WHAT DRIVES INEQUALITY?

RESEARCH ON ECONOMIC INEQUALITY

Series Editors

John A. Bishop
East Carolina University, USA

and

Juan Gabriel Rodriguez
Universidad Complutense de Madrid, Spain

RESEARCH ON ECONOMIC INEQUALITY, VOL. 27

WHAT DRIVES INEQUALITY?

EDITED BY

KOEN DECANCQ
University of Antwerp, Belgium

and

PHILIPPE VAN KERM
Luxembourg Institute of Socio-Economic Research (LISER) and University of Luxembourg, Luxembourg

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2020

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78973-378-5 (Print)
ISBN: 978-1-78973-377-8 (Online)
ISBN: 978-1-78973-379-2 (Epub)

ISSN: 1049-2585 (Series)

CONTENTS

GUEST EDITORS BIOS

Koen Decancq is an Associate Research Professor at the Herman Deleeck Centre for Social Policy at the University of Antwerp. He holds a PhD in Economics from the University of Leuven (Belgium). He has been a Visiting Professor at the Woodrow Wilson School of Public and International Affairs of the University of Princeton and has been the Ludwig M. Lachmann Research Fellow at the London School of Economics. He is currently a Fellow at the Department of Economics of the University of Leuven, The Centre for Philosophy of Natural and Social Science at LSE, ZEW Mannheim, and the Center for Operations Research and Econometrics at the University of Louvain.

His research interests are in welfare economics, the measurement of (multidimensional) inequality, poverty and well-being with a special focus on the incorporation of individual preferences and the role of social policies.

He is an Associate Editor of the *Journal of Economic Inequality* (since 2019) and Member of the Editorial Board of the Review of Income and Wealth (since 2018). He is currently the Country Team Leader for Flanders for the Survey of Health, Ageing and Retirement in Europe (SHARE).

Philippe Van Kerm is a Professor of Social Inequality and Social Policy at the University of Luxembourg on a joint appointment with the Luxembourg Institute of Socio-Economic Research (LISER). He holds a PhD in Economics from the University of Namur (Belgium). He has been Head of the Living Conditions department at LISER where he has also been scientific director (ad interim) and Head of the Graduate Studies Programme. He has been a Visiting Researcher at the London School of Economics and is currently a Fellow at the Institute for Social and Economic Research (University of Essex), ZEW Mannheim, and the Institute for New Economic Thinking (University of Oxford).

His research interests are in applied micro-econometrics, welfare and labour with particular reference to poverty and income distribution dynamics, wealth inequality, social mobility, wage, tax, social protection and social policy.

He is an Associate Editor of the *Journal of Economic Inequality* (since 2017) and of the *Stata Journal* (since 2018). He is currently the National Leader for Luxembourg in the NORFACE project 'The impact of childhood circumstances on individual outcomes over the life-course' (IMCHILD, 2018–2020).

CONTRIBUTORS BIOS

Eyal Bar-Haim is a Postdoctoral Researcher at the PEARL Institute for Research on Socio-Economic Inequalities at the University of Luxembourg. His research interests are mobility, stratification, inequality of opportunities and economic inequality.

Maurizio Bussolo is a Lead Economist at the World Bank, and previously at the OECD, and the Overseas Development Institute in London. He has extensively published in peer-reviewed journals on trade, growth, poverty and income distribution. He holds a PhD in Economics from the University of Warwick.

Begoña Cabeza obtained a research master's degree in Human Rights at the Universidad Carlos III de Madrid. She is currently a PhD candidate at the University of Antwerp (Herman Deleeck Centre for Social Policy). Her research focusses on beliefs about fairness, responsibility and redistribution.

Louis Chauvel is a Professor of Sociology at the University of Luxembourg where he leads the PEARL Institute for Research on Socio-Economic Inequality allocated by the Fonds National de la Recherche of Luxembourg. His scientific interests comprise income and wealth inequality, generations, middle class, demography, mortality, health and social policy.

Koen Decancq is an Associate Research Professor at the Herman Deleeck Centre for Social Policy at the University of Antwerp. His research interests are the measurement of (multidimensional) inequality, poverty and well-being with a special focus on the incorporation of individual preferences and the role of social policies.

Franziska Deutschmann is working as a Consultant in data science. She finished her PhD at the Department of Economics, at the University of Constance, Germany in 2018. Her research interest lie in the field of inequality, family economics, applied microeconomics and public economics.

Michael Förster is a Senior Economist and Policy Analyst at the OECD Directorate for Employment, Labour and Social Affairs. He has been leading successive OECD work on inequality, income distribution, poverty and social mobility over the past decade.

Antoine Genest-Grégoire is a Senior Research Associate of the Research Chair on Taxation and Public Finance and a Lecturer in Applied Political Studies and Taxation at Université de Sherbrooke, Canada. His research on perceptions of

taxes and transfers and their distributional impacts was showcased in various academic and policy forums and has been cited frequently in Canadian media.

Jean-Herman Guay is a Full Professor at the School of Applied Political Studies, Université de Sherbrooke. He is a recognised public opinion specialist and he has extensively contributed to media coverage of electoral politics. He currently oversees Perspective Monde, an online encyclopaedia of international trends, and recently published a book on the history of elections in Quebec, with Serge Gaudreau.

Luc Godbout is a Full Professor of Taxation and holds the Research Chair in Taxation and Public Finance, Université de Sherbrooke. He is a recognised public intellectual, his work is frequently cited in the media and in parliamentary debates. He Chaired an Expert Governmental Commission on tax reform in 2014–2015, and also sat on similar panels studying tax expenditures and the sustainability of pension plans, among others.

Anne Hartung is a Research Scientist at the Institute for Research on Socio-Economic Inequality at the University of Luxembourg. In the frame of the NORFACE/FNR project IMCHILD, she also investigates the effects that childhood conditions may have on later-life outcomes.

Carla Krolage is a PhD student and Junior Economist at the ifo Center for Macroeconomics and Surveys and the University of Munich. Her research focusses on empirical public economics, in particular on social policy and tax reforms.

Marko Ledić is a Postdoctoral Fellow at the Department of Economic Theory, University of Zagreb. His main research interests are in the field of empirical public, labour, health and welfare economics.

Mattia Makovec is an Economist at the World Bank, where he leads analytic activities on jobs and social protection in Europe and Central Asia. Previously, he held positions at Essex University, at the University of Chile and at the Ministry of Labor in Chile. He holds a PhD in Economics from Bocconi University.

Arip Muttaqien holds a PhD in Economics and Governance from Maastricht University, the Netherlands. Currently, he is an Evaluation Specialist at Secretariat of Accelerating Stunting Prevention at the National Team for the Acceleration of Poverty Reduction (TNP2K), under Secretariat of Indonesian Vice President. He is also an Affiliated Researcher at the United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT). His main research topics are health economics, labor economics, development economics, social protection, impact evaluation, poverty, inequality, microsimulation, and tax-benefit policy.

Cathal O'Donoghue is a Dean of Arts and Social Sciences and Professor of Public and Social Policy at the National University of Ireland, Galway (NUI Galway).

His research is mainly in the area of applied public economics, focussing on the impact of household sector, tax and social policy on poverty, inequality, work incentives and environmental pollution.

Andreas Peichl is the Head of the ifo Center for Macroeconomics and Surveys and Professor of Macroeconomics and Public Finance at the University of Munich. His research centres on empirical public economics, taxation and redistribution, and inequality. He holds a PhD in Economics from the University of Cologne.

Ivica Rubil is a Research Fellow at the Institute of Economics Zagreb. His research interests are at the intersection of welfare, public, labour and health economics, with a focus on empirical analysis of inequality, poverty and well-being.

Denisa Sologon is a Research Economist at Luxembourg Institute of Socio-Economic Research. Her research interests are welfare and labour economics, public policy analysis, social protection policy and applied econometrics.

Tsvetana Spasova is currently a Research Fellow at the University of Basel where she successfully completed a PhD in Economics and Statistics. Her research interests are mainly in the area of income distribution and inequality as well as econometric modelling.

Marc Stöckli is a PhD student and Junior Economist at the ifo Center for Macroeconomics and Surveys and the University of Munich. His research circles around various aspects of economic inequality with a focus on inequality of opportunity and redistributive justice.

Iván Torre is an Economist at the Office of the Chief Economist for Europe and Central Asia at the World Bank. His research focusses on political economy, income distribution and labour economics. He holds a PhD in economics from Sciences Po, Paris.

Philippe Van Kerm is a Professor of Social Inequality and Social Policy at the University of Luxembourg on a joint appointment with the Luxembourg Institute of Socio-Economic Research (LISER). His research interests are in applied micro-econometrics, welfare and labour with particular reference to poverty and income distribution dynamics, wealth inequality, social mobility, wage, tax, social protection and social policy.

Gerlinde Verbist is a Senior Researcher at the Centre for Social Policy Herman Deleeck, University of Antwerp. Her research mainly focusses on topics of inequality, tax-benefit policy evaluation and microsimulation modelling.

Christian Wittneben is a PhD student and Junior Economist at the ifo Center for Macroeconomics and Surveys and the University of Munich. His research interests are structural models of labour supply and automatic stabilisers.

PREFACE

This is the 27th volume of *Research on Economic Inequality* and concerns about inequality seem as vivid as ever since the start of this series. Researchers now dispose of a more mature methodological machinery and of new, increasingly rich, data sources. Inequality trends have become well-documented in many countries and the main determinants of recent trends are increasingly well-understood. Much less is known however about the driving forces behind international differences in inequality.

By soliciting contributions addressing the question 'What drives inequality?', we aimed to throw new light on the underlying drivers of inequality and to unpack the reasons for the wide variations in inequality across countries and over time. This endeavour may appear futile since these reasons may be so diverse and deep-rooted in the cultural, historical or geographical characteristics of countries that one can hardly expect comprehensive models or clear-cut causal inference. Yet, we are convinced that a better understanding of differences in inequality across countries (and over time) is important to address the next key question: 'What can be done?'

For this volume, we sought to attract papers discussing the role of labour markets, taxation, social protection and redistributive policies, but were also interested in papers studying the role of 'deeper drivers' such as political institutions, norms and attitudes and preferences for redistribution. Advancements to methodology and critiques on the cross-country comparability of inequality measures were also welcome. While the main discussion may be about income inequality, we hoped to attract contributions about wealth, consumption or other forms of inequalities.

The nine chapters collected in the volume address these dimensions. Chapters 1–4 examine income or expenditure inequality and discuss the role of tax policy and redistribution, demographics or labour market factors. Chapters 5–7 broaden the concept of welfare beyond income by incorporating measures of wealth, public goods and non-monetary dimensions in the analysis of inequality. Chapters 8 and 9 provide insights about individual perceptions, preferences and beliefs about inequality and redistribution.

Chapter 1, by Tsvetana Spasova, examines trends in income distributions and inequality in the European Union using data from the European Union Statistics on Income and Living Conditions. She uses the estimates of a parametric income distribution model to study the contribution of individual countries to inequality in broader regional aggregates – the 'old' and 'new' Member States – and shows that the 'new' EU countries have become richer and less unequal over the observed years, while the 'old' ones have experienced a small increase in inequality over the Great Recession years.

Chapter 2, by Maurizio Bussolo, Carla Krolage, Mattia Makovec, Andreas Peichl, Marc Stöckli, Iván Torre and Christian Wittneben, addresses the redistributive impact of taxes and benefits in 28 European Union countries across the Great Recession. Unlike most research on this topic, they examine impacts on both vertical and horizontal inequality. While they observe a significant degree of heterogeneity across countries, their results highlight horizontal inequality concerns as a dimension which policy-makers should take into account when reforming tax and transfer systems.

Chapter 3, by Franziska Deustchmann, zooms in on long-term income inequality trends and on the East–West contrast in Germany. Using various counterfactual methods, the paper quantifies the impact on inequality of differences in socio-economic characteristics over time and across East and West Germany. The prevalence of singlehood accounts to a large extent for the observed increase in inequality over time (along with a change of employment among males and single females). Differences in employment and household sizes also drive the difference in inequality observed between East and West Germany.

Chapter 4, by Arip Muttaqien, Cathal O'Donoghue and Denisa Sologon, offers a novel cross-national contrast with an analysis of differences in inequality in household expenditure between India and Indonesia (which together account for about 20% of the world population). Despite many similarities between the countries, Indonesia now exhibits higher inequality than India following a relatively sharp increase over the last 15 years. The decomposition of the difference across countries reveals that the gap is mostly accounted for by differences in education and the return to education, rather than by differences in work and employment structures. A large part of the gap remains 'unexplained' however.

The volume then moves beyond the sole income and expenditure dimensions.

Chapter 5, by Gerlinde Verbist and Michael Förster, examines the distributional implications of publicly provided free or subsidised services. Many important services are not provided (exclusively) through the market such as, for example, education, housing, health care, etc. They are not fully paid from household income but contribute to household welfare, so ignoring their contribution in the assessment of social inequality is potentially misleading, especially in international comparisons. This chapter reviews the main methodological approaches and presents empirical results for 27 OECD countries. The authors find that indicators of inequality based on extended income measures that add an imputed value of public services to household cash incomes can be up to a third smaller than inequality in cash income alone. This finding is important for cross-country policy comparisons.

Chapter 6, by Louis Chauvel, Anne Hartung, Eyal Bar-Haim and Philippe Van Kerm, brings wealth into the picture. The importance to study wealth inequality alongside income inequality is increasingly appreciated by economists and sociologists alike. The study exploits the 'isograph' as a tool to describe income and wealth distributions, and to present fine-grained information about the upper tail of these distributions. Using combined data from the Eurozone Household Finance and Consumption Survey and the US Survey on Consumer Finance, this chapter illustrates how much more unequal is the distribution of wealth, especially in the United States when compared to 16 European countries.

Chapter 7, by Marko Ledić and Ivica Rubil, introduces a multidimensional measure of well-being that incorporates unemployment, health, housing, crime and environment besides income. This chapter uses tools from the literature on taxation and redistribution to decompose the difference between the inequality in the multidimensional well-being measure and standard income in two parts: a vertical and reranking effect. The authors implement the decomposition with data from the European Quality of Life Survey for 27 European Union countries in 2011. They find that inequality is higher for the multidimensional measure and that the reranking effect accounts for a large part of the inequality difference, with health contributing most to both effects.

The last two chapters of the volume examine how people perceive inequality, and how perceptions and beliefs can shape attitudes towards redistribution policies.

Chapter 8, by Antoine Genest-Grégoire, Jean-Herman Guay and Luc Goodbout, studies who believes to belong to the middle class and how that affects their support for higher taxes on the rich. According to the so-called reference-group theory, most citizens perceive to be situated in the middle of their (non-representative) reference group. The authors test this theory with an online survey in the Canadian province of Quebec. They find that a sizeable share of objectively rich persons place themselves in the middle class. These respondents are found to support higher taxes on the rich, without realising that they are actually part of this group.

Chapter 9, by Begoña Cabeza and Koen Decancq, investigates how beliefs about the influence of effort have been affected by the Great Recession in Spain. The beliefs about the influence of effort have been found in the literature to be an important determinant of the demand for redistribution. The authors use a series of Spanish public opinion surveys between 2010 and 2018, matched with regional-level unemployment data and find that people attribute a larger role to luck in provinces where the unemployment rate increased more during the Great Recession. Moreover, lower educated individuals and those who position themselves as more left-wing, are found to have adjusted their beliefs more.

We trust the nine chapters collected in this volume provide useful contributions towards a better understanding of the question 'What drives inequality?', although they far from exhausted the theme! The chapters in this volume are steps forward and will hopefully help addressing the bigger 'What can be done?' challenge.

To conclude, we want to thank John Bishop and Juan Gabriel Rodriguez, the series editors, for having invited us to edit this 27th volume. We also thank all contributors and reviewers who made the compilation of this book possible. Their efficient, professional and timely work made our task as guest editors easy and stress-free. Funding from LISER, the University of Luxembourg and the University of Antwerp for the organisation of a thematic workshop in October 2018 at the early stage of preparing the volume is gratefully acknowledged.

Koen Decancq
University of Antwerp

Philippe Van Kerm
University of Luxembourg
and Luxembourg Institute of Socio-Economic Research

This page intentionally left blank

CHAPTER 1

REGIONAL INCOME DISTRIBUTION IN THE EUROPEAN UNION: A PARAMETRIC APPROACH

Tsvetana Spasova

ABSTRACT

This chapter studies trends in income distributions and inequality in the European Union using data from the European Union Statistics on Income and Living Conditions. The author models the income distribution for each country under a Dagum distribution assumption and using maximum likelihood techniques. The author uses parameter estimates to form distributions for regions defined as finite mixtures of the country distributions. Specifically, the author studies the groups of 'new' and 'old' countries depending on the year they joined the European Union. The author provides formulae and estimates for the regional Gini coefficients and Lorenz curves and their decomposition for all the survey years from 2007 through 2011. The estimates of this study show that the 'new' European Union countries have become richer and less unequal over the observed years, while the 'old' ones have undergone a slight increase in inequality which is however not significant at conventional levels.

Keywords: Income distribution; Dagum distribution; finite mixtures; inequality; Gini decomposition; European Union

JEL classifications: D31; D63; C13

What Drives Inequality?
Research on Economic Inequality, Volume 27, 1–18

ISSN: 1049-2585/doi:10.1108/S1049-258520190000027002

1. INTRODUCTION

The European Union experienced several enlargements since the establishment of its predecessor, the European Economic Community, in 1957, from only six founding states – Belgium, France, Italy, Luxembourg, the Netherlands and (West) Germany – to 28 Member States today. One of the major enlargements happened in 2004 when countries from Central and Eastern Europe joined the European Union. The change in composition modified the income distribution and inequalities within the European Union. Even though it is currently the world's largest economy, generating a nominal gross domestic product of approximately 14.303 trillion Euros according to International Monetary Fund (2014), if we look into its Member States individually, we see large differences in the income distribution between and within them.

This chapter adds to the literature on the distribution of income and inequality in the European Union, both for individual countries and regionally. The income distribution in European Union countries has been much researched either as part of the world distribution of income with inequality analyses based on grouped income data (for instance, Chotikapanich, Griffiths, & Prasada Rao, 2007; Chotikapanich, Griffiths, Prasada Rao, & Valencia, 2012; Milanovíc, 2002, 2005, 2012; Sala-i-Martín, 2006), or separately in inequality analyses at the country level (e.g. Filauro, 2017; Tóth & Medgyesi, 2011). Jenkins, Brandolini, Micklewright, & Nolan (2013) have studied the evolution of income distribution during the Great Recession in 21 rich OECD countries, Brzezinski (2018) has analysed the income inequality in Central and Eastern Europe, while Anderson, Pittau, Zelli, & Thomas (2018) have focussed on income classification in the Euro zone as an entity.

Like many recent studies, we use representative microdata from the European Union Statistics on Income and Living Conditions (EU-SILC) cross-sectional survey to study both inequality in individual countries, and in broader country groupings as 'new' and 'old' Member States. The 'new' countries are those which entered the European Union after 2004 and the 'old' are those which entered before 2004. We selected 2004 as the splitting year since this marked the largest expansion of the European Union.

This work makes three contributions to the literature. First, we provide new parametric model estimates for the income distribution of the European Union as a whole, for multicountry regional groupings, and for individual countries for each year from 2007 to 2011. We obtain the regional models as finite mixtures of the individual countries' distributions. To derive the models, we fit the Dagum distribution via maximum likelihood techniques to the income data available for each of the European Union countries. Second, we provide formulae for the Gini coefficient and the Lorenz curve implied by Dagum distribution mixture models. We introduce an efficient way for computing the total Gini coefficient numerically and decompose the regional Gini coefficients into within- and between-country contributions. Third, our results show that the region formed by the 'new' Member States is more unequal and less wealthy than the region formed by the 'old' ones and we observe that inequality in the 'new' countries

contributes substantially to overall inequality for the European Union as a whole. However, looking at the evolution of income distribution over time, we find that the 'new' Member States have become, on average, wealthier and more equal over time, while the 'old' Member States have undergone a slight increase in inequality. We provide **R** code for replicating estimations in an Online Supplementary Material (R Core Team, 2014).

Using parametric models for studying income distributions has several advantages. We can represent the income distribution of a country with a small number of estimated parameters (Chotikapanich et al., 2007, 2012; Hajargasht, Griffiths, Brice, Rao, & Chotikapanich, 2012), from which the distribution in larger entities, in our case regions, can be obtained in a straightforward way. We exploit this in Section 4 below. The model parameters often also possess an economic interpretation, which allows to gain insights about the causes of the evolution of income distribution over time or interpret the differences between income distributions across countries (Brzezinski, 2013). Explicit formulae are available for many poverty and inequality measures as functions of the parameters of the theoretical income distribution. Benefits of parametric models in terms of estimation stability are also put forward in Graf & Nedyalkova (2014).

Specifically, the Dagum distribution has been used successfully for fitting data from various sources (Dagum, 1977; García Pérez & Prieto Alaiz, 2011; Kleiber & Kotz, 2003). Dagum (1977) aimed to find a distribution that would capture the heavy tails present in wealth distributions as well as permitting interior modes, thereby outperforming the more classical Pareto and lognormal distributions. In a comprehensive empirical study involving 11 parametric models and 23 countries, Bandourian, McDonald, & Turley (2003) observed that the Dagum distribution was the best-fitting three-parameter distribution in more than 80% of the cases. Kleiber (2008) provides further references on the empirical performance of the Dagum distribution. The distribution may sometimes be outperformed by a distribution with additional parameters such as the generalised beta distribution of the second kind (GB2), but the effect is often marginal (Bandourian et al., 2003) at the cost of introducing significant empirical and analytical complexity. Our analysis confirms the good performance and the tractability of the Dagum distribution for modelling income distributions.

The chapter is structured as follows. The EU-SILC data are described in Section 2. Section 3 collects some basic properties of the Dagum distribution, describes model fitting via maximum likelihood and bootstrap inference, and provides an assessment of goodness-of-fit. Also, in Section 3.3, we give analytical expressions for the regional Lorenz curves and Gini coefficients. Country-specific and regional results appear in Section 4. Finally, Section 5 provides concluding remarks.

2. DATA

The EU-SILC provides nationally representative data on income, poverty, social exclusion and living conditions for all of the European countries. The EU-SILC

survey for each country is provided to the statistical office of the European Union (Eurostat) by the relevant national statistical institutes which collect the data according to a common overarching methodology suggested by Eurostat. EU-SILC is the basis for calculation of commonly agreed indicators on poverty and social inclusion in EU countries (Atkinson, Guio, & Marlier, 2017). EU-SILC data have also been used by academics for income modelling and inequality analysis across Europe (see e.g. Anderson et al., 2018; Aristei & Perugini, 2010; Filauro, 2017; Graf & Nedyalkova, 2014; Longford, Pittau, Zelli, & Massari, 2012; Tóth & Medgyesi, 2011) and for examining poverty measures (e.g. Fabrizi, Ferrante, Pacei, & Trivisano, 2011; Jenkins & Van Kerm, 2011).

We use EU-SILC cross-sectional survey data for the years 2007–2011. The income reference period is one year earlier than the year of the survey, since the total income collected in EU-SILC is the income for the calendar year previous to the interview (except for the UK and Ireland; see Online Supplementary Material B). We model and compare the distributions of personal income for each of the European Union countries except Ireland (as it is not included in the EU-SILC 2011 survey), Malta (since it is not included in the EU-SILC 2007 and 2008 surveys) and Croatia (since it entered the European Union in 2013). Table 5, in the Online Supplementary Material B, presents descriptive statistics for the 2011 data.

We focus on the equivalised disposable income computed in purchasing power parities and apply cross-sectional weights to account for population size. For more details on the variables used, see again Online Supplementary Material B.

Section 4 presents an analysis of European Union regions composed of 'new' and 'old' countries depending on the year they joined the European Union (after or before 2004). In Table 1, we provide the so-defined 'old' and 'new' European Union countries along with country codes in brackets as given by Eurostat (2011). Table 1 can be used as a reference for the 'old' and 'new' regions and their respective graphs and explanations provided later in this work. From now on, whenever we refer to the (whole) European Union in this work, we mean the countries listed in Table 1 under Old EU Member States plus the New EU Member States, excluding Croatia, Malta and Ireland.

Table 1. European Union Regional Classification.

Old EU	New EU	EU
Austria (AT)	Bulgaria (BG)	All without:
Belgium (BE)	Cyprus (CY)	Croatia (HR)
Denmark (DK)	Czech Republic (CZ)	Ireland (IE)
Finland (FI)	Estonia (EE)	Malta (MT)
France (FR)	Hungary (HU)	
Germany (DE)	Latvia (LV)	
Greece (EL)	Lithuania (LT)	
Italy (IT)	Poland (PL)	
Luxembourg (LU)	Romania (RO)	
Netherlands (NL)	Slovakia (SK)	
Portugal (PT)	Slovenia (SI)	
Sweden (SE)		
United Kingdom (UK)		

3. METHODOLOGY

This section describes the methodology which we applied for fitting the income data from EU-SILC using Dagum distributions. In Section 3.1, we provide some basic characteristics of the distribution. In Section 3.2, we explain how we employ the maximum likelihood approach for model fitting. Section 3.3 provides all the necessary components for regional analysis of income distribution and inequality with the Dagum distribution. It gives closed-form expressions for the regional densities, the regional Lorenz curves, the between-country and within-country Gini coefficients and explains how regional Gini coefficients were estimated. Finally, Section 3.4 describes a parametric bootstrap method that was used to obtain standard errors.

3.1 The Dagum Distribution

The Dagum distribution is a three-parameter distribution, $D(\eta)$, where η is the triple (a, b, p). We use a parametrisation of the Dagum distribution given in Kleiber and Kotz (2003) that slightly differs from the parameterisation originally used in Dagum (1977). Its density is

$$f(x;\eta)=\frac{apx^{ap-1}}{b^{ap}\left[1+(x/b)^{a}\right]^{1+p}},\quad x>0, \tag{1}$$

where a, b and p are positive real numbers. When η is obvious from the context, we write only $f(x)$.

The cumulative distribution function can be written in closed form as

$$F(x;\eta)=\left[1+\left(\frac{x}{b}\right)^{-a}\right]^{-p},\quad x>0. \tag{2}$$

The quantile function can also be written in closed form as

$$Q(u;\eta)=b\left[u^{-\frac{1}{p}}-1\right]^{-\frac{1}{a}},\quad 0<u<1. \tag{3}$$

The mean of the Dagum distribution equals

$$\mu=\frac{b\Gamma(p+1/a)\Gamma(1-1/a)}{\Gamma(p)} \tag{4}$$

where $\Gamma(p)$ is the gamma function.

The Lorenz curve of the Dagum distribution is

$$L(u)=I_{z}\left(p+\frac{1}{a},1-\frac{1}{a}\right),\quad 0\le u\le 1, \tag{5}$$

where $z=u^{1/p}$ and $I_{z}(p,q)$ is the incomplete beta function ratio defined as $I_{z}(p,q)=\frac{1}{B(p,q)}\int_{0}^{z}u^{p-1}(1-u)^{q-1}\,du$, $0\le z\le 1$, with $B(p,q)$ the beta function (Kleiber & Kotz, 2003).

The Gini coefficient is

$$G = \frac{\Gamma(p)\Gamma(2p+1/a)}{\Gamma(2p)\Gamma(p+1/a)} - 1. \tag{6}$$

3.2 Estimation

We employ maximum likelihood to estimate the parameters of the distribution. To account for unequal sampling probabilities, we weight the likelihood by the cross-sectional weights provided with the data. Let N be the number of people in the given sample, x_i the equivalised income of person i and w_i the cross-sectional weight of person i. The weighted log-likelihood $l(\eta)$, with $\eta = (a, b, p)$, is

$$l(\eta) = \sum_{i=1}^{N} w_i \log\big(f(x_i;\eta)\big), \tag{7}$$

where $f(x; \eta)$ is the Dagum density given in formula (1).

We maximise the log-likelihood function $l(\eta)$ with respect to the Dagum distribution parameters a, b and p for each country and year using the **R** programming language (R Core Team, 2014). For optimisation, we use the `nlminb` function. The initial values a_0, b_0 and p_0 for the parameters a, b and p are $a_0 = 2$ and $p_0 = 0.4$ for all countries, whereas for each country b_0 is set to the mean income of the respective country.

3.3 Regional Income Distribution and Inequality

Once we have estimated the three parameters of the Dagum distribution for each country, we form groups of countries and compute the regional income distribution and inequality for each region. This can be achieved by computing regional densities and distribution functions which are sums of the densities, or respectively distribution functions, of all countries in a given region weighted by their population sizes. Formally, given K countries each with parameter vector η_k, $k = 1,\ldots,K$, density functions $f_k(x) = f(x;\eta_k)$, and population shares $\pi_1, \pi_2, \ldots, \pi_K$, the regional density is given by (Chotikapanich et al., 2012)

$$f(x) = \sum_{k=1}^{K} \pi_k f_k(x), \tag{8}$$

with $f_k(x)$ as in equation (1). The regional cumulative distribution function is

$$F(x) = \sum_{k=1}^{K} \pi_k F_k(x), \tag{9}$$

with $F_k(x) = F(x;\eta_k)$ given in equation (2). The population shares $\pi_1, \pi_2, \ldots, \pi_K$ are computed using the total population sizes (see Online Supplementary Material B). The regional mean income is

$$\mu = \sum_{k=1}^{K} \pi_k \mu_k, \tag{10}$$

with $\mu_k(x)$ as given in equation (4).

The regional cumulative income shares $\psi(x)$ are analogous to the ones given by Chotikapanich et al. (2012) for the beta-2 distribution. Here, for the Dagum distribution the cumulative income shares, $\psi(x)$, are computed as

$$\begin{aligned}\psi(x) &= \frac{1}{\mu}\int_0^x zf(z)dz = \frac{1}{\mu}\sum_{k=1}^{K}\pi_k\int_0^x zf_k(z)dz \\ &= \frac{1}{\mu}\sum_{k=1}^{K}\frac{\pi_k b_k I_y\left(p_k+\frac{1}{a_k},1-\frac{1}{a_k}\right)\Gamma\left(p_k+\frac{1}{a_k}\right)\Gamma\left(1-\frac{1}{a_k}\right)}{\Gamma(p_k)}\end{aligned} \tag{11}$$

where $I_y(p, q)$ is the incomplete beta function ratio defined as above, now with $y = \frac{(x/b)^a}{1+(x/b)^a}$ and μ as given in equation (10). To graphically represent inequality, we obtain Lorenz curves by plotting the regional cumulative income shares $\psi(x)$ (given in equation 11) against the regional cumulative shares of population $F(x)$ (given in equation 9).

Finally, the regional Gini coefficient can be written as (Chotikapanich et al., 2012)

$$G = -1+\frac{2}{\mu}\sum_{j=1}^{K}\sum_{i=1}^{K}\pi_j\pi_i\int_0^{\infty} yF_j(y)f_i(y)dy \tag{12}$$

where μ is the regional mean income given in equation (10), $F_j(y)$ is the distribution function for country j given in equation (2), and $f_i(y)$ is the income density for country i given in equation (1). The integral appearing in (12) can be estimated numerically. We have split the integration into ranges and summed the results up, using the function `integrate` in **R**, which performs adaptive quadrature.

The regional Gini coefficient can be decomposed into a within-country and a between-country component (along with an interaction term) to capture how much aggregate inequality is driven by income differences across countries and how much is driven by income differences within countries: $G = G_B + G_W + I$ (see Lambert & Aronson, 1993).

The first term G_B captures how much differences in income between countries accounts for the aggregate inequality and is obtained if every income in every country is replaced with the mean income of the relevant country. We compute the between-country Gini coefficient G_B as (Chotikapanich et al., 2012; Lambert & Aronson, 1993)

$$G_B = \frac{1}{2\mu}\sum_{j=1}^{K}\sum_{i=1}^{K}\pi_j\pi_i\left|\mu_j-\mu_i\right|, \tag{13}$$

where μ_i is the mean income for country i given in equation (4), and μ is the regional mean income given in equation (10).

G_W measures the contribution of within-country inequality and is obtained as a weighted sum of the Gini coefficients for all countries (see Chotikapanich et al., 2012; Lambert & Aronson, 1993)

$$G_W = \sum_{j=1}^{K} \pi_j s_j G_j, \tag{14}$$

the weights are the products of the population shares π_j and income shares $s_j = \pi_j \mu_j / \mu$, and G_j is the Gini coefficient for country j as given in equation (6).

The interaction term I is the difference between the regional Gini coefficient and the between-country and the within-country Gini coefficients, namely $G = G_B + G_W + I$. I is zero if the income ranges for each country do not overlap. Recently, Anderson et al. (2018) used the interaction term to define a 'non-segmentation factor'.

Fig. 1 provides a graphical representation of the Gini decomposition for two imaginary countries A and B. The total Gini coefficient, G, for countries A and B is twice the area between the diagonal line of perfect equality and the Lorenz curve which is the solid black curve on the plot. The between-country Gini coefficient is twice the area between the diagonal line of perfect equality and the perfect equality lines for countries A and B when all their citizens receive incomes equal to the mean income of the respective country. The within-country Gini components are twice the area between the between-country Gini components perfect equality lines and the Lorenz curves corresponding to the weighted Gini

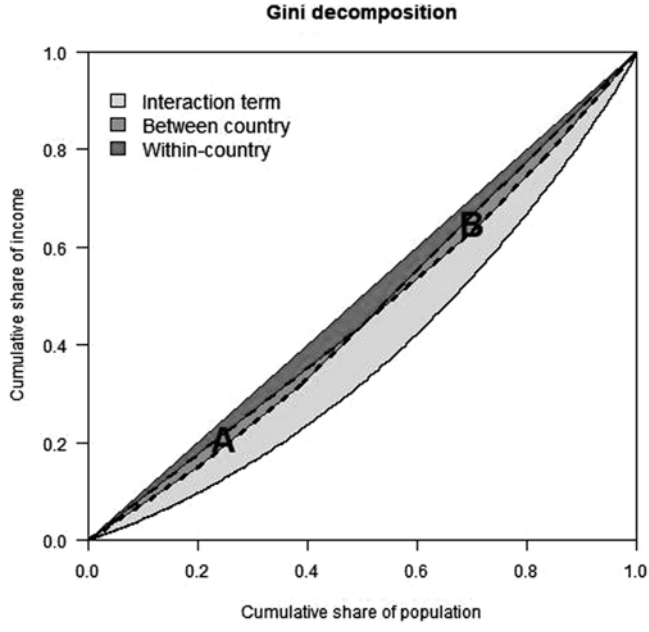

Fig. 1. Gini Decomposition.