Bremen, March 2026

Ivo Mossig
Hannes Lehmann
Marius Bode
Carmen Cordes
Nils Düpont

Filling Data Gaps in the Measurement of Income Inequality. A Complete Dataset of National GINI Coefficients 1995-2019

WeSIS — Technical papers

No. 24

Global Dynamics of Social Policy CRC 1342

Gefördert durch DFG Deutsche Forschungsgemeinschaft

Ivo Mossig, Hannes Lehmann, Marius Bode, Carmen Cordes, Nils Düpont
Filling Data Gaps in the Measurement of Income Inequality. A Complete Dataset of National GINI Coefficients 1995-2019
SFB 1342 Technical Paper Series, 24
Bremen: SFB 1342, 2026

in alphabetical order

Carmen Cordes 0009-0000-5498-0823
Nils Düpont 0000-0002-4766-9540
Hannes Lehmann 0009-0001-6430-6194
Ivo Mossig 0000-0002-2305-113X

SFB 1342 Globale Entwicklungsdynamiken von Sozialpolitik /
CRC 1342 Global Dynamics of Social Policy

Postadresse / Postal address:
Postfach 33 04 40, D - 28334 Bremen

Website:
https://www.socialpolicydynamics.de

[DOI https://doi.org/10.26092/elib/5604]
[ISSN 2700-0389]

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG)
Projektnummer 374666841 – SFB 1342

Ivo Mossig
Hannes Lehmann
Marius Bode
Carmen Cordes
Nils Düpont

Filling Data Gaps in the Measurement of Income Inequality. A Complete Dataset of National GINI Coefficients 1995-2019

SFB 1342
No. 24

Filling Data Gaps in the Measurement of Income Inequality. A Complete Dataset of National GINI Coefficients 1995-2019

Ivo Mossig*, Hannes Lehmann*, Marius Bode*, Carmen Cordes**, and Nils Düpont*

Index

About the authors:

*CRC 1342 Global Dynamics of Social Policy, University of Bremen; ***University of Bremen*

Corresponding authors:

Ivo Mossig, mossig@uni-bremen.de
Nils Düpont, duepont@uni-bremen.de

1. Abstract

Income inequalities are a major societal challenge (Grusky 2018; Polacko 2021). Despite the criticism that is being expressed, the Gini coefficient - especially income based - remains the most important indicator for measuring the extent and development of income inequality within a country. Unfortunately, Gini coefficients based on comparable methodologies are only available to a very limited extent. The most comprehensive data set available with consistent definitions for net income is the WIID Gini. With around 900 data points, this data set covers only 22% of the possible country-year combinations for the selected sample of 160 countries between 1995 and 2019.

We pursue two objectives: (1) to close existing data gaps through statistical imputation thereby creating a consistent and plausible dataset of Gini coefficients for 160 countries with over 1 Mio. inhabitants from 1995 to 2019 and (2) to identify the socioeconomic and political indicators that most strongly influence these imputations. To achieve this, missing data are estimated using a gradient boosting machine (GBM) drawing on over 1.400 socioeconomic and political indicators from the WeSIS database.

With this novel dataset, we enable researchers to broaden their inquiry into causes and effects of socio-economic inequality on a formerly unachievable scale.

Keywords: socio-economic inequality, gini coefficient, imputation, machine learning, gradient boosting, social policy

Introduction

Income inequalities are a major societal challenge. Accordingly, research on their causes, consequences, and policy solutions has become central to social science (Dabla-Norris 2015; Grusky 2018; Polacko 2021). Despite the criticism that is being expressed, the Gini coefficient remains the most important indicator for measuring the extent and development of income inequality within a country (De Maio 2007). Common sources include the OECD, World Bank, Luxembourg Income Study (LIS), and World Income Inequality Database (WIID). Unfortunately, Gini coefficients based on comparable methodologies are only available to a very limited extent. From geographical perspective, data gaps are especially prevalent for countries in the Global South. The data gaps increase with temporal distance. Even for the Global North, pre-2005 data are often incomplete or missing. Therefore, the technical paper pursues two objectives: (1) to close existing data gaps through statistical imputation and thereby create a consistent and plausible dataset of Gini coefficients for 163 countries with over 1 mio. inhabitants (1995-2019) and (2) to identify the socioeconomic and political indicators that most strongly influence these imputations. Missing data are estimated using a gradient boosting machine (GBM) drawing on over 1.400 socioeconomic and political indicators from the WeSIS database (www.wesis.org; Mossig and Obinger 2023).

The paper is structured along 6 chapters. The first and second chapter introduce and discuss the Gini coefficient as a concept, as well as already existing data. Chapter Three builds the theoretical foundation for explaining and measuring inequalities within countries. In chapter Four, our method is described, and the results are presented in Chapter Five. Chapter Six summarizes the results and concludes.

1. The Gini Coefficient

The Gini coefficient is *the* standard indicator to measure income inequality. It quantifies the deviation between a perfectly equal income distribution and the observed one, based on the Lorenz curve of cumulative income distribution. On this curve, the cumulative share of income (y-axis) is plotted against the cumulative share of the population (x-axis), ordered by income. The coefficient equals the ratio of the area between the Lorenz curve and the 45-degree line of perfect equality to the total area under that line. A value of 0 indicates perfect equality, whereas 1 represents maximum inequality. The coefficient was introduced in 1912 by Corrado Gini. A comprehensive discussion of its mathematical foundations, variations, and potential applications can be found in Yitzhaki and Schechtman (2013).

2. Critical Evaluation and Data Gaps

When employing Gini coefficients, it is important to be aware of their key limitations and to consider alternative measures of inequality when appropriate (Blesch et al. 2022; Neves Costa and Pérez-Duarte 2019). The Gini coefficients is not additively decomposable, meaning overall inequality cannot be expressed as within- and between-group inequalities. For analyses requiring such decomposability, the Theil index is preferable (Conceicao and Ferreira 2000). Moreover, the Gini coefficient is most sensitive to changes in the middle of the income distribution: small redistributions among middle-income groups affect it more strongly than larger shifts at the extremes (Gastwirth, 2017). To address this, alternative indices such as the Atkinson index have been developed (Atkinson 1970). Finally, the Gini coefficient does not account for the income levels. Identical values can mask substantial differences in living conditions if average incomes diverge. Firebaugh (2006) highlighted general challenges in the use of income data, such as purchasing power disparities and the problematic assumption that income directly reflects productivity.

Despite criticism regarding its limitations, the Gini coefficient remains the most widely used indicator to assess income inequality within countries (De Maio 2007). Common sources include the OECD, World Bank, Luxembourg Income Study (LIS), and World Income Inequality Database (WIID). However, Gini coefficients based on comparable methodologies are only available to a very limited extent. The most comprehensive data for 1995–2019, covering 970 country-years with consistent definition, can be found in the WIID dataset (UNU-WIDER 2025a). Though the WIID companion has 1759 datapoints for the same time period, the data series is a combination of different techniques and concepts to combine different sources and fill up missing data (UNU-WIDER 2025b). In our view, the companion thus lacks an internally consistent data series (the selection and advantages of this databases over alternative Gini datasets is explained in detail in Section 4).

3. Causes of Income Inequalities Within a Country

Before creating a new dataset by predicting missing Gini data, we review the literature explaining income inequality within countries. This ensures that the most important factors from existing research are taken into account in our methodological approach and the explanatory variables used from the WeSIS data information system, and that our results can be interpreted reliably.

Although drivers of income inequality differ across countries, common patterns emerge in the literature. At least four main causal complexes of income inequality can be distinguished: (i) economic factors, (ii) political and institutional factors, (iii) education and demographics, and (iv) further determinants e.g. external shocks or crises (Dabla-Norris et al. 2015; Huber and Stephens 2014; OECD 2011).

As economic causes, the literature gives various explanations that attribute income inequality to structural characteristics of labor markets, particularly the segmentation between highly paid, high-skilled occupations and low-skilled jobs (Osterman 1975). Similarly, high shares of informal employment and unemployment are associated with higher Gini coefficients (Chong and Gradstein 2007; Cysne 2009; Kum 2024). Globalization processes (Mossig and Lischka 2022) and their effects on income opportunities have also been extensively analysed. However, these effects differ across regions, resulting in a heterogeneous empirical picture. Globalization indicators such as trade openness, foreign direct investment, portfolio financial capital flows, or remittances show widely varying impacts on income inequality measured by Gini coefficients (Dorn et al. 2022; Eichengreen et al. 2021). The literature further addresses technological progress and its influence on income disparities. Rationalization potentials arising from technological developments are particularly pronounced in certain labor market segments and therefore put pressure on income earners in those areas (Hartmann et al. 2017). More recently, discussions have emerged regarding the potential effects of the growing use of generative AI on income inequality (Cazzaniga et al. 2024).

Common political and institutional causes of income inequality, mentioned in the literature, are frequently linked to a country's tax and economic policies and to the role of welfare state social policies. Studies have shown that progressive tax systems, with tax rates increasing with income, reduce income inequality and consequently the Gini coefficient. Duncan and Sabirianova (2016) confirm this, but also show that the effect is smaller when looking at long-term wealth inequality measured by consumption-based Gini coefficients. Both high average and marginal tax rates are of importance. The higher these rates, the lower the Gini coefficients of income inequality (Eydam and Qualo 2024). Progressive income taxation reduces income inequality especially when accompanied by an effective redistribution mechanism, which can be achieved primarily through social policy measures and therefore depends on the strength of the welfare state. Accordingly, it is argued that social policy reduces income inequality (Cammeraat 2020; Mossig and Düpont 2020). However, in OECD countries since the mid-1990s, market-oriented forces and austerity policies in the sense of a “race to the bottom” (Kvist 2004) have reduced the equalizing effect of social transfers, leading to increasing income inequality (Obst 2014; Wagmiller et al. 2020). Likewise, in formerly socialist states, previously low levels of income inequality increased as a result of market-oriented transformations with the expansion of private-sector activities and the retrenchment of the redistributive state following the shift in economic policy frameworks (Bandelj and Mahutga 2010; Henderson et al. 2008).

To explain existing income inequalities, social factors such as education, (anti)discrimination, and demography are also considered. Different educational opportunities lead to unequal incomes within a country. Longer and better education creates human capital that can be converted to disproportionately higher income opportunities in knowledge-based economic sectors (de Georio and Lee 2002; Yang and Gao 2018). Discrimination in access to well-paid jobs works in the opposite direction. Although an increasing number of states have joined the International Labour Organization’s (ILO) anti-discrimination conventions (Hahs 2022; Seitzer 2022), ethnic minorities still face disadvantages in job searches and existing employment relationships, as reflected in the ethnic pay gap (Ayaita 2023). Empirically, discrimination during the job search plays the largest role (Brynin and Güveli 2012). A similar pattern can be seen in the gender pay gap. The International Monetary Fund has found a strong correlation between Gini coefficients and the gender pay gap as well as other gender-related indices (Gonzales et al. 2015). Demographic factors can also contribute to income inequality. Income inequality may increase in an ageing society, as pension incomes are comparatively low. Moreover, public pension systems are put under pressure, especially when productivity growth is insufficient to offset the decline in the working-age population (Dolls et al. 2019). The extent to which income inequality is related to an ageing population depends on the pension formula of a given country (von Weizsäcker 1995).

Among other determinants of income inequalities within a country, external shocks (e.g., natural disasters or pandemics) and crises (e.g., financial and economic crises) have often been examined in response to contemporary events. With regard to the potential impacts of climate change, SenGupta and Atal (2024) demonstrated that rising climate change indices lead to a significant increase in Gini coefficients. This effect is particularly strong in countries that are already characterized by high levels of income inequality. Studies on the COVID-19 pandemic have shown that low-income population groups are disproportionately affected by the consequences of a pandemic, thereby increasing inequalities (Furceri et al. 2020; Esseau-Thomas et al. 2022). Equally, research on the Great Recession have shown how economic crises can intensify income inequality (Pfeffer et al. 2013), while other studies examining the 2007/08 financial crisis have reached similar conclusions (Bodea et al. 2021).

This overview of the main factors influencing Gini coefficients as an indicator of income inequality within a country is neither exhaustive, nor do we want to discuss how the various groups of factors are interlinked or mutually dependent. Instead, it highlights the multifaceted nature of income inequality and the importance of different causes within these causal complexes. This contextualizes the subsequent analysis and justifies our decision to not rely on selected single indicators to predict missing Gini coefficients as the suitability of any individual indicator is debatable in a heterogeneous and complex context. Instead, we adopt a data-driven approach, utilizing a broad set of available socio-economic and political country-level data to train an advanced model from Computational Social Science for estimating missing Gini coefficients. Nonetheless, the foundation of our approach lies in the connection between the causes of inequality explained in Chapter 3 and the data from WeSIS with income inequality.

4. Method

To reach our goal of creating a consistent Gini index for as many countries as possible, we leveraged the vast amount of data on global social policies in our Welfare State Information System (WeSIS). We opted for the time period 1995-2019 (25 years); an earlier starting point is not feasible because the data situation after the end of the Cold War in the early 1990s was very volatile, especially for the Central and Eastern European transition countries. This would have led to incomprehensible results. Similar to Windzio et al. (2022), our country sample comprises 163 countries. A complete dataset with Gini coefficients would therefore comprise 4,100 country-year combinations.

We ran a prediction of net income Ginis with every numerical indicator collected in WeSIS as our independent variables. As our prediction algorithm, we used gradient boosting, a weak-learning algorithm suited well for prediction tasks (Bahad and Saxena 2020). In addition to the data from WeSIS, we took data on Gini coefficients from the World income inequality Database Wiid (UNU-WIDER 2025a). However, some pre-processing was necessary to use both datasets as the basis for our prediction.

For our dependent variable we rely on the Wiid. Although several other databases like the World Bank or the World Inequality Database provide Gini coefficients as well, most of them show large fluctuation in the amount of homogeneous datapoints when filtered by type of measurement. Therefore, we chose the Wiid dataset because it provides the longest and most complete homogeneous data series. Since this dataset comprised a multitude of different sources, measuring concepts, and further discriminating variables, we had to filter out about 95% of observations to get an internally consistent time series. For this, we filtered out any resources other than net income. Additionally, we excluded any reference units other than "Person" and sharing units other than "households". Furthermore, we filtered the reference period by years to ensure consistent time series on a country-year basis. Even after these filtering steps, there were still country-years that were duplicated. To resolve this, the dataset was filtered by source frequency. For each country, the most frequent source was selected. For the scale variable in the Wiid dataset, the existing values were prioritized differently. First, the dataset was checked if "per capita" was included in the scale. If it was found, all other values for the duplicated country-years were removed from the dataset. If "per capita" was not found, the dataset was then checked for "No adjustment" and, if not found, for "Equivalized". Despite these efforts, there were still few duplicated country-years left, specifically for the countries New Zealand and Luxembourg. In the case of New Zealand, the same number of values came from different sources. This was due to the OECD providing both historical and updated data series. The new data series was retained. For Luxembourg, there were two different surveys that yield similar results. To remove the duplicated country-years, we calculated the mean of the duplicate values. Through this process, we created a training dataset without duplicated country-years and with only a single calculated mean value. The dataset with original Gini values available for training ranged from the year 1990 to 2020, incorporating 904 individual values in total.

Subsequently, we prepared the independent variables. For this, we relied on the vast number of indicators in WeSIS, which generally reflects the causal complexes laid out in Chapter 3. As pre-processing, some indicators had to be interpolated. We modified a subset of four indicators not collected annually by running a linear interpolation. Specifically, the share of the population that is Jewish, Christian, and Muslim. These indicators were only collected every five years. Since the differences between existing data point were rather small, linear interpolation is justified. The fourth interpolated indicator was the share of the population earning less than the median income. In contrast to the other interpolated indicators, the differences between data points varied across countries. The differences were still sufficiently small to allow for linear interpolation. Moreover, the relevance of this indicator provides an additional reason for applying linear interpolation. Further classifications demonstrated that the interpolated indicator increases the precision of our models. Finally, we filtered country–year combinations so that only those combinations that contained at least two of the three most influential indicators remained. As a result, our dataset consists of 1412 independent variables. The distribution of the most dominant indicators including their importance per cause is documented in Table 1.

With this dataset, we trained our gradient boosting machine (GBM). A GBM is a machine learning algorithm based on decision trees that is used for prediction. Each decision tree predicts the residuals of the previous tree. With this method, it iteratively reduces the prediction error through gradient boosting. Gradient boosting proceeds through the following steps: First, predictions are initialized using a simple decision tree. Next, the residuals are calculated. The residuals are defined as the differences between the actual values and the current predictions. A new decision tree is then fitted to these residuals using all available independent variables. Next, the original predictions are updated with the new prediction multiplied by learning rate. This process is repeated for a predetermined number of iterations. In addition to the number of trees, the accuracy of the prediction depends on the tree depth, the minimum number of observations in the terminal nodes and the shrinkage/learning rate (Ayyadevara 2018).

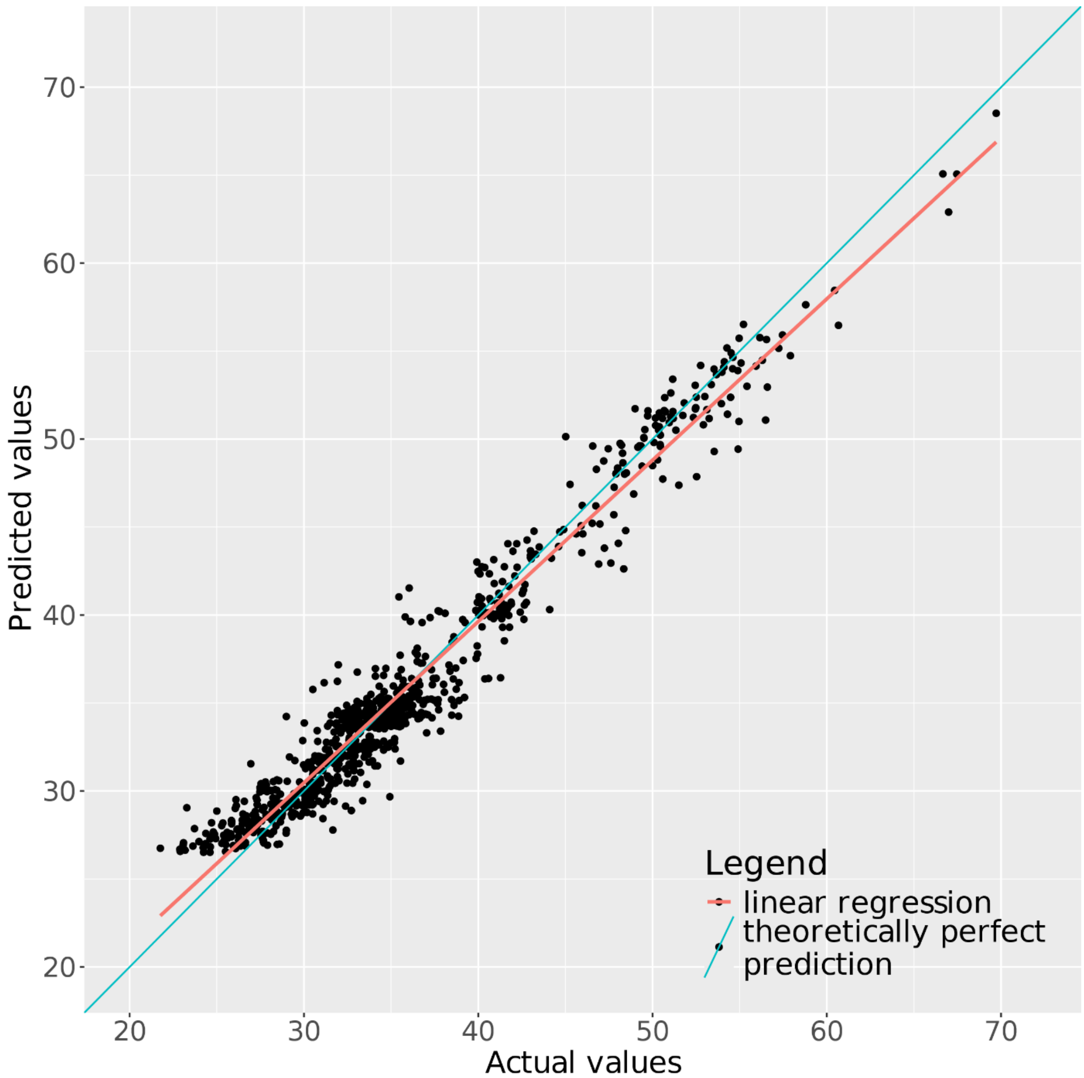
We chose this model because it has been shown to have great predictive capabilities (Bahad and Saxena 2020). In particular, its strengths lie in the stepwise optimization that iteratively reduces the residuals and its ability to combine many weak learners into a strong model (Bentéjac et al. 2021). In real-world applications, GBMs have shown high performance and reliability. They also outperformed Support Vector Machines (SVM) and improved predictive performance (Ray et al. 2023; Touzani et al. 2018).

5. Results

The Wiid Gini data have shown a very unequal distribution, with 731 country–year combinations for OECD countries and 195 for non-OECD countries. This amounts to a total of 924 existing data points, or 22.1% of the possible 4,100 country-year combinations, meaning that the prediction closes significant data gaps. Because of the weak representation of non-OECD countries, we ran several classifications using different weights of 2, 3, 4, 5, and 10 for non-OECD countries. Based on the analysis of the different results, we decided to use a weighting of 5 for non-OECD countries and unweighted predictions for OECD countries for our final dataset. The final training dataset consisted of data for 926 unique country-years.

With the original dataset of 926 unique cases, 859 of which between 1995 and 2019, we predictedGini coefficients for 2913 additional country-years. This resulted in a new set of Gini values containing 3772 data points. Our prediction performed very well with a root mean sqare error (RSME) of only 2.619315 (non-weighted GBM). The median of the residuals was -0.0146 indicating that the model was able to estimate the actual values without systematic bias. The interquartile range was relatively narrow (Q1 = -1.3108, Q3 = 1.1512), indicating that most prediction errors were small. Despite a few higher residuals (minimum -7.1377, maximum 9.3925), the model was overall capable of generating highly accurate predictions. This is further highlighted by the very small standard error (0.0122). Moreover, the p-value was $< 2 \cdot 10^{-16}$, allowing the null hypothesis of no relationship to be rejected with a very high degree of confidence. The multiple R^2 value of 0.8914 shows that approximately 89% of the variance in the dependent variable was explained by the model. Overall, the linear regression analysis revealed a very strong and highly statistically significant relationship between the observed and the predicted Gini values.

Figure 1. Predicted vs. original Gini values



Taking a more thorough look at our independent variables, they resemble our discussion of causes for income inequality in Chapter 3. Table 1 shows the twenty most important features for our prediction. Their importance is either split among or sorted into one of the causes discussed in chapter three. The importance can be read as the percentage of variance explained through variance in this indicator. The most important features were the Gender Inequality Index, the proportion of people living below 50% of the median income (interpolated), the total population ages 65 and above, the capital city coordinates (latitude) and the codetermination and information/consultation of workers.

The importance of the indicators Gender Inequality Index and Human Development Index cuts across the causal complexes discussed in Chapter 3. This is due to the fact that the indicators themselves consist of three dimensions. Specifically, the Gender Inequality Index is composed by the components (a) labour market (economical), (b) empowerment (political) as well as (c) reproductive health (social), while the Human Development Index is based on the three dimensions (i) enjoying a reasonable standard of living (economical), (ii) acquiring knowledge (political) and (iii) living a long and healthy live (social).

All in all, the distribution among the three set of causes is very balanced, highlighting the importance of each one. The low degree of importance gathered among the "other" category underlines the theoretical assumptions and expectations explained in Chapter 3 as well.

Table 1. Feature-importance split along causes for unweighted prediction

Indicator	Economic	Political-Institutional	Social	Other
Proportion of people living below 50% of median income (interpolated)	19,77			
Gender Inequality Index	6,20	6,20	6,20	
Total population ages 65 and above			17,92	
Total ODA for water supply and sanitation, by recipient countries		7,23		
Capital city coordinates (Latitude)				5,59
International financial flow in support of clean energy research		2,97		
Human Development Index	0,99	0,99	0,99	
Victims of intentional homicide per 100,000 population (bothsex)				2,65
Codetermination and information/consultation of workers		1,86		
Codetermination board membership		1,58		
Unemployment Insurance Law. Duration of Benefits		1,54		
Maternal mortality ratio			1,41	
State authority over territory				1,04
Poverty gap at dollar 3.65 a day (2017 PPP)	0,95			
De facto coverage of children by the child benefit for citizens/residents		0,93		
Share of population ages 0 to 14			0,79	
School life expectancy, primary, male (UIS)		0,77		
Behavioral conditions in the child benefit system		0,69		
Out-of-pocket expenditure per capita in US$ (WHO)		0,69		
Domestic autonomy				0,67
Sum	27,91	25,45	27,31	9,96
	90.63			

6. Discussion

Despite the frequent use and importance of the Gini index in explaining income inequalities, a closer look at the "usual" resources revealed large data gaps. Our novel approach, taking advantage of the large amount of data in WeSIS, produced an internally consistent and encompassing Gini coefficient dataset. Though produced through prediction, our dataset fills gaps for country-years for which no data was available before. Figure 2a to 2c show the advancement of our predicted data compared to the original data series our analysis was based on. The original Gini data points are printed in blue and our predicted Gini values in red. The graphs highlight the tremendous improvement over the former data availability. Given that the gradient boosting machine reported high confidence in our results, this new dataset of Gini coefficients opens up formerly impossible lines of research while decreasing data biases between the Global North and the Global South.

Obviously, we cannot prove the correctness of our new data points directly – in ideal world there would have been data, and hence no need for imputations. The match between the expectations rooted in the literature on the causes of income inequality and the results our prediction produced still offer face and "construct validity" in Adcock and Collier's (2001) terms. In sum, we are confident that our Gini data – based on a "theory-connected data-driven approach" – is useful for future research.

The new data can be found here: https://wesis.org/indicators/2588.

Figure 2a: Country-timelines original (blue) and predicted (red) Gini values

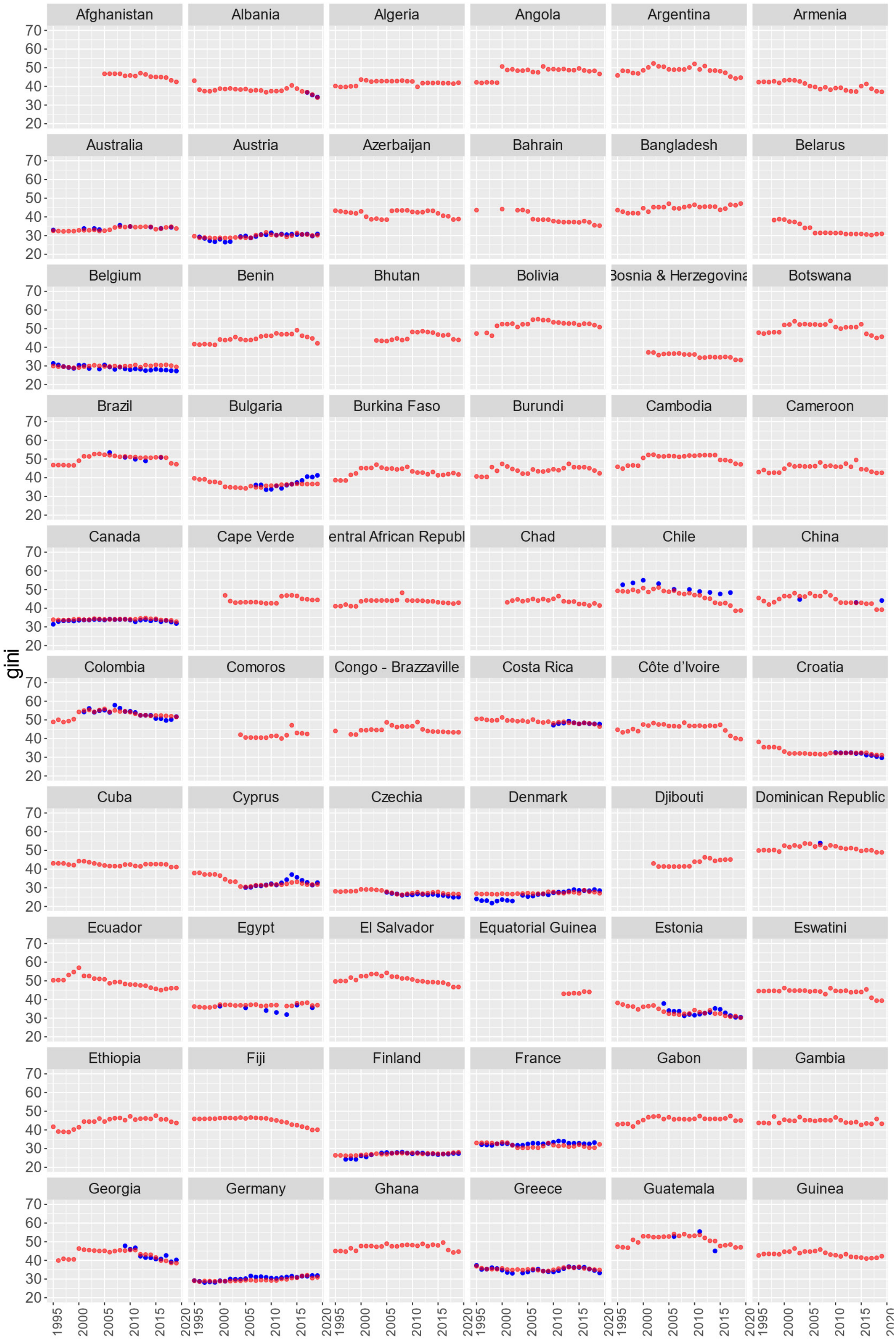


Figure 2b. Country-timelines original (blue) and predicted (red) Gini values

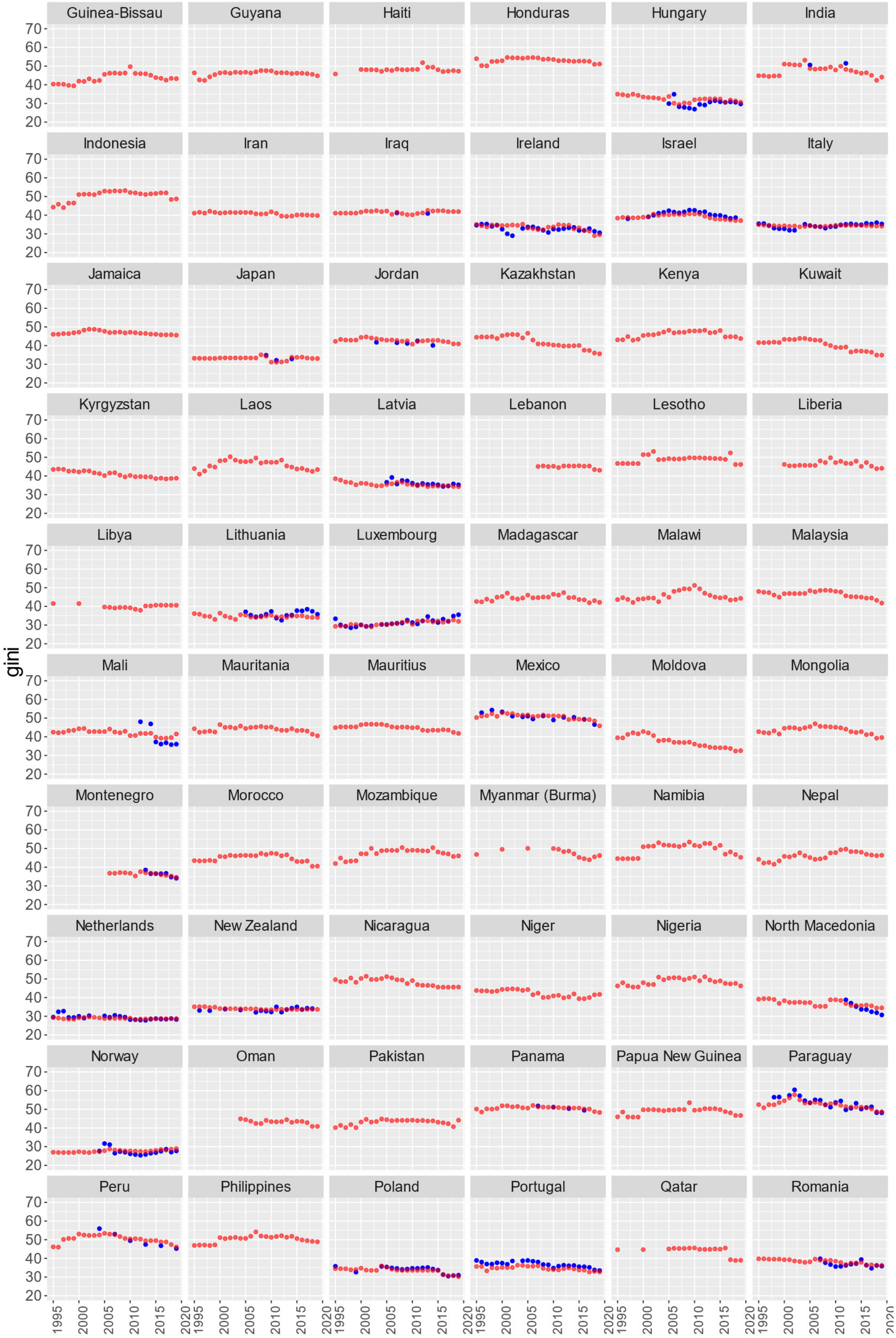
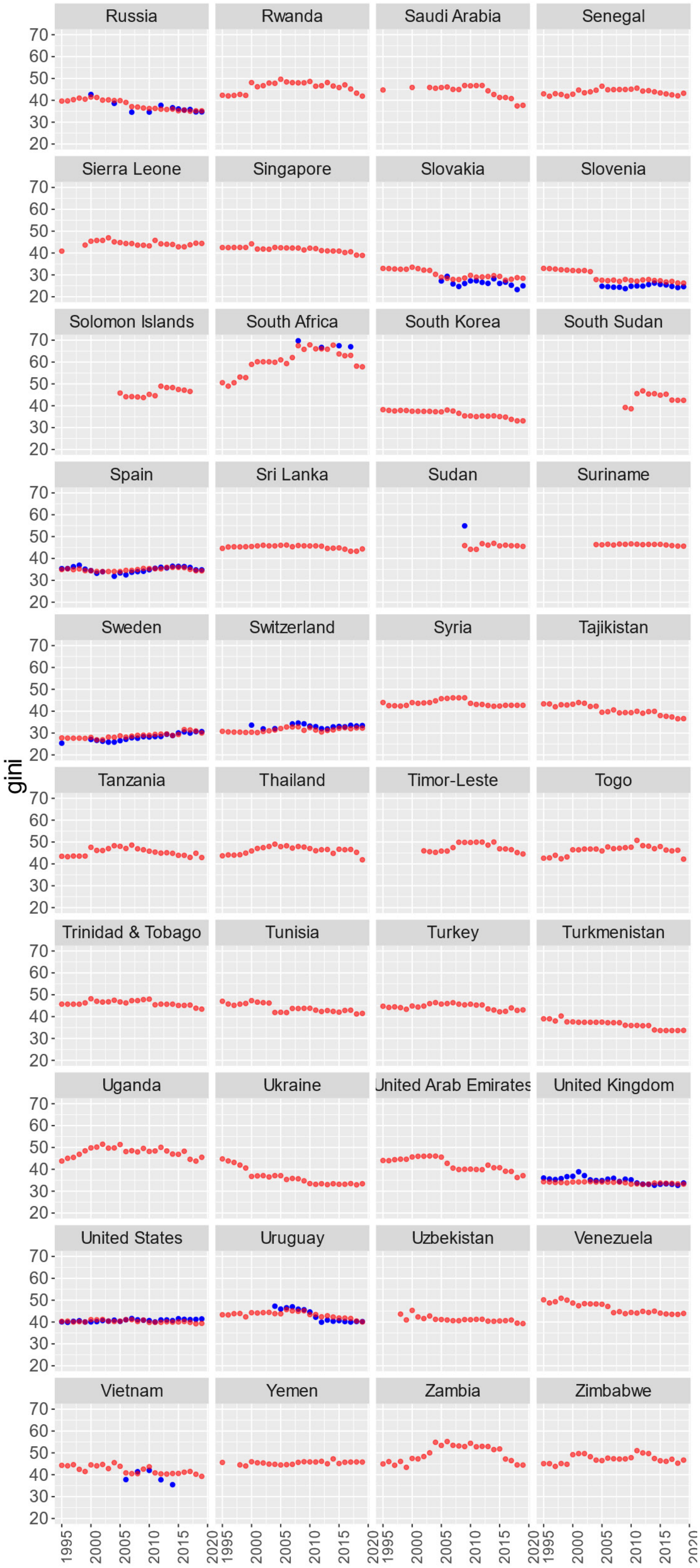


Figure 2c. Country-timelines original (blue) and predicted (red) Gini values



References

Adcock, Robert; Collier, David (2001): Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. In: American Political Science Review 95 (3), 529 – 546. https://doi.org/10.1017/S0003055401003100

Atkinson, Anthony B. (1970): On the measurement of inequality. In: Journal of Economic Theory 2 (3), 244-263. https://doi.org/10.1016/0022-0531(70)90039-6

Ayaita, Adam (2023): Is There an Ethnic Pay Gap in Germany? Evidence from a Representative Sample of the Adult Population, ZBW – Leibniz Information Centre for Economics: Kiel, Hamburg. https://hdl.handle.net/10419/267865

Ayyadevara, V. Kishore (2018): Gradient Boosting Machine. In: V. Kishore Ayyadevara (eds.): Pro Machine Learning Algorithms. Berkeley, CA: Apress, 117–134. https://doi.org/10.1007/978-1-4842-3564-5_6

Bahad, Pritika; Saxena, Preeti (2020): Study of AdaBoost and Gradient Boosting Algorithms for Predictive Analytics. In: Geetam Singh Tomar, Narendra S. Chaudhari, Jorge Luis V. Barbosa, Mahesh Kumar Aghwariya (eds.): International Conference on Intelligent Computing and Smart Communication 2019. Singapore: Springer Singapore (Algorithms for Intelligent Systems), pp. 235–244. https://link.springer.com/chapter/10.1007/978-981-15-0633-8_22

Bandelj, Nina; Mahutga, Matthew C. (2010): How Socio-Economic Change Shapes Income Inequality in Post-Socialist Europe. In: Social Forces, 88 (5), 2133-2161. https://doi.org/10.1353/sof.2010.0042

Bentéjac, Candice; Csörgő, Anna; Martínez-Muñoz, Gonzalo (2021): A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review 54, 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Blesch, Kristin; Hauser, Oliver P.; Jachimowicz, Jon M. (2022): Measuring inequality beyond the Gini coefficient may clarify conflicting findings. In: Nature Human Behaviour 6, 1525-1536. https://doi.org/10.1038/s41562-022-01430-7

Bodea, Cristina; Houle, Christian; Kim, Hyunwoo (2021): Do financial crises increase income inequality? In: World Development, 147, 1-13. https://doi.org/10.1016/j.worlddev.2021.105635

Brynin, Malcom; Güveli, Ayse (2012): Understanding the ethnic pay gap in Britain. Work, In: Employment and Society, 26(4), 574-587. https://doi.org/10.1177/0950017012445095

Cammeraat, Emile (2020): The relationship between different social expenditure schemes and poverty, inequality and economic growth. In: International Social Security Review 73 (2), 101-123. https://doi.org/10.1111/issr.12236

Cazzaniga, Mauro; Jaumotte, Florence; Li, Longji; Melina, Giovanni; Panton, Augustus J.; Pizzinelli, Carlo; Rockall, Emma; Tavares, Marina M. (2024): Gen-AI: Artificial Intelligence and the Future of Work. IMF Staff Discussion Note SDN2024/001, Washington: International Monetary Fund. https://doi.org/10.5089/9798400262548.006

Chong, Alberto; Gradstein, Mark (2007): Inequality and informality. In: Journal of Public Economics 91 (1-2), 159-179. https://doi.org/10.1016/j.jpubeco.2006.08.001

Conceicao, Pedro; Ferreira, Pedro M. (2000): The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications. UTIP Working Paper No. 14. https://doi.org/10.2139/ssrn.228703

Cysne, Rubens Penha (2009): On the positive correlation between income inequality and unemployment. In: The Review of Economics and Statistics, 91 (1), 218-226. https://www.jstor.org/stable/25651330

Dabla-Norris, Era; Kochhar, Kalpana; Suphaphiphat, Nujin; Ricka, Franto; Tsounta, Evridiki (2015): Causes and Consequences of Income Inequality: A Global Perspective. IMF Staff Discussion notes. https://www.imf.org/external/pubs/ft/sdn/2015/sdn1513.pdf

De Maio, Fernando (2007): Income inequality measures. In: Journal of Epidemology and Community Health, 61 (10), 849-852. https://www.jstor.org/stable/40665665

Dolls, Mathias; Doorley, Karina; Paulus, Alari; Schneider, Hilmar and Sommer, Eric (2019): Demographic change and the European income distribution. In: The Journal of Economic Inequality, 17, 337-357. https://doi.org/10.1007/s10888-019-09411-z

Dorn, Florian; Fuest, Clemens; Potrafke, Niklas (2022): Trade openness and income inequality: New empirical evidence. In: Economic Inquiry 60 (1), 202-223. https://doi.org/10.1111/ecin.13018

Duncan, Denvil; Sabrinova Peter, Klara (2016): Unequal inequalities: Do progressive taxes reduce income inequality? In: International Tax and Public Finance, 23, 762–783. https://doi.org/10.1007/s10797-016-9412-5

Eichengreen, Barry; Csonto, Balazs; El-Ganainy, Asmaa and Koczan, Zsoka (2021): Financial Globalization and Inequality: Capital Flows as a Two-Edged Sword. In: IMF Working Paper 21 (4), 1-37. https://doi.org/10.2139/ssrn.3772791

Esseau-Thomas, Chrys; Galarraga, Omar; Khalifa, Sherif (2022): Epidemics, pandemics and income inequality. In: Health Economics Review, 12 (7), 1-15. https://doi.org/10.1186/s13561-022-00355-1

Eydam, Ulrich; Qualo, Hannes (2024): Income inequality and taxes – an empirical assessment. In: Applied Economics Letters, 31 (18), 1828-1835, https://doi.org/10.1080/13504851.2023.2208328

Firebaugh, Glenn (2006): The New Geography of Global Income Inequality. Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/9780674036895

Furceri, Davide; Loungani, Prakash; Ostry, Jonathan D.; Pizzuto, Pietro (2020): Will Covid-19 affect inequality? Evidence from past pandemics. In: Covid Economics. Vetted and Real-Time Papers, 12 (1), 138-157. http://reparti.free.fr/furceri20.pdf

Gastwirth, Joseph L. (2017): Is the Gini Index of Inequality Overly Sensitive to Changes in the Middle of the Income Distribution? In: Statistics and Public Policy, 4 (1), 1-11. https://doi.org/10.1080/2330443X.2017.1360813

Gonzales, Christian, Jain-Chandra, Sonali; Kochhar, Kalpana, Newiak, Monique; Zeinullayev, Tlek (2015): Catalyst for Change: Empowering Women and Tackling Income Inequality. IMF Staff Discussion Note SDN/15/20. http://www.imf.org/external/pubs/ft/sdn/2015/sdn1520.pdf

Grusky, David (2018). The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender (2nd ed.). Routledge: New York. https://doi.org/10.4324/9780429494468

Hahs, Jenny, (2022): From Geneva to the World? Global network Diffusion of Antidiscrimination Legislation in Employment and Occupation: The ILO´s C111. In: Michael Windzio, Ivo Mossig, Fabian Besche-Truthe, Helen Seitzer (eds.): Networks and Geographies of Global Social Policy Diffusion. Culture, Economy, and Colonial Legacies. Palgrave Macmillan, Cham, 195-225. https://doi.org/10.1007/978-3-030-83403-6_8.

Hartmann, Dominik; Guevara, Miguel R.; Jara-Figueroa, Cristian; Aristarán, Manuel; Hidalgo, César A. (2017): Linking Economic Complexity, Institutions, and Income Inequality. In: World Development 93, 75-93. https://doi.org/10.1016/j.worlddev.2016.12.020

Henderson, David R.; McNab, Robert M.; Rozsas, Tamas (2008): Did Inequality Increase in Transition? An Analysis of Transition Countries of Eastern Europe and Central Asia. In: Eastern European Economies, 46 (2), 28-49. https://www.jstor.org/stable/27740058

Huber, Evelyne; Stephens, John D. (2014): Income inequality and redistribution in post-industrial democracies: demographic, economic and political determinants. In: Socio-Economic Review 12 (2), 245–267. https://doi.org/10.1093/ser/mwu001

Kum, Hakan (2024): The Relationship Between Informal Economy and Income Inequality: An Econometric Analysis for BRICS Countries. In: International Journal of Economics and Financial Issues 14 (1), 117-125. https://doi.org/10.32479/ijefi.15664

Kvist, Jon (2004): Does EU Enlargement Start a Race to the Bottom? Strategic Interaction among EU Member States in Social Policy. In: Journal of European Social Policy 14(3): 301-318. https://doi.org/10.1177/0958928704044625

Mossig, Ivo; Düpont, Nils (2020): Wirtschaftsgeographische Perspektiven auf Entwicklungsdynamiken von Sozialpolitik. In: Geographische Zeitschrift 108 (4), 224 – 248. https://doi.org/10.25162/GZ-2020-0004

Mossig, Ivo; Lischka, Michael (2022): Globalization, Economic Interdependencies and Economic Crises. In: Delia González de Reufels, Frank Nullmeier, Herbert Obinger (eds.): International Impacts on Social Policy. Short Histories in a Global Perspective. Cham: Palgrave Macmillan, pp. 289-304. https://doi.org/10.1007/978-3-030-86645-7_23

Mossig, Ivo; Obinger, Herbert (eds.) (2023): Mapping Global Dynamics of Social Policy. Socium, SFB 1342: Bremen. https://doi.org/10.26092/elib/2559

Neves Costa, Rita; Pérez-Duarte, Sébastien (2019): Not all inequality measures were created equal. The measurement of wealth inequality, its decompositions, and an application to European household wealth. In: Statistics Paper Series, 31, 1-58. https://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sps31~269c917f9f.en.pdf

Obst, Thomas (2014): Income Inequality and the Welfare State: – How redistributive is the public sector? AV Akademikerverlag: Saarbrücken.

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

Osterman, Paul (1975): An Empirical Study of Labor Market Segmentation. In: ILR Review 28 (4), 508-523. https://doi.org/10.2307/2521648

Pfeffer, Fabian T.; Danziger, Sheldon; Schoeni, Robert F. (2013): Wealth Disparities before and after the Great Recession. In: The Annals of the American Academy of Political and Social Science, 650 (1), 98–123. https://www.jstor.org/stable/24541678

Polacko, Matthew (2021): Causes and Consequences of Income Inequality – An Overview. In: Statistics, Politics and Policy 12 (2), 341-357. https://doi.org/10.1515/spp-2021-0017

Ray, Sourav; Rahman, Md Masnun; Haque, Mohaiminul; Hasan, M. Washif; Alam,M. Manjurul (2023): Performance evaluation of SVM and GBM in predicting compressive and splitting tensile strength of concrete prepared with ceramic waste and nylon fiber, Journal of King Saud University - Engineering Sciences 35 (2), 92-100. https://doi.org/10.1016/j.jksues.2021.02.009 .

Seitzer, Helen (2022): The Diffusion of Workplace Antidiscrimination Regulations for LGBTQ+ Community. In: Michael Windzio, Ivo Mossig, Fabian Besche-Truthe, Helen Seitzer (eds.): Networks and Geographies of Global Social Policy Diffusion. Culture, Economy, and Colonial Legacies. Palgrave Macmillan, Cham, 227-253. https://doi.org/10.1007/978-3-030-83403-6_9.

SenGupta, Swapnanil; Atal, Aakansha (2024): Income inequality in the face of climate change: an empirical investigation on unequal nations, vulnerable regions and India. In: SN Business & Economics, 4 (87). 1-33, https://doi.org/10.1007/s43546-024-00685-8

Touzani, Samir; Granderson, Jessica; Fernandes, Samuel (2018): Gradient boosting machine for modeling the energy consumption of commercial buildings. In: Energy and Buildings 158 (1), 1533-1543. https://doi.org/10.1016/j.enbuild.2017.11.039

UNU-WIDER (2025a): World Income Inequality Database (WIID). Version 29 April 2025. https://doi.org/10.35188/UNU-WIDER/WIID-290425

UNU-WIDER (2025b): World Income Inequality Database (WIID) Companion dataset (wiidcountry and/or wiidglobal). Version 29 April 2025. https://doi.org/10.35188/UNU-WIDER/WIIDcomp-290425

von Weizsäcker, Robert K. (1995): Public pension reform, demographics, and inequality. In: Journal of Population Economics 8, 205–221 (1995). https://doi.org/10.1007/BF00166652

Wagmiller, Robert L.; Schultz Lee, Kristen; Houston Su, Jessica (2020): The role of welfare in family income inequality: 1968–2016. In: Children and Youth Services Review, 119, 105615. https://doi.org/10.1016/j.childyouth.2020.105615.

Windzio, Michael; Mossig, Ivo; Besche-Truthe, Fabian, Seitzer, Helen (2022): Networks of Global Social Policy Diffusion: The Effects of Culture, Economy, Colonial Legacies, and Geographic Proximity. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-83403-6

Yang, Juan and Gao, Man (2018): The impact of education expansion on wage inequality. In: Applied Economics 50 (12), 1309-1323. https://doi.org/10.1080/00036846.2017.1361008

Yitzhaki, Shlomo; Schechtman, Edna (2013): The Gini Methodology. A Primer on a Statistical Methodology. Springer: New York, Heidelberg, Dordrecht London. https://doi.org/10.1007/978-1-4614-4720-7