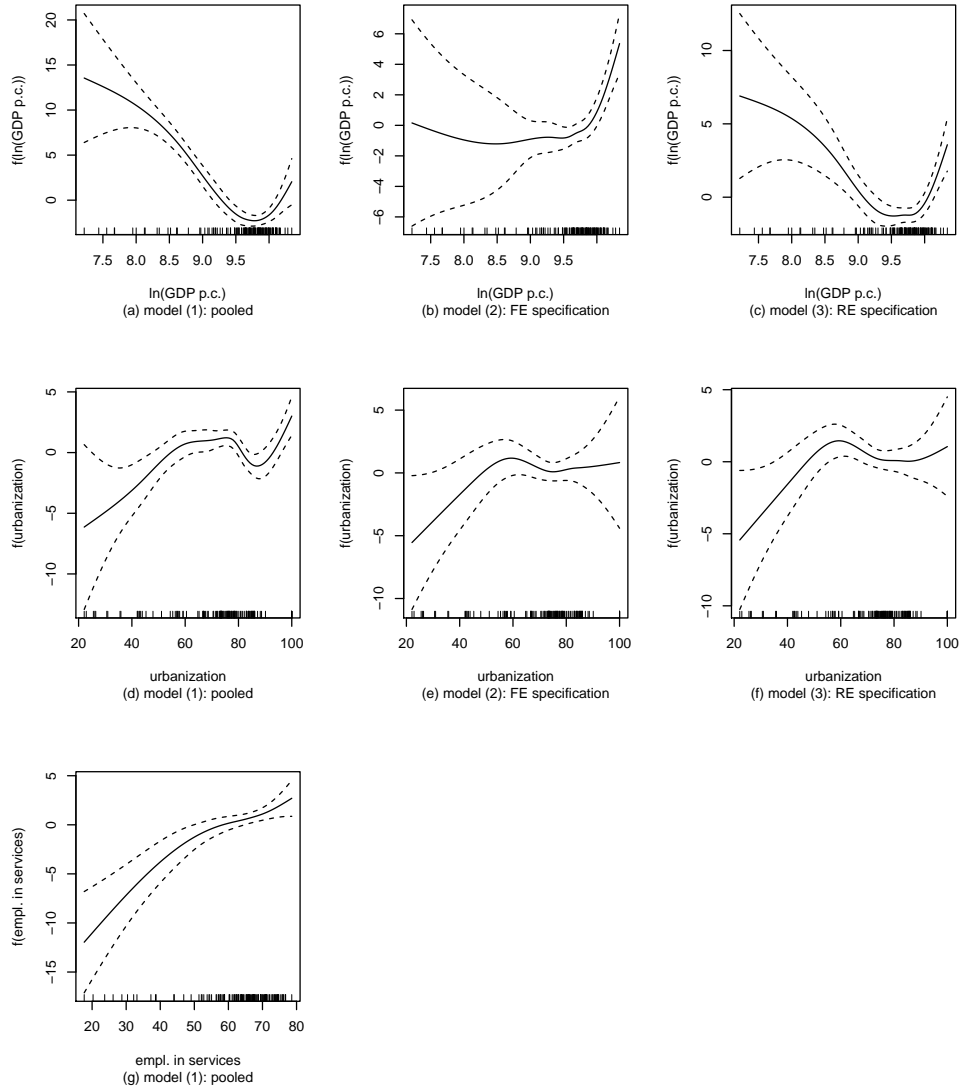


Figure 3: Illustrations of the smooths, using 5-year average data (years 1980–2009). See Table 2 for model details. The plots present only the smooth functions f_{\bullet} , so the reader should focus on the shapes of the smooths instead of the levels. The plots also show the 95% Bayesian credible intervals (dashed), and covariate values as a rug plot along the horizontal axis.

Source: author's calculations.

But the positive slope at the highest levels of GDP shows in all three specifications, and the ‘turning point’ is located close to $\ln(\text{GDP per capita}) \approx 9.5$. As a comparison, also Frazer (2006) controls for urbanization in the sensitivity checks of his pooled model and finds that the overall shape of the Gini–development relationship holds. Moreover, Table 2 and Figure 3 show that the results on sectoral variables are fairly uniform irrespective of the country-effect specification. Smooth of urbanization resembles an inverted-U curve (especially in plots (e) and (f) of Figure 3). The association between top 1% share and employment in services is positive which allows speculation whether this illustrates a new structural shift.

These results were also checked against leaving country groups out of the sample (one group at a time).¹³ Leaving the English-speaking countries out of the sample weakened the positive slope in $f(\ln(\text{GDP per capita}))$ at the highest levels of development. However, dropping the Nordic or the Continental/Southern European countries did not change the overall shape of $f(\ln(\text{GDP per capita}))$. Dropping the group of the ‘less-advanced’ countries had a more evident effect on $f(\ln(\text{GDP per capita}))$ but this was anticipated: with 30 years of recent data, dropping these countries means that the remaining $\ln(\text{GDP per capita})$ values were larger than 9. Thus, in this case the association between $\ln(\text{GDP per capita})$ and *top1* was mainly positive (slightly J-shaped curve). Moreover, the results on urbanization and service sector did not contradict the previous findings qualitatively. Thus, surprising changes did not occur even when considerably large proportions of observations were dropped from the sample, and details of this check are not reported.

Finally, an alternative measure for the service sector was tested. Data on *services, etc., value added (% of GDP)* (World Bank, 2014b) start already from the 1960s for some countries, but Swiss data are not available. Results related to $\ln(\text{GDP per capita})$ and urbanization did not change. The alternative service sector measure correlated positively with *top1* but it was not statistically significant (at 10% level) in specifications with country effects. However, these results were not in conflict with the models reported in Table 2. Thus, details are not reported.

The overall shape of the smooth $f(\ln(\text{GDP per capita}))$ in plots (a) and

¹³Categorization was the same as in the previous subsection and Table D.4. The checks were performed using specifications with decade dummies and random country effects.

(c) of Figure 3 resembles the shape shown in Figure 1 (and also in Figure E.9 in Appendix E). Especially, the result of a positive *top1*–development relation at the highest levels of development holds when two sectoral measures are included. Although this paper is not taking a strong stand in causality, the results advocate that sectoral shifts are only one side of the story in distributional changes.

5. Discussion

Kuznets (1955) suggested that inequality first increases during modernization, but later in the development process it starts to decrease. In actual data, he observed a plateau following a decline in inequality in some countries during the first half of the twentieth century. The results of the current study are based on an unbalanced longitudinal data from 26 countries covering 1900–2010. Various specifications in this paper suggest a negative association between top 1% income share and $\ln(\text{GDP per capita})$ after a certain point in the development process. In addition, this study finds that this relationship turns positive at even higher levels of economic development. Thus, the data suggest reversal of the famous Kuznets curve in the ‘advanced’ economies. The current study also shows that more research on ‘less-advanced’ economies is needed when new data become available.

In an additional analysis covering 1980–2009, this paper takes a broad interpretation of the Kuznets process. The discovered positive association between top 1% share and development is robust to inclusion of two sectoral measures. The empirical results advocate that sectoral shifts are not to be taken as the sole explanation behind the evolution of the top 1% income shares. Thus, the findings are in line with the ideas by Atkinson et al. (2011) and Roine and Waldenström (2015) who discuss the problems in fitting changes in top income shares into the story of structural shifts in the economy. It seems that various competing forces deserve more attention in the future studies.

Appendix A. Top 1% income share series

Table A.3: Top 1% income share series (years 1900–2010). For better comparability, series excluding capital gains have been selected whenever possible. The series are plotted in Figure A.4 below.

Country (abbreviation)	N	Source
Argentina (ARG)	39	Alvaredo et al. (2013b)
Australia (AUS)	90	Alvaredo et al. (2013b)
Canada (CAN)	91	Alvaredo et al. (2013b) ^b
China (CHN)	18	Alvaredo et al. (2013b)
Colombia (COL)	18	Alvaredo et al. (2013b)
Denmark (DNK)	95	Alvaredo et al. (2013b)
Finland (FIN)	90	Alvaredo et al. (2013b) ^c
France (FRA)	96	Alvaredo et al. (2013b) ^d
Germany (DEU)	47	Alvaredo et al. (2013b)
India (IND)	71	Alvaredo et al. (2013b)
Indonesia (IDN)	28	Alvaredo et al. (2013b)
Ireland (IRL)	37	Alvaredo et al. (2013b)
Italy (ITA)	34	Alvaredo et al. (2013b)
Japan (JPN)	110	Alvaredo et al. (2013b)
Mauritius (MUS)	56 ^a	Alvaredo et al. (2013b)
Netherlands (NLD)	55	Alvaredo et al. (2013b)
New Zealand (NZL)	83	Alvaredo et al. (2013b)
Norway (NOR)	69	Alvaredo et al. (2013b)
Portugal (PRT)	24	Alvaredo et al. (2013b)
Singapore (SGP)	62 ^a	Alvaredo et al. (2013b)
South Africa (ZAF)	71 ^a	Alvaredo et al. (2013b)
Spain (ESP)	30	Alvaredo et al. (2013b)
Sweden (SWE)	79	Alvaredo et al. (2013b)
Switzerland (CHE)	74	Alvaredo et al. (2013b) ^e
United Kingdom (UK)	60	Alvaredo et al. (2013b)
United States (USA)	98	Alvaredo et al. (2013b)
total: 1625 ^a		

^aAvailability of GDP data limits the number of observations in the estimated models to N=1609 with Mauritius N=52, Singapore N=59, and South Africa N=62.

^bTwo overlapping series available. Here: series up to 1981 is based on tax data, and series from 1982 is based on Longitudinal Administrative Database.

^cTwo overlapping series available. Here: series up to 1989 is based on tax data, and the series from 1990 is based on the Income Distribution Survey.

^dThe figure for 1905 is for 1900–1910 averaged.

^eFor all years except 1933, the estimates relate to income averaged over the year shown and the following year. Thus, repeated value for two consecutive years is used in this study.

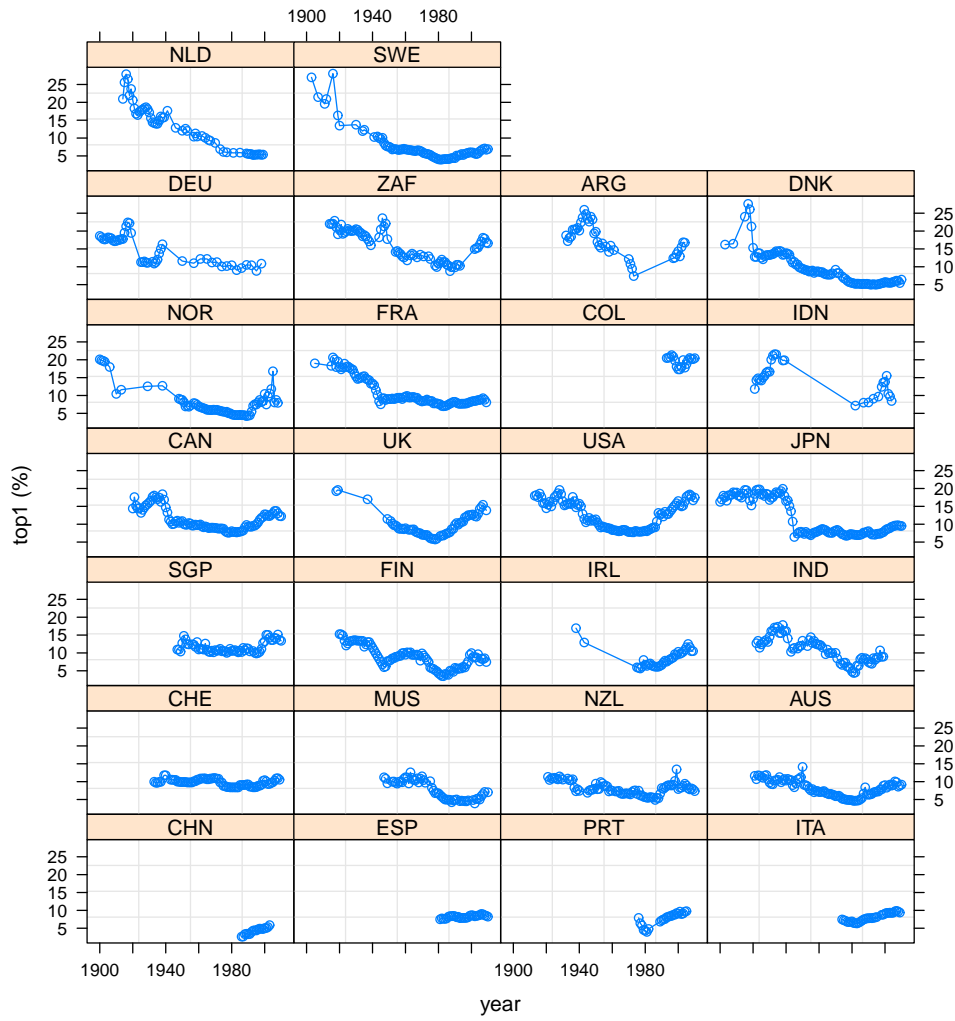


Figure A.4: Top 1% income share series for each country (years 1900–2010). See Table A.3 for details and country abbreviations. Data source: Alvaredo et al. (2013b).

Appendix B. GDP per capita series

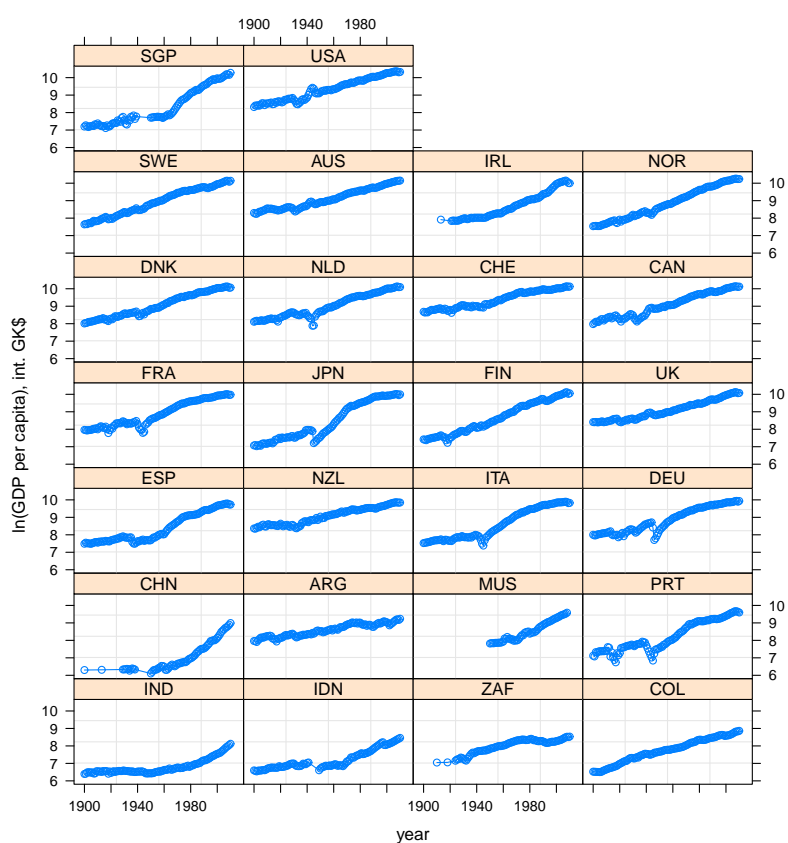


Figure B.5: Natural logarithm of GDP per capita in international 1990 Geary-Khamis dollars for each country (years 1900–2010). See Table A.3 for country abbreviations. Data source: update of Maddison Project (Bolt & van Zanden, 2013).

Appendix C. Sectoral variable series

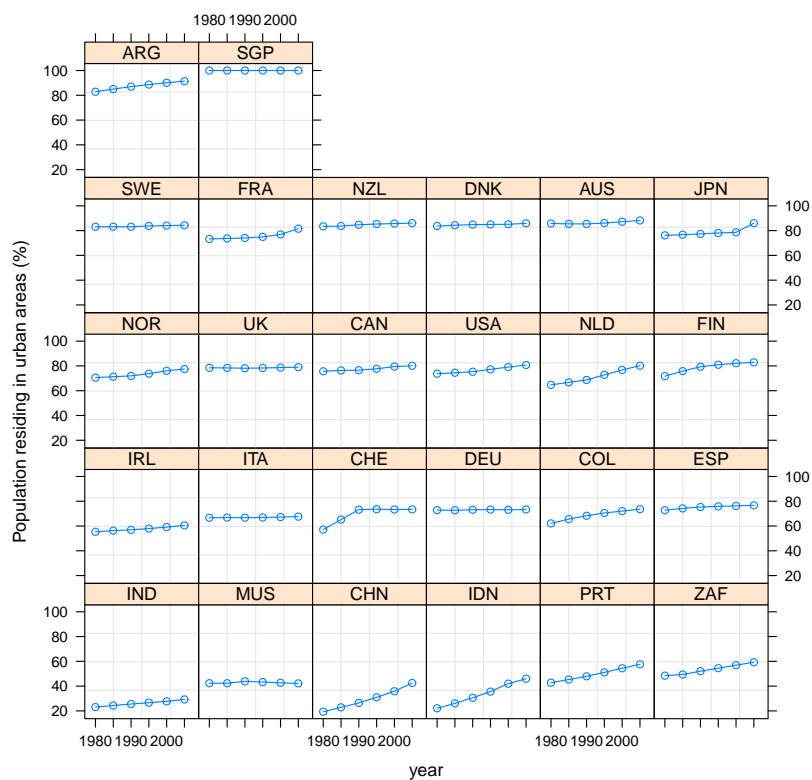


Figure C.6: Urbanization, years 1980–2005. See Table A.3 for country abbreviations. Data source: United Nations (2012).

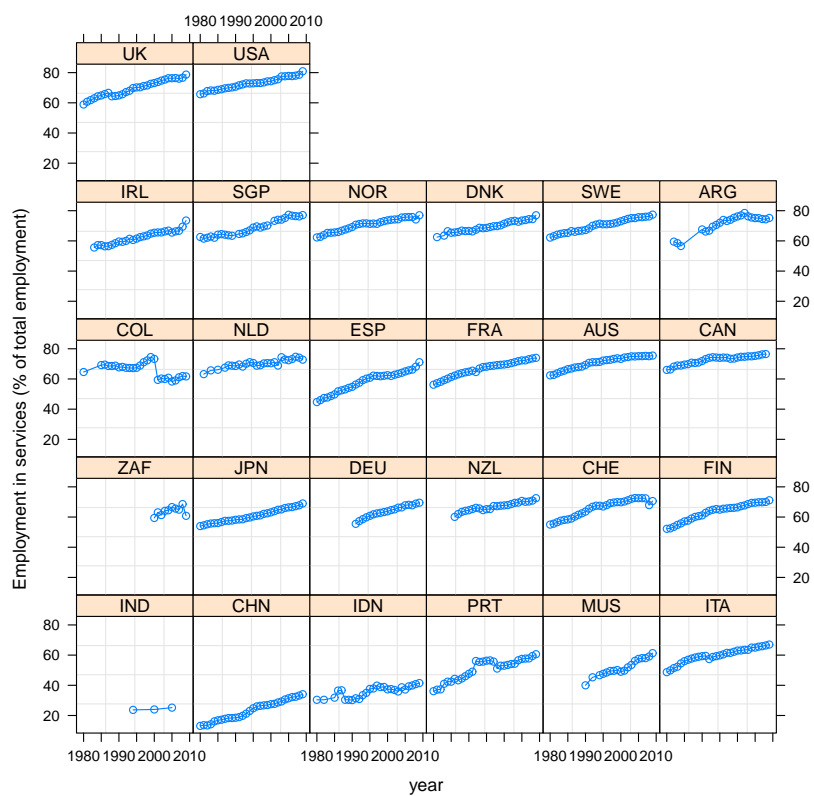


Figure C.7: Employment in services, years 1980–2009. See Table A.3 for country abbreviations.

Data source: World Bank (2014a).

Appendix D. Model details: subsets of data

Table D.4: Subsets of data. Results on models with random country effects, using annual data (years 1900–2010): effective degrees of freedom for each smooth. Intercepts, country effects and time effects^a are not reported.

$top1_{it} = \alpha + f(\ln(GDP \text{ per capita})_{it}) + \delta_{decade} + u_i + \epsilon_{it}$	N	smooth f
In Figure 2:		
English-speaking ^b	459	[edf=7.41 ^f]***
Nordic ^c	333	[edf=5.85 ^f]***
Continental and Southern Europe ^d	360	[edf=6.94 ^f]***
‘Less-advanced’ ^e	288	[edf=5.37 ^f]***
In Figure D.8:		
Without English-speaking ^b	1150	[edf=10.02 ^g]***
Without Nordic ^c	1276	[edf=10.03 ^g]***
Without Continental/Southern Europe ^d	1249	[edf=9.78 ^g]***
Without ‘less-advanced’ ^e	1321	[edf=9.47 ^g]***

Significance levels: (***) < 0.01, (**) < 0.05, (*) < 0.10.

These are calculated using the Bayesian estimated covariance matrix of the parameter estimators. The smooth terms’ p-values are approximate and based on F statistics.

^aTime effects are simple dummy variables for each decade. The decades are considered as follows: 1900–1909 is the 1900s, 1910–1919 is the 1910s, etc. However, all observations 2000–2010 are considered in the ‘last’ decade.

^bAustralia, Canada, Ireland, New Zealand, UK, and USA.

^cDenmark, Finland, Norway, and Sweden.

^dFrance, Germany, Italy, Netherlands, Portugal, Spain, and Switzerland.

^eArgentina, China, Colombia, India, Indonesia, Mauritius, and South Africa.

^fThe basis dimension of the smooth before imposing identifiability constraints is $k = 10$.

^gThe basis dimension of the smooth before imposing identifiability constraints is $k = 15$.

Variables are described in Section 2 and Appendices A–B.

Source: author’s calculations.

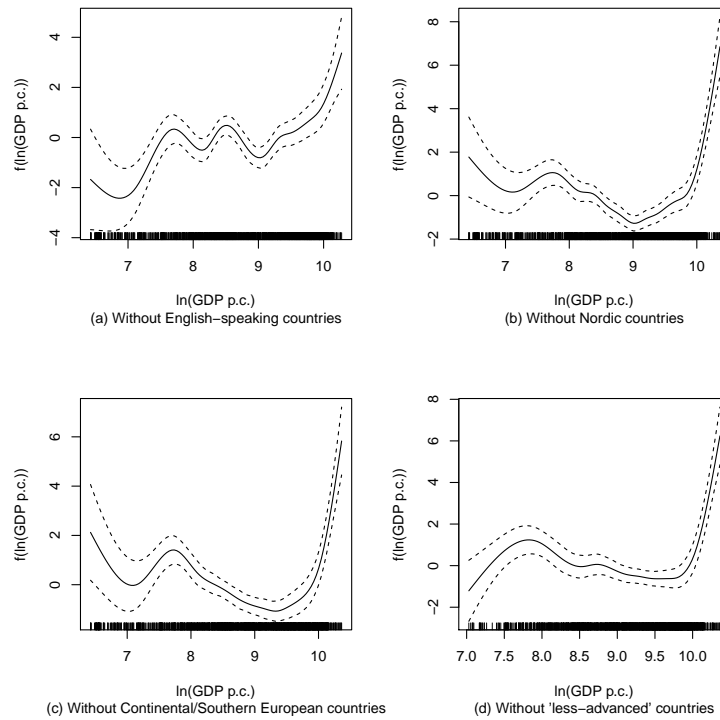


Figure D.8: The effect of leaving countries out of sample (annual data 1900–2010). The models include decade dummies and random country effects. See Table D.4 for model details. The solid line represents the smooth function $f(\ln(\text{GDP per capita}))$. Note that the shape of f is of interest, not the level. The plots also show the 95% Bayesian credible intervals (dashed), and covariate values as a rug plot along the horizontal axis. The shapes of these smooths can be compared to Figure 1(c) which illustrates the corresponding random-effect specification with the whole sample. Source: author's calculations.

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

The baseline models with the 5-year average data (discussed at the end of the Sub-section 4.2) are of the form:

$$top1_{it} = \alpha + f(\ln(\text{GDP per capita})_{it}) + \delta_{decade} + u_i + \epsilon_{it},$$

where α is a constant, f is a smooth function that is described using a penalized cubic regression spline, δ_{decade} is a fixed time effect, u_i is a country effect (fixed or random) and ϵ_{it} is the conventional error term. Here i refers to country and t refers to each 5-year period (1900–04, 1905–09, ..., 2005–09). For each period t , the values for $top1$ and $\ln(\text{GDP per capita})$ refer to period averages.

Figure E.9 below describes the smooths. The obtained shapes of $f(\ln(\text{GDP per capita}))$ are close to the corresponding ones in Figure 1. Thus, changing the modeling strategy from annual to 5-year average data does not influence the overall shapes of the corresponding smooths.

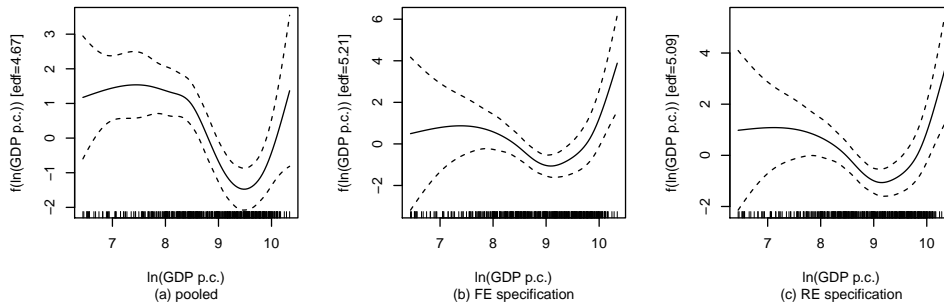


Figure E.9: Illustration of the $top1$ –development relation, using 5-year average data (years 1900–2009, here $N=376$). The solid line represents the smooth function $f(\ln(\text{GDP per capita}))$. In estimation, the basis dimension of the smooth f before imposing identifiability constraints is $k = 10$. The figure also shows the 95% Bayesian credible intervals (dashed), and covariate values as a rug plot along the horizontal axis. Plot (a) represents a model without country effects, plot (b) illustrates a model with country-specific fixed effects, and plot (c) represents a model with country-specific random effects. All models include decade dummies.

Variables are described in Section 2 and Appendices A–B.

Source: author’s calculations.

References

- Aghion, P., Bolton, P., 1997. A Theory of Trickle-Down Growth and Development. *Review of Economic Studies* 64(2), 151–172.
- Ahluwalia, M.S., 1976. Inequality, Poverty and Development. *Journal of Development Economics* 3(4), 307–342.
- Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2013a. The Top 1 Percent in International and Historical Perspective. *Journal of Economic Perspectives* 27(3), 3–20.
- Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2013b. The World Top Incomes Database. <http://g-mond.parisschoolofeconomics.eu/topincomes> (Sep 12, 2013).
- Anand, S., Kanbur, S.M.R., 1993. Inequality and development: A critique. *Journal of Development Economics* 41(1), 19–43.
- Atkinson, A.B., 1995. *Incomes and the Welfare State: Essays on Britain and Europe*. Cambridge University Press, Cambridge.
- 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 and New York, 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 and New York.
- Atkinson, A.B., Piketty, T. (Eds.), 2010. *Top Incomes - A Global Perspective*. Oxford University Press, Oxford and New York.
- Atkinson, A.B., Piketty, T., Saez, E., 2011. Top incomes in the Long Run of History. *Journal of Economic Literature* 49(1), 3–71.
- Barro, R.J., 2000. Inequality and Growth in a Panel of Countries. *Journal of Economic Growth* 5(1), 5–32.
- Bolt, J., van Zanden, J.L., 2013. The First Update of the Maddison Project; Re-Estimating Growth Before 1820. Maddison Project Working Paper 4. Data downloaded from website: http://www.ggdc.net/maddison/maddison-project/data/mpd_2013-01.xlsx (Sep 12, 2013).

- Dahan, M., Tsiddon, D., 1998. Demographic Transition, Income Distribution, and Economic Growth. *Journal of Economic Growth* 3(1), 29–52.
- Deininger, K., Squire, L., 1996. A New Data Set Measuring Income Inequality. *World Bank Economic Review* 10(3), 565–591.
- Deininger, K., Squire, L., 1998. New ways of looking at old issues: Inequality and growth. *Journal of Development Economics* 57(2), 259–287.
- Desbordes, R., Verardi, V., 2012. Refitting the Kuznets curve. *Economics Letters* 116(2), 258–261.
- Frazer, G., 2006. Inequality and Development Across and Within Countries. *World Development* 34(9), 1459–1481.
- Galor, O., Tsiddon, D., 1996. Income Distribution and Growth: the Kuznets Hypothesis Revisited. *Economica* 63(250), S103–S117.
- Hastie, T., Tibshirani, R., 1986. Generalized additive models (with discussion). *Statistical Science* 1(3), 297–318.
- Hastie, T.J., Tibshirani, R.J., 1990. *Generalized Additive Models*. Chapman & Hall/CRC, New York.
- Huang, H.-C.R., 2004. A flexible nonlinear inference to the Kuznets hypothesis. *Economics Letters* 84(2), 289–296.
- Huang, H.-C.R., Lin, S.-C., 2007. Semiparametric Bayesian inference of the Kuznets hypothesis. *Journal of Development Economics* 83(2), 491–505.
- Kanbur, R. (2011). Does Kuznets Still Matter? Paper written for a Festschrift in honor of Montek Ahluwalia. <http://kanbur.dyson.cornell.edu/papers/DoesKuznetsStillMatter.pdf> (Nov 16, 2014).
- Kanbur, R., Zhuang, J., 2013. Urbanization and Inequality in Asia. *Asian Development Review* 30(1), 131–147.
- Kuznets, S., 1953. Shares of Upper Income Groups in Income and Saving. NBER Publication No. 55, New York.
- Kuznets, S., 1955. Economic Growth and Income Inequality. *American Economic Review* 45(1), 1–28.
- Leigh, A., 2007. How Closely Do Top Income Shares Track Other Measures of Inequality? *Economic Journal* 117(524), F589–F603.
- Lin, S.-C., Huang, H.-C.R., Weng, H.-W., 2006. A semi-parametric partially linear investigation of the Kuznets' hypothesis. *Journal of Comparative Economics* 34(3), 634–647.

- List, J.A., Gallet, C.A., 1999. The Kuznets Curve: What Happens After the Inverted-U? *Review of Development Economics* 3(2), 200–206.
- Piketty, T., 2001. *Les Hauts revenus en France au 20e siècle: inégalités et redistribution, 1901–1998*. Paris: B. Grasset.
- Piketty, T., 2003. Income Inequality in France 1901–1998. *Journal of Political Economy*, 111(5), 1004–1042.
- Piketty, T., Saez, E., 2003. Income Inequality in the United States 1913–1998. *The Quarterly Journal of Economics* 118(1), 1–39.
- Piketty, T., Saez, E., 2006. The Evolution of Top Incomes: A Historical and International Perspective. *American Economic Review* 96(2), 200–205.
- Robinson, S., 1976. A Note on the U Hypothesis Relating Income Inequality and Economic Development. *American Economic Review* 66(3), 437–440.
- Roine, J., Vlachos, J., Waldenström, D., 2009. The long-run determinants of inequality: What can we learn from top income data? *Journal of Public Economics* 93(7-8), 974–988.
- Roine, J., Waldenström, D., 2015. Long-Run Trends in the Distribution of Income and Wealth, in: Atkinson, A.B., Bourguignon, F. (Eds.), *Handbook of Income Distribution*, vol. 2A, North-Holland, Amsterdam, pp. 469–592.
- United Nations, 2012. Department of Economic and Social Affairs, Population Division. *World Urbanization Prospects: The 2011 Revision (POP/DB/WUP/Rev.2011/1/F2)*. Website: http://esa.un.org/unpd/wup/CD-ROM/WUP2011-F02-Proportion_Urban.xls (Nov 28, 2013).
- Wood, S.N., 2006. *Generalized Additive Models - An Introduction with R*. Chapman & Hall/CRC, Boca Raton FL.
- World Bank, 2014a. World Development Indicators, Employment in services (% of total employment). Website: <http://data.worldbank.org/indicator/SL.SRV.EMPL.ZS/> (Feb 25, 2014).
- World Bank, 2014b. World Development Indicators, Services, etc., value added (% of GDP). Website: <http://data.worldbank.org/indicator/NV.SRV.TETC.ZS/> (Feb 25, 2014).
- Zhou, X., Li, K.-W., 2011. Inequality and development: Evidence from semiparametric estimation with panel data. *Economics Letters* 113(3), 203–207.