Multidimensional Inequality: An Empirical Analysis of its Social and Economic Determinants and Implications

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von

Laura Radatz

Tübingen
2023

Tag der mündlichen Prüfung:	26.07.2023
Dekan:	Prof. Dr. Ansgar Thiel
1. Gutachter:	Prof. Dr. Jörg Baten
2. Gutachter:	Prof. Dr. Manuel Llorca-Jaña

Acknowledgements

During the time I have written this dissertation, I have received support from many people to whom I am very grateful.

First of all, I would like to thank my supervisor, Prof. Dr. Joerg Baten, for giving me the opportunity to pursue my PhD, and for his valuable feedback and constant support with his positive attitude. He also encouraged and supported me to present my research at international conferences and workshops and to submit my work to academic journals. This provided me with the opportunity to present my research to an international audience and helped me to gain valuable experience and insight into academia and research.

I am also very grateful to my co-examiner, Prof. Dr. Manuel Llorca-Jaña. Thank you for your willingness to examine and evaluate this thesis. I also greatly appreciate the willingness of Prof. Dr. Markus Pudelko to chair the disputation.

I also wish to thank my co-authors Prof. Dr. Alberto Batinti and Prof. Dr. Joan Costa-Font for your engagement, creative ideas and our kind cooperation, as well as for your help and support at all times.

We have a wonderful, international, fun and supportive research group. I would like to thank you all for the academic and moral support, but also for the coffee, the table tennis tournaments in the sun, the tricky games of hangman, the beanbags for much-needed breaks, and the delicious cakes and sweets from all over the world. Special thanks go to Dr. Elisabeth Kempter, Sarah Ferber and Sophia Jung, for our discussions and your inputs, and for reviewing parts of the thesis and providing me with critical feedback and helpful comments. I would also like to thank Martina Mayer for her constant administrative support.

I am deeply grateful to my family, parents, sisters and friends, and to Marco Dell'Oso, for their constant moral support and encouragement, for proofreading—a special thanks to Niroja Sivanesan and Rebecca Radatz—and for believing in me.

Laura Radatz
Tübingen, April 2023

Contents

A Introduction **1**
A.1 Inequality – drivers and implications . . . 1
A.2 On the measurement of inequality: dimensions and anthropological indicators . . . 3
A.3 Aim and outline of the dissertation . . . 6
A.4 References . . . 10

B Measuring Multidimensional Inequality and Conflict in Africa and in a Global Comparison **16**
B.1 Introduction . . . 17
B.2 Measuring economic inequality . . . 20
B.2.1 Health inequality . . . 21
B.2.2 Land inequality . . . 25
B.2.3 Calculation of the multidimensional joint inequality index . . . 26
B.2.4 Cross-Validation of the new measure: how our anthropological inequality measure correlates with income measurements . . . 28
B.2.5 Conflict data . . . 28
B.2.6 Control variables . . . 30
B.3 Does inequality fuel violent conflict? . . . 31
B.3.1 Regression results . . . 31
B.3.2 Robustness check . . . 33
B.3.3 Instrumental variable regression . . . 34
B.4 Conclusion . . . 36
B.5 References . . . 37
B.6 Tables and figures . . . 41
B.7 Appendix . . . 56
B.7.1 Variable definitions . . . 56
B.7.2 The construction of multidimensional inequality . . . 57
B.7.3 Health inequality . . . 57
B.7.4 Land inequality . . . 57
B.7.5 Joint inequality index . . . 58
B.7.6 References . . . 60
B.7.7 Tables and figures . . . 61

C Social and Economic Disparities and the 1918 Influenza Pandemic: Lessons for Today **76**
C.1 Introduction . . . 77
C.2 Mechanism . . . 80
C.3 Literature review: empirical evidence from previous research . . . 83

C.4 Data and method . . . 86
C.4.1 Measuring the severity of a pandemic . . . 86
C.4.2 Measuring inequality in income and health . . . 87
C.4.3 Methodology . . . 89
C.5 Descriptive statistics and regression results . . . 90
C.5.1 Descriptive statistics . . . 90
C.5.2 Regression results . . . 91
C.6 Discussion and implications for today . . . 92
C.7 Conclusion . . . 94
C.8 References . . . 95
C.9 Tables and figures . . . 100
C.10 Appendix . . . 108
C.10.1 Variable description and sources . . . 108
C.10.2 Country coverage . . . 109
C.10.3 References . . . 110

D The Heights of Medical Care: Health Insurance and Inequality in Adult Stature 112
D.1 Introduction . . . 113
D.2 Health insurance, inequality and heights . . . 115
D.3 Data and methods . . . 117
D.3.1 Estimating inequality using the Gini coefficient of height . . . 117
D.3.2 Health insurance . . . 119
D.4 Descriptive statistics and results . . . 122
D.4.1 Descriptive statistics . . . 122
D.4.2 Regression analysis . . . 124
D.5 Instrumental variable estimates . . . 125
D.6 Robustness check . . . 126
D.7 Conclusion . . . 128
D.8 References . . . 129
D.9 Tables and figures . . . 133
D.10 Appendix . . . 142
D.10.1 Height inequality . . . 142
D.10.2 Universal Health Coverage: measurement and sources . . . 142
D.10.3 Controls . . . 143
D.10.4 References . . . 144
D.10.5 Tables and figures . . . 145

E Summary and Outlook 148
E.1 References . . . 152

List of Figures

Figure B.1. Relationship between height Gini and income Gini . . . 41
Figure B.2. Distribution of our joint inequality index by the percentage of world population covered . . . 42
Figure B.3. Comparison of different inequality measures . . . 43
Figure B.4. Joint inequality index, most recent data available . . . 43
Figure B.5. Inequality by world region . . . 44
Figure B.6. Number of civil war outbreaks over time and by world region . . . 44
Figure B.7. Civil war outbreaks worldwide for the decade 2000 (2000–2009) . . . 45
Figure B.8. Civil war outbreaks by type . . . 45
Figure B.9. Data selection . . . 46
Figure C.1. Severity of the influenza pandemic 1918 . . . 100
Figure C.2. Differences in inequality between the (birth) decade 1910 and 1920 by income group . . . 101
Figure C.3. Scatterplot between inequality and excess flu mortality rate . . . 102
Figure D.1. Height inequality worldwide: 1810–1900 . . . 133
Figure D.2. Height inequality worldwide: 1960–2000 . . . 133
Figure D.3. Development of height inequality by world region . . . 134
Figure D.4. Differences in height inequality by Universal Health Coverage . . . 135
Figure D.5. Percentage of different levels of height inequality by the (non-)presence of Universal Health Coverage . . . 136

List of Tables

Table B.1. Comparison between income Gini and height Gini: regressions based on selected different datasets . . . 47
Table B.2. Correlation between income Ginis and top 1% income share from WID on a yearly basis . . . 47
Table B.3. Overview of our joint inequality index – time and region coverage as a percentage of the world region's population . . . 48
Table B.4. Correlation between our joint inequality index and income Ginis from different sources . . . 49
Table B.5. Summary statistics . . . 49
Table B.6. Regression of civil war onset . . . 50
Table B.7. OLS regression of civil war onset . . . 51
Table B.8. OLS regression with region and country fixed effects . . . 52
Table B.9. OLS regression by different civil war types and measures . . . 53
Table B.10. Robustness check: regression for low- and high-income countries . . . 54
Table B.11. Instrumental variable regression . . . 55
Table B.12. Additional sources of height Gini . . . 61
Table B.13. Sources for income Gini coefficients . . . 64
Table B.14. The average effect of a land reform . . . 73
Table B.15. Sensitivity test for the construction of the joint inequality index . . . 74
Table C.1. Summary statistics . . . 103
Table C.2. First-difference estimation . . . 104
Table C.3. First-difference estimation including controls . . . 105
Table C.4. First-difference estimation by income group . . . 106
Table C.5. Robustness checks . . . 107
Table D.1. Summary statistics . . . 137
Table D.2. Correlation between height inequality and Universal Health Coverage . . . 137
Table D.3. OLS regression including controls: correlates of height inequality . . . 138
Table D.4. Instrumental variable approach: determinants of height inequality . . . 139
Table D.5. Correlation between height inequality and the distance to the Soviet Union . . . 140
Table D.6. Robustness check for height inequality and healthcare measurements . . . 141
Table D.7. Sources for health insurance legislation . . . 145

A Introduction

A.1 Inequality – drivers and implications

The unequal distribution of resources and public goods, not only seems unfair, but also has concrete negative economic, social, and political consequences for individuals and the state as a whole. At an individual level, this could include reduced access to education and medical services for the poorer members of society, and at a macroeconomic level, hindered economic growth and greater political instability (United Nations Department of Economic and Social Affairs [UN-DESA], 2020).

Although we observe a recent decline in income inequality between countries, disparities within countries have been rising in almost all regions of the world (e.g., Bourguignon, 2018; Chancel and Piketty, 2021; Lakner and Milanovic, 2016; Milanovic, 2016; Ravallion, 2018). Historically, Latin America stands out as a region with high inequality countries, such as Brazil[1], followed by sub-Saharan Africa (e.g., Piketty, 2013). In the latter, South Africa and Namibia are known for being some of the most unequal countries in the world (World Bank, 2022), as well as the United States (US), which have shown increasing discrepancies in income since the 1980s (Piketty et al., 2018). Globally, high inequalities are a major concern, especially since the COVID-19 pandemic. As a result, there has been a resurgence of interest in understanding inequalities, both in the literature and the public, about how inequality affects our societies—and what can be done about it.

Scholars worldwide have extensively, and often controversially, discussed the causes and implications of inequality. High levels of inequality are of concern in many ways: people belonging to a disadvantaged group or minority, or coming from a low-income background, have lower educational attainment (e.g., Heckman, 2011; Narayan et al., 2018), which leads to poorer adult outcomes, such as limited intergenerational mobility and lower earnings (e.g., Breen and Müller, 2020; Chetty

[1] In recent years, from 2000 to 2010, we could observe a slight reduction in inequality in Latin American countries (Chancel, 2022; Clifton et al., 2020), even in Brazil (Bourguignon, 2018). However, since 2013, there are some signs that the Gini coefficient is increasing again (UN-DESA, 2020).

et al., 2014; Corak, 2013; Heckman, 2011). Disparities in socioeconomic status also show a negative (although not direct) relationship with individual health outcomes, such as life expectancy or mortality rates (even more so during a health crisis such as the COVID-19 pandemic) (e.g., Currie, 2011; Deaton, 2002, 2003; Kawachi et al., 1997; Singu et al., 2020). In these high-inequality societies, it is not just the individual who is at higher risk: the health and well-being of the whole population could be improved by promoting equality (Pickett & Wilkinson, 2015). At the macroeconomic level, it is further assumed that high inequalities hamper poverty reduction efforts (e.g., Banerjee and Duflo, 2011; Bergstrom, 2022; Besley and Burgess, 2003; Ferreira et al., 2008; Fosu, 2017; Ravallion, 2005), and undermine social cohesion, especially trust (Alesina et al., 2016). Moreover, inequality is seen as a threat to social and political stability (Piketty, 2013), as it is positively associated with homicide rates and crime in general (e.g., Choe, 2008; Kelly, 2000; Neumayer, 2005; Nivette, 2011) and could fuel protests and violent conflict (e.g., Acemoglu and Robinson, 2006; Jensen and Sørensen, 2012; Vergolini, 2011). In the well-studied relationship with economic development, rising inequality is mainly associated with slower economic growth in the subsequent period (Banerjee & Duflo, 2003; Castelló-Climent, 2010; Deininger & Squire, 1998; Ostry et al., 2014; Stiglitz, 2012).[2]

Most of those negative implications described above, however, could be hindered by decreased levels of inequality. Appropriate (national) policy interventions have the potential to significantly reduce even persistent inequalities (Ravallion, 2018; UN-DESA, 2020). Piketty (2013) even argues that governments determine the level of inequality themselves, as they would be able to combat it through redistribution and social programmes, such as education, public health, and employment policies (UN-DESA, 2020). However, this requires a thorough understanding of inequality and its drivers, and how to tackle it properly.

But what causes these disparities in the first place? Possible explanations in the literature for the different levels of inequality across countries include the quality of political institutions, as famously stated in the book by Robinson and Acemoglu (2012). They argue that the ability of the

[2]These findings remain controversial, mainly because of different ways of measuring and conceptualizing inequality, and the usage of different data sources, countries and time periods. For example, Barro (2000) finds no evidence of a relationship between overall inequality and economic growth. Instead, he states that the relationship depends on a country's income level.

government and the willingness of the party in power to implement specific welfare reforms, the degree of democratisation and the level of corruption are the main factors influencing a country's inequality level (see also Furceri and Ostry, 2019; Gupta et al., 2002; Iversen and Soskice, 2006; Rodrik et al., 2004). In addition to the political system, Roser and Cuaresma (2016) focus on technological progress and international trade as important drivers of inequality. Globalisation and the introduction of new technologies created a shift from agriculture to the manufacturing and service sector and increased demand for high-skilled workers and lower demand for low-skilled workers. When the supply on the labour market for high-skilled workers is low, wages increase, which in turn leads to an increase in inequality between high- and low-skilled workers (e.g., Acemoglu, 2003; Krugman, 2000; Kuznets, 1955; and see review in Helpman, 2016). Other explanations include a country's culture and its geographical location (e.g., Sachs, 2006; Steckermeier and Delhey, 2019) and ethnic, religious or language fractionalisation (Alesina et al., 2003; Casey & Owen, 2014). Unequal access to and the quality of the healthcare and school systems are seen as key causes of inequality (United Nations Economic and Social Commission for Asia and the Pacific et al. [UN-ESCAP], 2018). In their cross-country study, de Gregorio and Lee (2002) analyse the variation of educational attainment in a country and its relation to income inequality. They conclude that education is one main determinant of the distribution of income (see also Sylwester, 2003). Finkelstein et al. (2012) document evidence that public health insurance expansion does improve self-reported health. Finally, climate change could exacerbate inequalities between and within countries: rising temperatures and extreme weather events such as drought and famine disproportionately affect people at the bottom of the income distribution (Chancel, 2022; UN-DESA, 2020).

A.2 On the measurement of inequality: dimensions and anthropological indicators

When we talk about inequality in an economic context, we mainly refer to the concept of income inequality or wealth inequality.[3] Both measures are widely used as proxies for well-being, often measured as the Gini coefficient, and are publicly available from various databases.[4] During the

[3] Income refers to the flow of monetary benefits from work or an asset, whereas wealth describes a stock of values.

[4] Databases include those of the Organisation for Economic Co-operation and Development (OECD), the World Bank, the World Income Inequality Database (WIID), the World Inequality Database (WID), and the United Nations (UN), but individual estimates are also widely used in empirical analyses.

last decade, country and time coverage has improved considerably, allowing greater comparability over time and between developed and developing countries. Nevertheless, consistent data exists mainly for developed countries and the post-1960 period. The lack of data available for developing countries and historical data for long-run analysis may explain why many empirical studies tend to find different results in the inequality nexus, either in the direction of causality or in the level of significance.

Dimensions. Other reasons for conflicting empirical findings could be the use of different concepts of inequality. Inequality has many dimensions beyond income and wealth: disparities between countries and people can be manifested and driven by different aspects such as belonging to a particular group, class or minority, age, gender or underlying living conditions (e.g., Sen, 2005; Stiglitz et al., 2009). Stiglitz et al. (2009) therefore argue that we need to take further dimensions of inequality into account to measure and address individual's well-being. Amartya Sen's approach also goes beyond measuring the inequality of outcomes. He advocates for looking at a person's opportunities, such as their circumstances, rather than measuring the wealth of an economy (Sen, 2005).[5] Alkire and Foster (2010) differentiate outcomes like income, which are outsides one's control, such as health or living conditions, and other factors that one can influence through talent or effort. They argue that policy intervention should focus on the aspect for which no individual can be held responsible, which would reduce inequality; especially as these 'inequalities of opportunities' influence 'inequalities of outcomes', such as income or wealth (Atkinson, 1980). For example, Banerjee and Duflo (2011) demonstrate that a child's health can make a tremendous difference to later performance and income level: the use of bed nets in Kenya as a child can lead to an increase of later income of 15%. They conclude that unequal opportunities as a child can therefore determine your outcomes during adulthood.

Appropriate policies to reduce inequality, should thus not only target income inequality with appropriate taxes but also include education and health reforms to provide equal opportunities for everyone (Stiglitz et al., 2009). Consequently, the analysis of inequality (and the study of its effects

[5] With his approach to human development, Sen mainly influenced the development of the Human Development Index (HDI), which was developed by Mahbub ul Haq and is now used by the United Nations (see United Nations Development Programme [UNDP], 2023).

and drivers) should not only look at income or wealth as a proxy for well-being. Other dimensions should also be taken into consideration.

Alternative inequality measures. The literature already proposes several alternative measures to account for the multifaceted concept of inequality and some of them further provide us with data for earlier time periods and developing countries. One famous example includes the Human Development Index (HDI), which goes beyond the measurement of economic outcomes. To construct one single indicator it further takes health and education into account (UNDP, 2023). To measure different dimensions of inequality, Rohde and Guest (2018) construct a single indicator using household income, education, health status and leisure time from household surveys to compare multidimensional inequality levels across three developed countries. Neumayer and Plümper (2016) provide a broad approach to income inequality that considers other measurements of human capital like longevity. Ramos and Van de Gaer (2021) discuss different measurements of inequality of opportunity and apply them to European countries. Developing a new measure, Kuhn and Weidmann (2015) study the relationship between inequality and the risk of conflict using night light data on ethnic settlement areas and estimated population data as a proxy for inter-group inequality. They state that it highlights the importance of different dimensions of economic inequality.

The use of anthropological data, such as the distribution of height within a country's population, is a common alternative to measure inequality, especially for historical time periods (see reviews in Ayuda and Puche-Gil, 2014; Blum, 2014, and more recently: Choi, 2020; Llorca-Jaña et al., 2021; Schwekendiek and Baten, 2019). Height is linked to the social and economic circumstances a person experiences during childhood and adolescence as access to food, shelter and healthcare (Pradhan et al., 2003). The distribution of heights thus reflects the inequality level within a country (Baten & Blum, 2011). Another important indicator for structural inequality is land inequality (Easterly, 2007). High levels of land inequality contribute to long-term negative impacts on human capital (Baten & Juif, 2014; Hippe & Baten, 2012; Qasim et al., 2020), as large landowners need unskilled workers and are therefore not interested in educational reforms (Galor et al., 2009).

A.3 Aim and outline of the dissertation

As outlined in the previous section, high levels of inequality can have several negative effects, which have already been studied in detail. However, long-term evidence, particularly for developing countries, is still lacking. As policy interventions could help to reduce prevailing disparities in income and health, this doctoral dissertation aims to provide a comprehensive long-term view of inequality, its dimensions, drivers, and implications, for both historical and recent developments. This is done by extending existing data on inequality using the anthropometric indicator of height. Using a birth cohort approach, this thesis analyses the evolution of inequality within countries and across decades since the 19^{th} century. For developing countries in particular, we can now provide new evidence and global comparisons with data that was not available before.

The first study included as Chapter B in this dissertation is titled "Measuring Multidimensional Inequality and Conflict in Africa and in a Global Comparison" and is authored by Joerg Baten and me. We have constructed an indicator that takes into account the multidimensional concept of inequality by considering not only income inequality but also health inequality and land inequality. For all three dimensions, we calculated the Gini coefficient and combined them into a joint inequality index by using a normative approach. By doing so, we were able to construct a broad dataset for our joint index from 1810 to 2010 for 193 countries worldwide, which covers 77% of the world population of the last 200 years on average. We find that the risk of a civil war is consistently higher, the higher the level of inequality within a county. We further examine the causal relationship by using an instrumental variable (IV) approach between inequality and the probability of a civil war outbreak. We hereby include countries worldwide, but focus on sub-Saharan Africa, as this region is mainly affected by civil wars: in 2021, 46% of the ongoing conflicts worldwide were located in Africa (Davies et al., 2022). At the same time, this region suffers from high inequality levels, not only in income but also in terms of education or health (UN-DESA, 2020).

Our results suggest that economy-wide inequality increases the risk of civil war. Although this question has been addressed in the literature, no consensus has been reached (see review in Cramer et al., 2005). With our study, we are the first to provide a broad and consistent dataset for over 200 years and, in addition, to overcome data scarcity for developing countries. This allows us to examine the evolution of inequality over time and to conduct a long-run analysis of the impact of multidimensional inequality on civil wars, providing important new insights into the inequality-conflict debate. This is relevant not only for sub-Saharan African countries but also for high-income countries such as the US or the United Kingdom (UK), which have experienced rising inequality in recent decades.

Motivated by the outbreak of the COVID-19 pandemic in early 2020, the single-authored paper in Chapter C deals with the 1918 influenza pandemic and its relationship with inequality. The title of this study is “Social and Economic Disparities and the 1918 Influenza Pandemic: Lessons for Today”. A global disease that affects everyone could still have a disproportionate impact on the poorest in society or on people in developing countries (UN-DESA, 2020). I aim to contribute to this discussion by empirically investigating the relationship between such a pandemic and inequality, in terms of income and health inequality. To this end, I look at one major past pandemic 100 years ago, which might be compared to the recent one: the influenza pandemic of 1918, considering the conditions prevailing at the time. I explore several pathways, such as asymmetric health risks, employment and the demand shock caused by the pandemic. The empirical analysis is based on a sample of 29 countries worldwide. For the measurement of income inequality data is derived from different databases and sources. To overcome data problems for developing countries, I again use anthropometric data as a proxy indicator for health inequality, as updated and used in Chapter B. For the empirical analysis, I run first-difference estimations with the change in inequality from the 1910 birth decade to the birth decade of 1920 as the outcome variable. The findings suggest a positive but statistically insignificant correlation between the pandemics’ mortality rate and income and health inequality. Although I observe an increase in inequality from 1910 to 1920 in most world regions, the main empirical result suggests that this cannot be linked to the 1918 influenza pandemic.

So far, the focus of economists’ research on historical pandemics and inequality has been on country-specific studies in developed countries. This study provides cross-country evidence on the relationship between a major past pandemic and income and health inequality, and can also provide important evidence for a significant number of developing countries. The findings of a positive, but insignificant association between the severity of the 1918 influenza pandemic and disparities in social and economic outcomes are consistent with a global study by Vollmer and Wójcik (2017), but contradict the significant findings for other epidemics (Alfani, 2022; Furceri et al., 2022), and the Black Death (Alfani, 2015; Alfani & Murphy, 2017).

Finally, the third study in Chapter D finally examines possible determinants of inequality; in particular, if the introduction of social insurance mitigates inequalities in height. This paper titled "The Heights of Medical Care: Health Insurance and Inequality in Adult Stature" is written together with my co-authors Alberto Batinti, Joan Costa-Font and Joerg Baten. In this long-run study, we globally investigate if the introduction of Universal Health Coverage (UHC) in a country led to a reduction in the level of height inequalities in the following decades. For the empirical analysis, we use height inequality as the dependent variable, measured as the Gini coefficient. Using a birth cohort approach, we compiled a sample of 134 countries worldwide for the birth decades 1810 to 2000. Our main independent variable, UHC, refers to the decade in which a country first legally introduced a health insurance scheme. In addition, we check whether a country has already achieved access to UHC for 90% or more of its population. With the expansion of health insurance, we observe a positive and substantial reduction in a country's level of height inequality. To assess the causal relationship, we apply an instrumental variable estimation. Our estimates suggest that within-country differences in height inequality declined with the expansion in health insurance.

Previous research has mainly focused on health outcomes for adults and children, rather than inequalities, and historical data has been limited; we contribute to the literature by retrospectively examining the effects of major health insurance expansions for countries worldwide, providing new evidence on the link with inequality. We conclude that reforms that reduce financial barriers to accessing health care can have a significant impact on reducing disparities in health, an call for the adaptation of appropriate social policies.

Taken together, these three papers provide a comprehensive overview of the different dimensions of inequality, its effects and the factors that contribute to it. By compiling and using an updated inequality dataset with anthropological data, it provides new evidence especially for developing countries and for global comparison. The long-term analyses conducted in this dissertation can help inform policy interventions aimed at reducing income and health inequalities and improving the well-being of people worldwide.

Finally, the last chapter provides a summary of the three chapters and their findings and concludes with possible further research and policy implications.

A.4 References

Acemoglu, D. (2003). Patterns of skill premia. *The Review of Economic Studies*, *70*(2), 199–230.

Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic growth*, *8*, 155–194.

Alesina, A., Michalopoulos, S., & Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, *124*(2), 428–488.

Alfani, G. (2015). Economic inequality in northwestern italy: A long-term view (fourteenth to eighteenth centuries). *The Journal of Economic History*, *75*(4), 1058–1096.

Alfani, G. (2022). Epidemics, inequality, and poverty in preindustrial and early industrial times. *Journal of Economic Literature*, *60*(1), 3–40.

Alfani, G., & Murphy, T. E. (2017). Plague and lethal epidemics in the pre-industrial world. *the Journal of economic History*, *77*(1), 314–343.

Alkire, S., & Foster, J. E. (2010). Designing the inequality-adjusted human development index.

Atkinson, A. (1980). Income distribution and inequality of opportunity. *IHS Journal*, *4*, 65–80.

Ayuda, M.-I., & Puche-Gil, J. (2014). Determinants of height and biological inequality in mediterranean spain, 1859–1967. *Economics & Human Biology*, *15*, 101–119.

Banerjee, A. V., & Duflo, E. (2003). Inequality and growth: What can the data say? *Journal of economic growth*, *8*, 267–299.

Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.

Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of economic growth*, 5–32.

Baten, J., & Blum, M. (2011). Anthropometric within-country inequality and the estimation of skill premia with anthropometric indicators. *Review of Economics*, *62*(2), 107–138.

Baten, J., & Juif, D. (2014). A story of large landowners and math skills: Inequality and human capital formation in long-run development, 1820–2000. *Journal of Comparative Economics*, *42*(2), 375–401.

Bergstrom, K. (2022). The role of income inequality for poverty reduction. *The World Bank Economic Review*, *36*(3), 583–604.

Besley, T., & Burgess, R. (2003). Halving global poverty. *Journal of economic perspectives*, *17*(3), 3–22.

Blum, M. (2014). Estimating male and female height inequality. *Economics & Human Biology*, *14*, 103–108.

Bourguignon, F. (2018). World changes in inequality: An overview of facts, causes, consequences, and policies. *CESifo Economic Studies*, *64*(3), 345–370.

Breen, R., & Müller, W. (2020). *Education and intergenerational social mobility in europe and the united states*. Stanford University Press.

Casey, G. P., & Owen, A. L. (2014). Inequality and fractionalization. *World Development*, *56*, 32–50.

Castelló-Climent, A. (2010). Inequality and growth in advanced economies: An empirical investigation. *The Journal of Economic Inequality*, *8*, 293–321.

Chancel, L. (2022). Global carbon inequality over 1990–2019. *Nature Sustainability*, *5*(11), 931–938.

Chancel, L., & Piketty, T. (2021). Global income inequality, 1820–2020: The persistence and mutation of extreme inequality. *Journal of the European Economic Association*, *19*(6), 3025–3062.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, *129*(4), 1553–1623.

Choe, J. (2008). Income inequality and crime in the united states. *Economics Letters*, *101*(1), 31–33.

Choi, S.-j. (2020). Height inequality and socioeconomic implications in korea: Analysis of individuals born between 1890 and 1919. *Journal of biosocial science*, *52*(4), 504–513.

Clifton, J., Díaz-Fuentes, D., & Revuelta, J. (2020). Falling inequality in latin america: The role of fiscal policy. *Journal of Latin American Studies*, *52*(2), 317–341.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, *27*(3), 79–102.

Cramer, C., et al. (2005). *Inequality and conflict: A review of an age-old concern*. United Nations Research Institute for Social Development Geneva.

Currie, J. (2011). Inequality at birth: Some causes and consequences. *American Economic Review*, *101*(3), 1–22.

Davies, S., Pettersson, T., & Öberg, M. (2022). Organized violence 1989–2021 and drone warfare. *Journal of Peace Research*, *59*(4), 593–610.

de Gregorio, J., & Lee, J. (2002). Education and income inequality: New evidence from cross-country data. *Review of income and wealth*, *48*(3), 395–416.

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health affairs*, *21*(2), 13–30.

Deaton, A. (2003). Health, inequality, and economic development. *Journal of economic literature*, *41*(1), 113–158.

Deininger, K., & Squire, L. (1998). New ways of looking at old issues: Inequality and growth. *Journal of development economics*, *57*(2), 259–287.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of development economics*, *84*(2), 755–776.

Ferreira, F. H., Leite, P. G., & Litchfield, J. A. (2008). The rise and fall of brazilian inequality: 1981–2004. *Macroeconomic Dynamics*, *12*(S2), 199–230.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., & Group, O. H. S. (2012). The oregon health insurance experiment: Evidence from the first year. *The Quarterly journal of economics*, *127*(3), 1057–1106.

Fosu, A. K. (2017). Growth, inequality, and poverty reduction in developing countries: Recent global evidence. *Research in Economics*, *71*(2), 306–336.

Furceri, D., Loungani, P., Ostry, J. D., & Pizzuto, P. (2022). Will covid-19 have long-lasting effects on inequality? evidence from past pandemics. *The Journal of Economic Inequality*, 1–29.

Furceri, D., & Ostry, J. D. (2019). Robust determinants of income inequality. *Oxford Review of Economic Policy*, *35*(3), 490–517.

Galor, O., Moav, O., & Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *The Review of economic studies*, *76*(1), 143–179.

Gupta, S., Davoodi, H., & Alonso-Terme, R. (2002). Does corruption affect income inequality and poverty? *Economics of governance*, *3*, 23–45.

Heckman, J. J. (2011). The economics of inequality: The value of early childhood education. *American Educator*, *35*(1), 31.

Helpman, E. (2016). *Globalization and wage inequality* (tech. rep.). National Bureau of Economic Research.

Hippe, R., & Baten, J. (2012). Regional inequality in human capital formation in europe, 1790—1880. *Scandinavian Economic History Review*, *60*(3), 254–289.

Iversen, T., & Soskice, D. (2006). Electoral institutions and the politics of coalitions: Why some democracies redistribute more than others. *American political science review*, *100*(2), 165–181.

Jensen, P. S., & Sørensen, T. V. (2012). Land inequality and conflict in latin america in the twentieth century. *Defence and Peace Economics*, *23*(1), 77–94.

Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *American journal of public health*, *87*(9), 1491–1498.

Kelly, M. (2000). Inequality and crime. *Review of economics and Statistics*, *82*(4), 530–539.

Krugman, P. R. (2000). Technology, trade and factor prices. *Journal of international Economics*, *50*(1), 51–71.

Kuhn, P. M., & Weidmann, N. B. (2015). Unequal we fight: Between-and within-group inequality and ethnic civil war. *Political Science Research and Methods*, *3*(3), 543–568.

Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, *45*(1), 1–28.

Lakner, C., & Milanovic, B. (2016, July). *Global income distribution: From the fall of the berlin wall to the great recession*. https://doi.org/10.1596/29118

Llorca-Jaña, M., Navarrete-Montalvo, J., Araya-Valenzuela, R., Droller, F., Allende, M., & Rivas, J. (2021). Height in twentieth-century chilean men: Growth with divergence. *Cliometrica*, *15*, 135–166.

Milanovic, B. (2016). *Global inequality: A new approach for the age of globalization*. Harvard University Press.

Narayan, A., Van der Weide, R., Cojocaru, A., Lakner, C., Redaelli, S., Mahler, D. G., Ramasubbaiah, R. G. N., & Thewissen, S. (2018). *Fair progress?: Economic mobility across generations around the world*. World Bank Publications.

Neumayer, E. (2005). Inequality and violent crime: Evidence from data on robbery and violent theft. *Journal of peace research*, *42*(1), 101–112.

Neumayer, E., & Plümper, T. (2016). Inequalities of income and inequalities of longevity: A cross-country study. *American journal of public health*, *106*(1), 160–165.

Nivette, A. E. (2011). Cross-national predictors of crime: A meta-analysis. *Homicide Studies*, *15*(2), 103–131.

Ostry, J., Berg, A., & Tsangarides, C. (2014). Redistribution, inequality, and growth. *Revista de Economía Institucional*, *16*(30), 53–81.

Pickett, K. E., & Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social science & medicine*, *128*, 316–326.

Piketty, T. (2013). *Le capital au xxie siècle*. Média Diffusion.

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts: Methods and estimates for the united states. *The Quarterly Journal of Economics*, *133*(2), 553–609.

Pradhan, M., Sahn, D. E., & Younger, S. D. (2003). Decomposing world health inequality. *Journal of health economics*, *22*(2), 271–293.

Qasim, M., Pervaiz, Z., & Chaudhary, A. R. (2020). Do poverty and income inequality mediate the association between agricultural land inequality and human development? *Social Indicators Research*, *151*, 115–134.

Ramos, X., & Van de Gaer, D. (2021). Is inequality of opportunity robust to the measurement approach? *Review of Income and Wealth*, *67*(1), 18–36.

Ravallion, M. (2005). A poverty-inequality trade off? *The Journal of Economic Inequality*, *3*, 169–181.

Ravallion, M. (2018). Inequality and globalization: A review essay. *Journal of Economic Literature*, *56*(2), 620–642.

Robinson, J. A., & Acemoglu, D. (2012). *Why nations fail: The origins of power, prosperity and poverty*. Profile London.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of economic growth*, *9*, 131–165.

Rohde, N., & Guest, R. (2018). Multidimensional inequality across three developed countries. *Review of Income and Wealth*, *64*(3), 576–591.

Roser, M., & Cuaresma, J. C. (2016). Why is income inequality increasing in the developed world? *Review of Income and Wealth*, *62*(1), 1–27.

Sachs, J. D. (2006). *The end of poverty: Economic possibilities for our time*. Penguin.

Schwekendiek, D., & Baten, J. (2019). Height development of men and women from china, south korea, and taiwan during the rapid economic transformation period of the 1960s–1980s. *Economics & Human Biology*, *34*, 169–180.

Sen, A. (2005). Human rights and capabilities. *Journal of human development*, *6*(2), 151–166.

Singu, S., Acharya, A., Challagundla, K., & Byrareddy, S. N. (2020). Impact of social determinants of health on the emerging covid-19 pandemic in the united states. *Frontiers in public health*, *8*, 406.

Steckermeier, L. C., & Delhey, J. (2019). Better for everyone? egalitarian culture and social well-being in europe. *Social Indicators Research*, *143*(3), 1075–1108.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

Stiglitz, J. E., Sen, A., Fitoussi, J.-P., et al. (2009). Report by the commission on the measurement of economic performance and social progress.

Sylwester, R. (2003). *A biological brain in a cultural classroom: Enhancing cognitive and social development through collaborative classroom management*. Corwin Press.

United Nations Department of Economic and Social Affairs. (2020). *World social report 2020: Inequality in a rapidly changing world*. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/02/World-Social-Report2020-FullReport.pdf.

United Nations Development Programme. (2023). Human development. https://hdr.undp.org/about/human-development.

United Nations Economic and Social Commission for Asia and the Pacific, UNESCWA, UNECA, & UNECLAC. (2018). Promoting equality: An interregional perspective.

Vergolini, L. (2011). Social cohesion in europe: How do the different dimensions of inequality affect social cohesion? *International Journal of Comparative Sociology*, *52*(3), 197–214.

Vollmer, S., & Wójcik, J. (2017). The long-term consequences of the global 1918 influenza pandemic: A systematic analysis of 117 ipums international census data sets. *Available at SSRN 3083584*.

World Bank. (2022). *Inequality in southern africa: An assessment of the southern african customs union*.

B Measuring Multidimensional Inequality and Conflict in Africa and in a Global Comparison[6]

Abstract

We construct a multidimensional inequality index covering 193 countries worldwide with a specific focus on Africa. For a substantial and unprecedented number of countries, we can trace the long-term evolution of inequality over 200 years, from 1810 to 2010. The inequality index includes not only post-tax income inequality but also health and land inequalities. We observe that the risk of a civil war outbreak is consistently increasing with high levels of within-country inequality. By applying an instrumental variable approach, we discover that the impact of multidimensional inequality on civil war is most likely causal. This finding is not only relevant for unstable low- and middle-income countries like Chad or South Sudan but also has implications for high-income countries, such as the United States and the United Kingdom, for which we predict an increased likelihood of civil war.

[6]This chapter was co-authored with Joerg Baten.

B.1 Introduction

Among the terrible conflicts that a society can experience, civil war is the most atrocious. Large numbers of killings within a country, often even within families or the same neighbourhood, is a horrible and almost unimaginable experience by those who did not suffer from it. Economists can identify risk factors, which increase the likelihood of civil war occurrence and can devise strategies to reduce this risk. Consequently, new studies that suggest better and more complete risk factor measurements have a substantial value added.

Does within-country inequality imply a high risk of conflict? Previous studies cannot confirm the positive link between nationwide income inequality and conflict (Collier & Hoeffler, 2004). One of the reasons for this non-result may be that most studies can only include evidence on income inequality on a larger scale from the 1980s onwards, particularly for African countries; hence, the available datasets are small (see review of Cramer et al., 2005). Recently, a new generation of studies has argued that only inequality between groups ('horizontal inequality') correlates with the outbreak of civil wars, rather than nationwide inequality (Koubi & Böhmelt, 2014). However, Wucherpfennig et al. (2016) have criticized these approaches, as they have not been able to establish causality via instrumental variable (IV) techniques.[7] The hypothesis that we study partially contradicts and complements these views, as we assess whether within-country inequality predicts civil war, if a long-term perspective is adopted, and if a comprehensive inequality indicator that covers more than income inequality is employed as an explanatory variable of main interest.

Our study therefore takes a long-term perspective on nationwide inequality in 193 countries and its relationship with the outbreak of conflicts over two centuries, with a particular focus on African countries. For the first time, evidence on developing countries is available with sufficient quality for early decades on a broad scale.

As inequality is a heterogeneous concept with several dimensions, we expand the measurement of this variable by including three components. We include not only post-tax income inequality but also health and land inequalities. Evidently, health is an important dimension of welfare, as human

[7]Hence, Wucherpfennig et al. (2016) have suggested the ethnic identity of the first post-independence government as an instrument—an exception in this literature. Baten and Mumme (2013) are also exceptions. They have instrumented nationwide inequality in a similar way as we do but with a much smaller dataset. They also restrict inequality to health inequality.

beings are more interested in an additional healthy year of life compared to a substantial unit of additional income if they already have a decent income level (Sen, 2005). We use height inequality as a proxy indicator for health inequality, which is now an established indicator of inequality in long-run studies that include the developing world (Blum, 2014; Fogel et al., 1982; Moatsos et al., 2014; Moradi and Baten, 2005; van Zanden et al., 2014b; see Section B.2.1). As a third dimension apart from post-tax income and health, we also include land inequality. Land inequality is crucial for agricultural economies, especially as we adopt a long-term perspective over the past 200 years and include developing countries.

As a preview of our findings, we observe a positive and robust relationship between inequality and the outbreak of civil war: high inequality increases the probability of civil wars. By applying an IV approach, we determine that the effect of multidimensional inequality on civil war is causal.

Which mechanism do we have in mind? Inequality of welfare is clearly a major source of dissatisfaction among those who receive less income, health, and other welfare-providing items. Social groups are inclined to join a rebel group if they are deprived of important resources that insure the well-being of themselves and their families. Collier and Hoeffler (2004) have summarized this mechanism as a grievance that can be one of the major mechanisms in initiating civil wars. Part of this grievance mechanism is caused by land inequality, which is usually measured as the share of landholders to total land area (Baten & Juif, 2014; Galor et al., 2009; Hippe & Baten, 2012; Qasim et al., 2020).[8] Land inequality is seen as one of the main drivers of structural inequality (Easterly, 2007). Given that land estates are often inherited, land inequality is perceived as particularly unjust and 'undeserved' (Baten & Hippe, 2018). Hence, the driving force effect for civil conflict may be particularly strong from this source of inequality, as illustrated by the Russian October Revolution, which occurred due to phenomenal land inequality.

The impact of inequality on civil war is not limited to the poorest world regions like sub-Saharan Africa (SSA). Given that within-country inequality has risen substantially over the past four decades in the United States (US) and the United Kingdom (UK), one obvious question is whether the likelihood of civil war in the US, the UK and in Russia has also increased. According

[8] The inequality of land ownership contributes to the divergence of per capita incomes in several countries (Easterly, 2007).

to our findings, civil war risk has surged dramatically, from 10% to 21% in the US. The conflicts in the last decade might have been only the first signals of an intense civil conflict.

We contribute to a number of important strands of the literature. Our main contribution is to the field of civil war studies. The studies of Fearon and Laitin (2003), and Collier and Hoeffler (2004) have been among the most cited; they observed that standard measures of income inequality do not determine civil war onset over the last decades. By contrast, studies that focused on between-group inequalities have observed positive correlations between inequality and conflict (Bartusevičius & Gleditsch, 2019; Koubi & Böhmelt, 2014; Østby, 2008; Stewart, 2016; Stewart et al., 2008). Inequality between different groups is defined along ethnic, regional, or religious boundaries, and the degree of inequality is correlated with the outbreak of conflicts (Stewart et al., 2008). Between-group differences are obviously an important part of overall inequality. Our study strongly revises the dominant negative view of the literature about economy-wide inequality effects on civil war.

Second, we contribute to the literature about measuring inequalities that may exhibit a civil conflict effect. Cramer et al. (2005) have discussed the challenges related to the cross-country comparability of inequality data. For example, the coverage of early surveys and the units of observations considerably vary (Dollar & Kraay, 2002). McGregor et al. (2019) have indicated that one crucial failure of traditional income surveys is the under-reporting of impoverished and high-income households. Finally, a main limitation of inequality data for studying its effect on conflict lies in the availability of data. For example, Fearon and Laitin (2003) have stated that due to insufficient data on inequality, especially for developing countries, the direct relationship between inequality and the emergence of civil wars has not been accurately studied. Cramer et al. (2005) have concluded that the data availability of inequality is severely lacking.

Third, inequality is a heterogeneous concept with several dimensions; hence, we must consider heterogeneous living conditions. Sen (2005) has provided an approach that examines the spectrum of possibilities a person has rather than the aggregated level of well-being. This strategy considers different dimensions of human development, which also differ between and within countries (Sen, 2005). As a consequence, inequality analysis (and the study of its implications) must not only consider income as a proxy for well-being but must also include other dimensions, such as life circumstances at birth (Stiglitz, 2012). Banerjee and Duflo (2011) have described how life without proper access to the health care system is for the poor and how severe the health consequences of early life conditions are. To address this aspect of inequality, we also consider circumstances that are beyond one's control, such as nutrition, health care, and social circumstances during childhood.

We also contribute to the inequality literature by combining available data on income distribution in contemporary and early societies with anthropological measures. This literature offers enormous potential in analyzing the development and impact of inequality over time. Hence, one of the main contributions of this paper is to provide a broad dataset of inequality that goes beyond income inequality in terms of its time coverage and dimension while still being correlated to it.

In Section B.2, we first present our data and methods; we then show a cross-validation of our measurement by comparing different inequality measures. In the next step, we empirically analyse the relation between inequality within countries and the probability of civil war outbreaks (using a whole battery of different models, control variables, and specifications). We then assess the robustness of our results in Section B.3, circumventing potential endogeneity via an IV estimation. We conclude with a discussion of our results and provide policy recommendations.

B.2 Measuring economic inequality

Previous studies have mainly used income inequality as a proxy for the inequality of well-being, which they measured by the Gini coefficient of income. However, inequality is a heterogeneous concept with several dimensions. Therefore, we calculate a multidimensional measure of economic inequality. Furthermore, although crucial for a long-term analysis of civil war determinants, income inequality estimates for developing countries before the 1980s are almost non-existent. For developed countries, income inequality estimates have been largely undocumented for the last 200 years; however, using height and land inequality increases coverage.

B.2.1 Health inequality

The average height of populations is now a well-established indicator of the quality of diet and health care of past populations (Baten & Komlos, 1998; Fogel et al., 1982; Steckel, 1995). Insufficient or low-quality nutrition, medical care, or shelter during the childhood period determines the growth of an individual. Specifically, family background and social status matter for the final height of an individual. While genetic factors may play a strong role in individual height variation, at the population level, this variable exhibits low relevance if large samples are used. Banerjee and Duflo (2011) have concluded that genetic differences in height between populations are minimal.[9] Baten and Blum (2011) have argued that the distribution of height between individuals shows unequal access to food, health care, and social circumstances during childhood and adolescence. In unequal societies, relatively poor individuals receive less or qualitatively worse nutrition, housing, and medical care. These differences lead to an increase in variation of heights when a cohort reaches adulthood (Baten & Blum, 2011). We conclude that the literature interprets height variation as a reflection, to some extent, of general inequality within a country. Height inequalities are related to income and health inequalities, and they also mirror unofficial family income, such as that based on farming (Choi, 2020).

Using anthropological data—i.e. the distribution of heights—as an indicator for the level of inequality has already been widely used in empirical studies (Ayuda & Puche-Gil, 2014; Baten & Blum, 2014; Baten & Komlos, 1998; Baten & Llorca-Jaña, 2021; Baten & Mumme, 2013; Blum, 2014; Choi, 2020; Guntupalli & Baten, 2006; López-Alonso, 2007; Moradi & Baten, 2005; Schwekendiek & Baten, 2019; van Zanden et al., 2014b).[10]

In our study, we use height distribution data as a proxy for health inequality. To obtain a comprehensive view of several facets of inequality, we combine health inequality, post-tax income inequality, and land inequality as the three components of a joint index, using the appropriate weights discussed below. As the average height of a population is an output-oriented measure that reflects the circumstances experienced early in life (Blum & McLaughlin, 2019), the advantages of health inequality data may outweigh those of income inequality for the following reasons: as

[9] Genetic factors are important for the determination of an individual height. This property shows that deviation can considerably expound inequality in a country.

[10] See also literature review in Blum (2014).

Sen (1987) argues, income cannot measure poverty. By contrast, height may directly reflect social circumstances, family background, and access to medical care or food (Pradhan et al., 2003). In addition, data on height is readily available, as it is included in many family surveys, which provide access to poor households. Therefore, height inequality data has some important advantages (Pradhan et al., 2003).

Our dataset is partly based on the data collection of heights by Baten and Blum (2011), which is available via the website of Clio infra.[11] Moreover, we substantially extend this dataset in its coverage of countries and years. One important source for height is the Demographic and Health Surveys (DHS) Program. The DHS are household surveys conducted at the national level in developing countries. The goal is to monitor and analyze representative data in the fields of population, health, and nutrition. Data on height is mainly available for women.[12] In addition to the DHS data for developing countries, the main surveys we include are the European Social Survey for European countries from 1930 to 1990 and the East Asian Survey for China, South Korea, and Taiwan. Furthermore, we include male height from North Africa, Asia, and Oceania collected in nation-specific anthropological studies and compiled by Grasgruber et al. (2016), as well as other individual height studies (see Appendix).

We calculate health inequality as a coefficient of variation (CV) of the final heights of adults, which is measured in centimetres. We exclude individuals aged below 22 years or older than 50 years from our sample. Young adults may have not yet reached their final height, and some individuals may be old enough to start shrinking. To avoid upward and/or downward bias, we restrict the samples to the above-mentioned age span. Following the methods of Baten and Blum (2011), we initially calculate the CV, which is the standard deviation divided by the mean and expressed as percentages. We transform the calculated CV into the height Gini values using the formula from Moradi and Baten (2005), which is $Height\ Gini = 33 + 25CV$. They based this formula on a multi-country sample of developing countries. This formula has been confirmed by other close estimates (van Zanden et al., 2014a).

[11] For additional information on this data collection, see https://clio-infra.eu/Indicators/HeightGini.html and https://datasets.iisg.amsterdam/dataset.xhtml?persistentId=hdl:10622/IAEKLA.

[12] Demographic and Health Surveys (DHS) Program: https://dhsprogram.com/Methodology/Survey-Types/DHS.cfm.

For our analysis, we concentrate on ten-year periods to eliminate year-specific random fluctuations. We calculate the height CV for every birth cohort of the respective country; for example, the birth decade of 1910 includes the years 1910–1919. After applying the restrictions described above and dropping observations with missing or obviously incorrect information, we construct a dataset of 928 Gini height values for the period 1810–2000 for 127 countries worldwide. Most observations are available after 1950. However, in the first decades until 1870, a total of 54 countries are available, including developing countries from sub-Saharan Africa, Latin America, and other continents. A detailed overview of the countries and time periods in our sample, as well as their sources, is provided in the Appendix, Table B.12.

When comparing our anthropometric inequality measure to income inequality data, we anticipate a positive correlation. However, we do not expect a perfect correlation between these measures. Health inequality, rather than income, reflects living conditions in a very broad sense, such as access to public health services (e.g., hospitals), nutrition, or other services during childhood. Sometimes poorer individuals receive additional income, for example, as development aid transfer (Moradi & Baten, 2005).

In the analyses that follow, we compare various inequality databases and indicators. For this comparison, we use data from the Organisation for Economic Co-operation and Development (OECD), World Income Inequality Database (WIID), World Inequality Database (WID), the World Bank (Milanovic, 2013), and from van Zanden et al. (2014b). Firstly, for the OECD, the data refers to the distribution of gross household income across individuals (Gini coefficients). The OECD data covers the time span of 1976–2019 (Organisation for Economic Cooperation and Development [OECD], 2020). Second, we include data for income Gini coefficients from the World Bank.[13] As the data is based on household surveys, it only covers the periods starting from 1974 (World Bank, 2021). The 'All the Ginis Dataset' compiled by Branko Milanovic combines eight existing databases into one comprehensive one, covering the period of 1950–2012. This dataset is available from the World Bank and includes 187 economies on a yearly basis (Milanovic, 2013).

[13]The World Bank, Development Research Group, receives data directly from different national statistical agencies, in addition to its own country departments. Annual data is available for 170 economies from the World Bank Poverty and Equity Database.

In their book, van Zanden et al. (2014b) have provided a long-run dataset on income inequality to study within-country inequality with observations from 1820 until 2000. They use a variety of sources to construct a broad dataset, including 'Williamson' estimates based on the proportion of GDP to real wages of unskilled workers (van Zanden et al., 2014a). The WID, created by the Piketty group, collects historical income data. By using tax statistics and different surveys, the WID team has built a database that provides long-run data on income and wealth distribution with a large country and time coverage (Alvaredo et al., 2020).[14] Data is available for top income shares for up to 11 countries from 1870 to 1910 and 74% of all countries after 1990.[15] However, Alvaredo et al. (2020) have mentioned that several countries are not fully covered by this data collection; hence, some imputations are necessary to reach a high coverage.

Our collection of height Gini data fills many remaining gaps and allows us to check several dimensions of inequality beyond income. We now assess how the measurement of health inequality is related to these income inequality indicators. As shown in Figure B.1, we observe a close relationship between income and health inequality. As expected, it shows a positive correlation, with only a few observations deviating from the trend line. These observations include Scandinavian countries at the bottom left of the graph, which exhibit an equitable distribution of health in the 1960s and 1970s. In the upper right corner of this figure, Mexico emerges with very high differences in height distribution and high levels of income inequality in the 1970s and 1980s. Even with high economic growth in the previous decades, high health inequality is still observed today (López-Alonso & Condey, 2003). The correlation coefficient of health inequality and income inequality is significant ($\rho = 0.34$, $P<.000$). As displayed in Table B.1, health inequality is significantly correlated when using measurements of the Ginis coefficient of income from different sources. Gini coefficients of income and height are positively correlated and significant for data from the top 10% income share from the WID. Given that only a few Gini coefficients are calculated in the WID dataset, we add the data based on their top 1% shares that are transformed into income Gini coefficients for comparison (Table B.2). We also compare the Gini coefficients of income only from van Zanden et al. (2014a).

The strength of height (inequality) data is to provide evidence for developing countries, espe-

[14] See World Inequality Database https://wid.world/wid-world/. Also see Thomas Piketty (2001, 2003), Piketty and Saez (2003), and the two multi-country volumes on top incomes edited by Atkinson and Piketty (2007, 2010); Atkinson et al. (2011).

[15] Other databases include Milanovic, Lindert, and Williamson (2008); and the WIID, with earliest data from 1867.

cially during early periods. By contrast, in rich and highly developed countries, income inequality data offers a particularly informative source on inequality, whereas nutrition levels or health care at a basic level are already available for poorer parts of the population (and hence height distribution is less informative for these rich countries). It is therefore particularly important to obtain good coverage of our 'income inequality' component for the richest and most developed countries, which is fortunately available.

Another important contribution to overall inequality is land inequality, which is usually measured as the Gini coefficient of all agricultural holdings. The inequality of land ownership contributes to the divergence of per capita incomes in many countries but is most pronounced in agricultural countries (such as Russia before the mid-20$^{\text{th}}$ century). A combination of all three measures therefore provides a unique and broad coverage.

B.2.2 Land inequality

High land inequality values indicate the degree to which large landowners dispose and produce on their land.[16] In the case of Latin America, Frankema (2005) has shown that the unequal distribution of land caused by colonial rule is accompanied by high income inequality (see also Baten and Juif, 2014; Qasim et al., 2020).

Our sample for land inequality is based on those of Frankema (2005, 2010)[17] and Baten and Juif (2014). Land inequality is measured as the Gini coefficient of plot sizes of estates. The basic data processed by Frankema (2005) is obtained from the census of agriculture from the Food and Agriculture Organization (FAO) Census in 1980–2000. The FAO report is published with a ten-year interval and includes data on land holding for over 127 countries (Food and Agriculture Organization [FAO], 2019).

Land inequality is typically not changing significantly over time unless a successful land reform or substantial industrialization development has occurred, during which labourers move from agriculture to industry and services (Baten & Juif, 2014). Therefore, if we do not observe substantial interventions, such as land reforms or industrialization, we anticipate that land inequality is stable over time. Following the adjustments made by Baten and Juif (2014) and building on their

[16] Landowners are defined as those who produce on their own or on rented land.
[17] Frankema (2010) is the updated and corrected version.

collection of land reforms, we interpolate the data on Gini coefficients of land backwards in time if we have a minimum of two observations for a specific country. In addition, if we have data on land reforms, we can calculate the estimated effect of land reform. As we extend our sample on land Ginis in its temporal and geographical coverage, we also expand the dataset on land reforms: we include data from World Conference on Agrarian Reform and Rural Development (WCARRD, 1988) to fill missing decades for several countries and add some missing countries, covering a total of 143 countries. Next, we estimate the average effect of a land reform, arriving at an effect on the reduction of land inequality of 4.47 Gini points (following Baten and Juif, 2014). A detailed calculation is shown in Appendix B.7.2 and Table B.14. The slightly smaller average effect of a land reform compared with that of Baten and Juif (2014) (-5.57 points) may be explained by newly added data on recent decades. We mainly extend the sample with data for recent years starting in 2000, during which land reforms may not have had such a large effect on land inequality compared with the period, for example, around 1900. The backward projection approach allows us to gain many observations for our analysis.

B.2.3 Calculation of the multidimensional joint inequality index

This study aims to construct a broad dataset of inequality, which goes beyond post-tax income inequality. After identifying health and land inequality as suitable inequality measures, the main challenge is to decide about an appropriate weighting strategy: to what extent must each of the alternative inequality measures contribute to the joint inequality index?

First, we compile the income inequality component from the different datasets. Our strategy is as follows: if the data on post-tax income inequality measured by the Gini coefficient is available, we take this value for the income component. As the WIID provides the highest amount of income data, we derive post-tax income data from this source. This dataset provides data for up to 170 countries from 1990 to 2019 (United Nations University [UNU-WIDER], 2021). The WID provides historical data for the top 1% and top 10% income shares for a broad number of countries and times but does provide Gini coefficients only for selected countries. To gain a profound comparison, we calculate the missing Gini coefficients by regressing the top 1% income share on post-tax income Ginis for the countries with both values. We use the formula calculated from the regression displayed in

Table B.2. Using this technique, we estimate the missing Gini coefficients for 1870 to 2010 and for additional six countries.

Given the high number of estimated income Ginis combined with our newly collected height Ginis and land inequality data, we construct our multidimensional inequality index as follows:

$$JointInequalityIndex = Income\ Gini_{it}\alpha_1 + Height\ Gini_{it}r_{it}\alpha_2 + Land\ Gini_{it}(1 - r_{it})\alpha_3,$$

where r_{it} reflects the urbanisation rate in country i at time t. We weight the different dimensions of inequality as $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$.[18] Land inequality matters less for highly urbanized and less agricultural societies. We take into account the degree of urbanisation, as it varies considerably from country to country. In Sierra Leone, for example, more than half of the six million inhabitants live in rural areas. According to data from 2013, over 60% of the labour force works in agriculture, with women being responsible for harvesting and processing the cassava crop, while men are typically involved in rice cultivation and tree crops (FAO, 2021). By contrast, in Chile in 2013, only 10.8% of the population lived in the countryside and agriculture accounted for 10.3% of employment (FAO, 2021). By combining data on income inequality with health and land inequality data, we construct a dataset for 193 countries and cover a period from 1810 to 2010. This data reflects an average of 77% of the world's population over the last 200 years, as shown in Figure B.2. The coverage of our joint index as a percentage of the world population is displayed in Table B.3. It is much larger than the coverage of any preceding study. For example, for Latin America, we reach 60% coverage already in the 1840s. For the Middle East and North Africa, we reach 40% coverage of the region's population in the birth decade of 1870 and 30% coverage for the sub-Saharan African population in the birth decade of 1890. When only health inequality values are available, we include those to fill gaps (controlling these cases with an appropriate dummy variable strategy). In doing so, we are addressing some of the concerns mentioned above by providing a broad dataset for long-term analysis and various countries.

[18] An explanation of the weighting procedure and alternative weights are given in the Appendix.

B.2.4 Cross-Validation of the new measure: how our anthropological inequality measure correlates with income measurements

In Table B.4, we show how the new joint inequality index is related to the Ginis coefficients of income from different sources. The correlation is highly significant for the income Ginis compared, which is expected given that income Gini is one of the components; however, the correlation is never close to one, which implies that the new index has some remaining value added.

In Figure B.3, we show the development of inequality over time for different measurements. The level of our measure is comparable with other inequality data, which are located between high inequality of WIID and low inequality of OECD estimates. The development of global within-country inequality in the 19$^{\text{th}}$ century is quite stable. Inequality decreased in the early and mid-20$^{\text{th}}$ century and started to increase again after the 1980s (consistent with Lakner and Milanovic, 2016).

In Figure B.4, we show the coverage for our multidimensional inequality index per country with the most recent data available. These data show the (almost) full coverage of countries and the level of inequality worldwide. Figure B.5 displays the development of inequality for selected world regions. Sub-Saharan Africa stands out as the region with the highest levels of inequality, followed by Latin America, with a tendency to decrease slightly since 2000 (similar: López-Calva and Lustig, 2010). However, inequality in North America has been increasing since the 1970s until 2010 (mainly driven by the US). Piketty et al. (2018) have confirmed this development. They argue that since the 1980s, the US has shown a growing discrepancy between the income growth of the poor and the rich.

B.2.5 Conflict data

In this section, we analyse the distribution of civil war onsets worldwide. Data on conflict is available from the Correlates of War Project (COW). This database tracks different types of violent conflicts worldwide and provides data on non-state wars, intrastate wars, and interstate and extra-state wars. As we analyse the impact of inequality on conflicts on a country level, we focus on conflicts that occur within national borders, namely, intrastate conflicts. Conflicts arise due to complex socio-economic interactions and motivations (Raleigh & Kniveton, 2012). In our analysis, we therefore distinguish between different types of civil wars as provided by the COW, which are civil wars over central control and over local issues.[19] The COW defines a threshold of 1,000 conflict-related deaths per conflict per year to be included in their database (Sarkees, 2010). The most recent COW dataset on intra-state wars (v5.1) covers the period from 1816 to 2014. This dataset provides the broadest coverage for conflict over the long run. We exclude the decade of the 2010s from our analyses, as some conflicts occurred after 2015.[20] For our regression analysis, we include civil war onset from all types as a dummy variable that takes a value of one if a new civil war occurred in this country and decade, and zero if not. This factor leads us to 177 observations of civil war outbreaks in 73 different countries from 1810 to 2010, with China as the country with the highest reported number of civil war outbreaks (9), followed by Mexico (7), Argentina, Colombia, Ethiopia Iraq, Russia, and Turkey (6). In Figure B.6, we report the number of civil war outbreaks over time from our sample and the unequal distribution of civil war outbreaks by different world regions on the right side of the graph.

When we look at the whole period regarded, from 1810 to 2010, except for Ethiopia, African countries do not stand out in terms of having high numbers of civil war outbreaks, compared to other countries. But we observe that the number of civil war outbreaks from the 1960s to the 1990s has considerably increased, which is associated with decolonization after 1945 and the sudden presence of many unstable states, mainly in Africa (Fearon & Laitin, 2003). Also when the regional distribution of conflict is examined, two regions stand out: Latin America and SSA. Sub-Saharan African conflicts were most frequent between 1960 and 2000, whereas in Latin America, most conflicts occurred during the 1890s and 1970s. In Figure B.7, we present a global map of civil war outbreaks in the 2000s. We observe the emergence of new civil wars mainly in African and Middle Eastern countries, which is for example the ongoing civil war in Sudan, starting in early 2003 or the First Ivorian civil war in the Ivory Coast. In Figure B.8, we control for the occurrence of different types of civil wars. We observe that most civil wars are about central control, followed by civil war over local issues. To test the relationship between inequality and different aspects of a conflict, we also measure civil war by its severity, in terms of conflict-related deaths over a decade.

[19]We do not include regional internal and intercommunal wars as these were very few and, by definition, not civil wars.

[20]Exemplary conflicts are the second Yemeni Civil war, which is ongoing since 2015; the Anglophone crisis in Cameroon; and the insurgency in Cabo Delgado in Mozambique since 2017; or several conflicts emerging in Latin America (e.g., in Colombia, Venezuela or the prison riots in Brazil).

B.2.6 Control variables

Obviously, inequality is not the only variable that matters. Hence, we include other factors that may determine if, how, and when a conflict occurs. Similarly, Raleigh and Kniveton (2012) have concluded that conflicts arise due to specific circumstances and complex socioeconomic interactions and motivation (See also Nygård, 2018).

We include control variables for population size, and the quality of institutions reflected by the polity2 index[21], colonial background, the history of wars, ethnic fractionalization, and diamond deposits. First, population size is a necessary control variable, since a large country like China almost automatically has a higher likelihood of civil conflict in some corners of the country than a smaller country like Portugal. Second, the quality of institutions and democratic decision-making processes may matter, as Collier and Hoeffler (2004) or Fearon and Laitin (2003) have found that consolidated democracies face a low risk of civil war. Against this background, colonisation and, in sub-Saharan Africa in the 1960s, the process of decolonisation was accompanied by political instability (Acemoglu et al., 2002), which may have increased the likelihood of civil wars. We control for the colonial history of a country by including a categorical variable. A highly fractionalised country in terms of ethnicity, language, or religion may be at greater risk of the outbreak of a civil war; therefore, we include a measure of ethnic fractionalisation in our analysis. We include GDP per capita as a control for the economic development of a country (and for similar purposes, we also include height growth). We are aware that the inclusion of these additional variables in our regression analysis, especially the inclusion of GDP per capita, could be seen as 'bad controls'. We address this concern, first, by including additional variables in a stepwise fashion to show that the relationship between civil war and our main independent variable, inequality, remains unchanged. Second, we base our selection of controls on common control variables in the inequality-conflict literature (see Cramer et al., 2005).

A detailed description of the control variables and their sources is included in the Appendix. In Table B.5, we present the summary of statistics if the main variables for *civil war* and *inequality* are available.

[21] The polity2 index indicates the regime type of a country, from full autocracy to a highly consolidated democracy. For more information, see Polity5 Project, https://www.systemicpeace.org/polityproject.html.

B.3 Does inequality fuel violent conflict?

High inequality tends to undermine social cohesion and to fuel protests and violent conflict (e.g., Vergolini, 2011), but does it increase the likelihood of civil war breaking out? Three main views are discussed on the economic causes of civil wars. Collier and Hoeffler (2004) favour the greed argument: if the benefits to join a rebellion exceed the costs, then the motivation to join a rebellion may be sufficiently high. These benefits may include individual economic situations, such as financial enrichment or control over natural resources in a country, especially if they are 'easily lootable', such as diamond mine products. In countries with low income, the opportunity cost of joining a violent movement is lower than in wealthier societies. This factor provides armed opposition groups with a larger number of people with low opportunity costs in poorer economies (Collier & Hoeffler, 2004). The grievance argument, on the other hand, argued that the motivation for civil war is based on inequalities. The motivation for people to change the status quo must be sufficiently high to engage in violent conflict to resolve these issues (Gurr, 2000). Fearon and Laitin (2003) have supported the view that civil wars tend to occur in countries with weak institutions. They therefore argue that state capacity matters more than the motivation of the people.

Civil wars in sub-Saharan Africa are often explained by the grievance argument that land and income inequalities are high and crucial for understanding the conflict in this region. Fjelde and Østby (2014) tested the association between socio-economic inequality and civil conflict using a panel of 34 SSA countries over the period 1990–2008. Their empirical study reports a significant relationship, which is explained by intergroup grievances. Their findings confirm the impact of inequality on conflict (also consistent with the model of Acemoglu and Robinson, 2006).

B.3.1 Regression results

The regression results of the likelihood of civil war onset are displayed in Table B.6. In nine regression analyses, we identified the correlation between inequality and civil war. First, we have chosen to use the standard model selection strategy to evaluate a bivariate regression (column 1). We add time-fixed effects in column 2 and then include world region- and time-fixed effects in column 3. Finally, we add control variables in columns 3 to 9 and assess different econometric models. We use pooled logit, panel logit, and rare events logit models with different control variables. We include time-fixed effects and world region-fixed effects, as mentioned in the table (except for the rare event model). As logistic models may influence the analysis of a sample where the number of civil wars is small, we also run a logistic regression for rare events data (King & Zeng, 2001). To address heteroskedasticity, we include robust standard errors in all of our models. We lag inequality by one decade, as the civil war response cannot be expected immediately. Moreover, this strategy reduces the endogeneity problem caused by contemporaneous correlation.

We observe a consistently significant positive coefficient of inequality. The coefficients of inequality are very stable, around 0.50 to 0.89, except in the rare events logit model, which is econometrically specified in a different way. The consistency of the inequality effect is the core result here. Moreover, population size is significant and positive in all models, which is in line with Collier and Sambanis (2002). The larger the population of a country, the higher the likelihood of a civil war onset. As expected, a higher level of democratisation lowers the risk of civil war in a country. However, this relation may be non-linear; Collier and Sambanis (2002) have argued that autocratic systems can be quite stable, whereas states that are transitioning to democracy and young and inexperienced democracies should fear violent civil conflicts. Hence, we include the squared terms as well. The colonial background shows no significant coefficient here, neither positive nor negative (different results: Baten and Mumme, 2013). The same holds for the 'greed' dummy variable diamond, which is included in models 4, 6 and 8. The 'greed' proxy for (low) income is significantly correlated with civil war (low income allows easy recruitment of rebels), although low income is also partly a grievance variable, as absolute poverty also reinforces the inequality motivations of rebellion. Regarding the theory that a high diversity of ethnic, linguistic, or religious groups increases the risk of civil war, we observe a positive association with ethnic fractionalisation. In addition, a country's previous wars (*history of wars*) seem to be positively linked with the outbreak of future civil wars. Conflicts over local issues or political power can always reemerge if remain unresolved, as seen in Israel/Palestine, where civil wars arise repeatedly. However, we mainly observe these significant coefficients in our rare event model (column 9). In model 7, we include height growth as a proxy for economic growth. We have found no evidence that this variable is associated with lower or higher risk of civil war.

In order to assess potential selectivity, we study how well our dataset, used in several regression specifications in Table B.6, covers low-income, lower-middle, upper-middle and high-income countries. For three of the four categories, we can obtain a coverage of 25–30%, and for the—always least documented—low-income countries, we still have a respectable 13–15% of all possible country-decade combinations (Figure B.9).

B.3.2 Robustness check

To show the robustness of our results, we use different models and include or exclude time and region-fixed effects, as well as different control variables. Our results remain robust when a linear probability model estimation is applied, as shown in Table B.7. We also compare the results using country-fixed effects as opposed to region-fixed effects. The coefficient of inequality is virtually identical (Table B.8).[22]

Does inequality affect different types of civil wars differently? In Table B.9 we consider several definitions of intrastate wars as well as alternative measures of civil wars. In model 1 we display the regression results from civil wars over central control, and over local issues in model 2. Our findings suggest a positive correlation between inequality and civil war, both over central control and over local issues, whereas the coefficient for the latter is slightly larger. If we multiply with the standard deviations of both variables, the beta coefficient for the local issues is 13% of a standard deviation of the dependent variable, while the beta coefficient of the central control is 8%. We therefore observe a relatively higher association of inequality in fueling civil wars over local issues rather than over central control. In model 3 we test for the relationship between inequality and the severity of a civil war, measured as average civil war deaths in the respective decade and country. We find that higher intra-country inequality is linked with significantly more severe conflicts.

We also control for differences between low- and high-income countries. In Table B.10, we omit poor countries in model 1 and very rich countries in model 2. In both models, the estimated coefficients of inequality are highly significant and positive. Finally, we assessed whether the significant relationship between inequality and civil war depends only on our composite measure of inequality. We used health inequality alone in Appendix Table B.15 (because health inequality covers the

[22] We admit that the country FE model provides less robust results if the sample is reduced by adding other control variables that have missing values for some of the country-decade units and hence reduce the sample size in other specifications.

largest amount of observations) and we find that the significance of inequality is robust using only this inequality concept.

B.3.3 Instrumental variable regression

To circumvent potential endogeneity issues, we apply an IV approach. For example, reverse causality may be an issue: civil war may affect inequality in a country, as Bircan et al. (2017) have noted that economic activity decreases significantly during and after wars. This circumstance affects schooling, health, access to food, and other factors related to equality. Moreover, reverse causality is conceivable, especially in many developing countries, where the family income of the poor mainly relies on physical labour in agriculture, which may be weakened during a civil war.

One possible instrument to address the endogeneity of inequality is suggested by Easterly (2007). His IV, called *wheat-sugar ratio*, refers to the suitability of the soil for sugar divided by the suitability of the soil for planting wheat. This approach is implemented in several studies, sometimes with further modifications (Baten & Juif, 2014; Baten & Mumme, 2013).[23] The use of this variable is based on the observation that the minimum efficient scale of wheat, as well as rice, is small. Hence, farmers can efficiently grow wheat on small farm units. By contrast, the production of sugar requires large plantations and a huge number of workers to be efficient. These sugar plantations were often based on slaves as the primary labour force in earlier times; for example, in the early 19^{th} century Brazil, as the land was suitable and had high return potential, and later by unskilled agricultural workers. This aspect has typically resulted in high inequality. Easterly (2007) has therefore concluded that the ratio of wheat and sugar suitability of soil may be a good instrument for the current inequality level.

Following this approach in Table B.11, we instrument inequality by using the wheat-rice-sugar ratio of soil suitability. We perform an IV approach in the form of a two-stage least square and limited information maximum likelihood in models 1–2 and 3–4, respectively. The F-test shows that the wheat-rice-sugar ratio is a strong instrument following the methods of Staiger and Stock (1994). In the second stage, the inequality effect on civil war onset is still consistently observable.

As for any reasonable instrument we need to discuss the exclusion restriction for the ratio of

[23] For example, Baten and Mumme (2013) use an interaction term of low population density of 1,500 with southern latitude in addition to the wheat-rice-sugar ratio instrumental variable.

crop suitability of wheat, rice and sugar. The exclusion restriction assumes that the relationship between our instrument and civil war outbreaks is fully reflected by inequality and has no direct influence itself. One could imagine an impact of the crop suitability of soil on civil war outbreaks via their production revenues. Exporting sugar cane might have generated exceptionally high revenues in the 19$^{\text{th}}$ century, therefore functioning like gas or oil as a natural resource curse (Frankel, 2010). However, Easterly (2007) argued that commodity wealth not necessarily violates the exclusion restriction if this mechanism is affecting inequality.[24] To test this concern, we include exports as an additional variable in our regression, following Baten and Juif (2014), Baten and Mumme (2013), and Easterly (2007). The variable *export* is defined as a country's exports of raw material and mining in relation to its total exports. As the export coefficient is insignificant and the inequality effect is not affected, this would not suggest a violation of the exclusion restriction (see column 5 of Table B.11).

[24]Easterly (2007) and Baten and Juif (2014) have carefully studied and rejected other causal channels that may imply a violation of the exclusion restriction, such as resource curse effects of sugar plantations. Hence, soil suitability is unlikely to have a direct effect on the outbreak of civil wars other than through inequality.

B.4 Conclusion

Does economy-wide inequality influence the occurrence of civil war conflicts? Although this question was widely addressed in the literature, no consensus was established about whether and which type of inequality increases the likelihood of conflicts. This circumstance was mainly due to the lack of data about inequality, particularly for sub-Saharan Africa. Moreover, the early waves of studies used quite narrow concepts of inequality.

Examining different dimensions of inequality and constructing a broad measurement of economic inequality allowed us to study factors other than income, for example, access to nutrition or healthcare. We constructed a multidimensional inequality index by combining income, as well as health and land inequalities. The inclusion of anthropometric measures enabled us to overcome the data-availability problems for developing countries and allowed us to build a broad dataset for over 200 years, from 1810 to 2010, for a maximum of 193 countries.

In our global long-run analysis of the impact of inequality on the risk of a civil war outbreak, we found that higher within-country inequality significantly increases the risk of a civil war outbreak in a country. Our results remained robust to the application of different models and estimation strategies, including various sets of control variables and time- and region-fixed effects. We also considered the robustness when using only health inequality and observe a very consistent influence of inequality. We addressed the concerns of endogeneity by applying an IV approach.

Traditionally, world regions, such as sub-Saharan Africa and Latin America—where inequality is high in terms of health, income, and social status—were seen as being most at risk of the observed relationship. However, over the past decades, inequality has also substantially risen in the US, the UK and in Russia. Hence, given the relationship that we observed between inequality and civil war, one question is how likely civil conflicts can occur in these high-income countries. Calculating the increase of post-tax inequality in the US, for example, as an increase from a minimum of 27 in the 1970s to a post-tax Gini coefficient of 48 in 2019, we estimated that the likelihood grew from 10% to 20%.[25] Political strategies to reduce this high civil war risk would obviously be progressive taxation efforts, even if we are aware that this measure is not popular among many economic advisors and high-income level taxpayers.

[25] We admit that we infer this from partly cross-sectional evidence.

B.5 References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly journal of economics*, *117*(4), 1231–1294.

Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Alvaredo, F., Atkinson, A. B., Blanchet, T., Chancel, L., Bauluz, L., Fisher-Post, M., Flores, I., Garbinti, B., Goupille-Lebret, J., Martínez-Toledano, C., et al. (2020). *Distributional national accounts guidelines, methods and concepts used in the world inequality database* [Doctoral dissertation, PSE (Paris School of economics)].

Ayuda, M.-I., & Puche-Gil, J. (2014). Determinants of height and biological inequality in mediterranean spain, 1859–1967. *Economics & Human Biology*, *15*, 101–119.

Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.

Bartusevičius, H., & Gleditsch, K. S. (2019). A two-stage approach to civil conflict: Contested incompatibilities and armed violence. *International Organization*, *73*(1), 225–248.

Baten, J., & Blum, M. (2011). Anthropometric within-country inequality and the estimation of skill premia with anthropometric indicators. *Review of Economics*, *62*(2), 107–138.

Baten, J., & Blum, M. (2014). Why are you tall while others are short? agricultural production and other proximate determinants of global heights. *European Review of Economic History*, *18*(2), 144–165.

Baten, J., & Hippe, R. (2018). Geography, land inequality and regional numeracy in europe in historical perspective. *Journal of Economic Growth*, *23*(1), 79–109.

Baten, J., & Juif, D. (2014). A story of large landowners and math skills: Inequality and human capital formation in long-run development, 1820–2000. *Journal of Comparative Economics*, *42*(2), 375–401.

Baten, J., & Komlos, J. (1998). Height and the standard of living. *The Journal of Economic History*, *58*(3), 866–870.

Baten, J., & Llorca-Jaña, M. (2021). Inequality, low-intensity immigration and human capital formation in the regions of chile, 1820-1939. *Economics & Human Biology*, *43*, 101030.

Baten, J., & Mumme, C. (2013). Does inequality lead to civil wars? a global long-term study using anthropometric indicators (1816–1999). *European Journal of Political Economy*, *32*, 56–79.

Bircan, Ç., Brück, T., & Vothknecht, M. (2017). Violent conflict and inequality. *Oxford Development Studies*, *45*(2), 125–144.

Blum, M. (2014). Estimating male and female height inequality. *Economics & Human Biology*, *14*, 103–108.

Blum, M., & McLaughlin, E. (2019). Living standards and inequality in the industrial revolution: Evidence from the height of university of edinburgh students in the 1830s. *Economics & Human Biology*, *35*, 185–192.

Choi, S.-j. (2020). Height inequality and socioeconomic implications in korea: Analysis of individuals born between 1890 and 1919. *Journal of biosocial science*, *52*(4), 504–513.

Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563–595.

Collier, P., & Sambanis, N. (2002). Understanding civil war: A new agenda. *Journal of Conflict Resolution*, *46*(1), 3–12.

Cramer, C., et al. (2005). *Inequality and conflict: A review of an age-old concern*. United Nations Research Institute for Social Development Geneva.

Dollar, D., & Kraay, A. (2002). Growth is good for the poor. *Journal of economic growth*, *7*(3), 195–225.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of development economics*, *84*(2), 755–776.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American political science review*, *97*(1), 75–90.

Fjelde, H., & Østby, G. (2014). Socioeconomic inequality and communal conflict: A disaggregated analysis of sub-saharan africa, 1990–2008. *International Interactions*, *40*(5), 737–762.

Fogel, R. W., Engerman, S. L., & Trussell, J. (1982). Exploring the uses of data on height: The analysis of long-term trends in nutrition, labor welfare, and labor productivity. *Social Science History*, *6*(4), 401–421.

Food and Agriculture Organization. (2019). *Main results and metadata by country (2006–2015). world programme for the census of agriculture 2010.*

Food and Agriculture Organization. (2021). *Gender and Land Rights Database*. Retrieved June 9, 2021, from http://www.fao.org/gender-landrights-database/country-profiles/en/.

Frankel, J. A. (2010). *The natural resource curse: A survey* (tech. rep.). National Bureau of Economic Research.

Frankema, E. H. (2005). *The colonial origins of inequality: Exploring the causes and consequences of land distribution* (tech. rep.). IAI Discussion Papers.

Frankema, E. H. (2010). The colonial roots of land inequality: Geography, factor endowments, or institutions? *The Economic History Review*, *63*(2), 418–451.

Galor, O., Moav, O., & Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *The Review of economic studies*, *76*(1), 143–179.

Grasgruber, P., Sebera, M., Hrazdíra, E., Cacek, J., & Kalina, T. (2016). Major correlates of male height: A study of 105 countries. *Economics & Human Biology*, *21*, 172–195.

Guntupalli, A. M., & Baten, J. (2006). The development and inequality of heights in north, west, and east india 1915–1944. *Explorations in Economic History*, *43*(4), 578–608.

Gurr, T. R. (2000). *Peoples versus states: Minorities at risk in the new century*. US Institute of Peace Press.

Hippe, R., & Baten, J. (2012). Regional inequality in human capital formation in europe, 1790—1880. *Scandinavian Economic History Review*, *60*(3), 254–289.

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political analysis*, *9*(2), 137–163.

Koubi, V., & Böhmelt, T. (2014). Grievances, economic wealth, and civil conflict. *Journal of Peace Research*, *51*(1), 19–33.

Lakner, C., & Milanovic, B. (2016, July). *Global income distribution: From the fall of the berlin wall to the great recession*. https://doi.org/10.1596/29118

López-Alonso, M. (2007). Growth with inequality: Living standards in mexico, 1850–1950. *Journal of Latin American Studies*, *39*(1), 81–105.

López-Alonso, M., & Condey, R. P. (2003). The ups and downs of mexican economic growth: The biological standard of living and inequality, 1870–1950. *Economics & Human Biology*, *1*(2), 169–186.

López-Calva, L. F., & Lustig, N. C. (2010). *Declining inequality in latin america: A decade of progress?* Brookings Institution Press.

McGregor, T., Smith, B., & Wills, S. (2019). Measuring inequality. *Oxford Review of Economic Policy*, *35*(3), 368–395.

Milanovic, B. (2013). All the ginis dataset [Data derived from World Bank Group, 2021-5-19, https://datacatalog.worldbank.org/search/dataset/0041738/All-the-Ginis-Dataset].

Moatsos, M., Baten, J., Foldvari, P., van Leeuwen, B., & van Zanden, J. L. (2014). Income inequality since 1820.

Moradi, A., & Baten, J. (2005). Inequality in sub-saharan africa: New data and new insights from anthropometric estimates. *World development*, *33*(8), 1233–1265.

Nygård, H. M. (2018). Inequality and conflict–some good news. *Development for Peace*.

Organisation for Economic Cooperation and Development. (2020). Income inequality [Accessed: 2020-13-12, https://www.oecd-ilibrary.org/content/data/459aa7f1-en.]. https://doi.org/https://doi.org/https://doi.org/10.1787/459aa7f1-en

Østby, G. (2008). Polarization, horizontal inequalities and violent civil conflict. *Journal of Peace Research*, *45*(2), 143–162.

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts: Methods and estimates for the united states. *The Quarterly Journal of Economics*, *133*(2), 553–609.

Pradhan, M., Sahn, D. E., & Younger, S. D. (2003). Decomposing world health inequality. *Journal of health economics*, *22*(2), 271–293.

Qasim, M., Pervaiz, Z., & Chaudhary, A. R. (2020). Do poverty and income inequality mediate the association between agricultural land inequality and human development? *Social Indicators Research*, *151*, 115–134.

Raleigh, C., & Kniveton, D. (2012). Come rain or shine: An analysis of conflict and climate variability in east africa. *Journal of peace research*, *49*(1), 51–64.

Sarkees, M. R. (2010). The cow typology of war: Defining and categorizing wars (version 4 of the data). *Note with version 4 of the Correlates of War Data*.

Schwekendiek, D., & Baten, J. (2019). Height development of men and women from china, south korea, and taiwan during the rapid economic transformation period of the 1960s–1980s. *Economics & Human Biology*, *34*, 169–180.

Sen, A. (1987). Gender and cooperative conflicts.

Sen, A. (2005). Human rights and capabilities. *Journal of human development*, *6*(2), 151–166.

Staiger, D. O., & Stock, J. H. (1994). Instrumental variables regression with weak instruments.

Steckel, R. H. (1995). Stature and the standard of living. *Journal of economic literature*, *33*(4), 1903–1940.

Stewart, F. (2016). *Horizontal inequalities and conflict: Understanding group violence in multiethnic societies*. Springer.

Stewart, F., Brown, G. K., & Langer, A. (2008). Major findings and conclusions on the relationship between horizontal inequalities and conflict. In F. Stewart (Ed.), *Horizontal inequalities and conflict: Understanding group violence in multiethnic societies* (pp. 285–300). Palgrave Macmillan UK. https://doi.org/10.1057/9780230582729_13

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

United Nations University. (2021, June). World income inequality database (wiid) companion dataset [Available online at https://doi.org/10.35188/UNU-WIDER/WIIDcomp-310521; Accessed: 2021-5-19].

van Zanden, J. L., Baten, J., Foldvari, P., & van Leeuwen, B. (2014a). The changing shape of global inequality 1820–2000; exploring a new dataset. *Review of income and wealth*, *60*(2), 279–297.

van Zanden, J. L., Baten, J., Mira d'Ercole, M., Rijpma, A., Smith, C., & Timmer, M. (2014b). *How was life?: Global well-being since 1820*. OECD publishing.

Vergolini, L. (2011). Social cohesion in europe: How do the different dimensions of inequality affect social cohesion? *International Journal of Comparative Sociology*, *52*(3), 197–214.

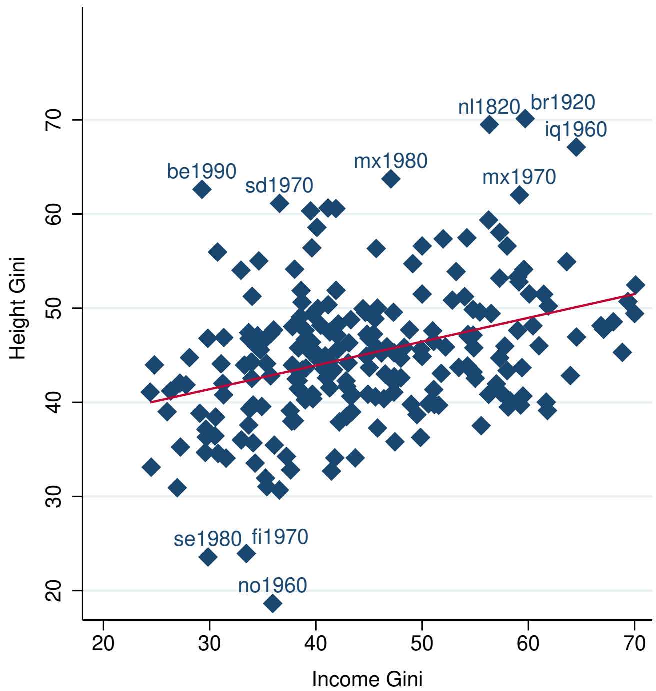
World Bank. (2021). Poverty and equity database [Available online at https://datacatalog.worldbank.org/dataset/poverty-and-equity-database.; Accessed: 2021-5-19].

World Conference on Agrarian Reform and Rural Development. (1988). Guidelines on socio-economic indicators for monitoring and evaluating agrarian reform and rural development.

Wucherpfennig, J., Hunziker, P., & Cederman, L.-E. (2016). Who inherits the state? colonial rule and postcolonial conflict. *American Journal of Political Science*, *60*(4), 882–898.

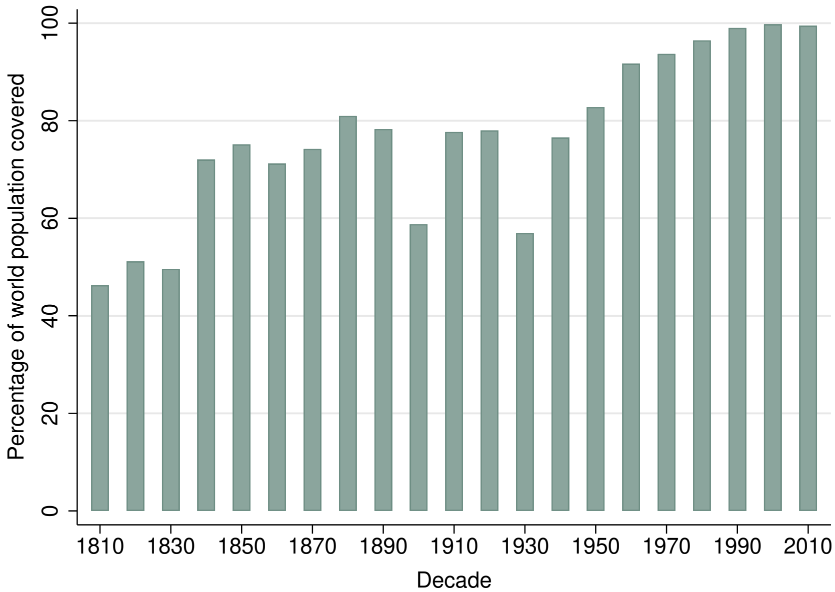
B.6 Tables and figures

Figure B.1: Relationship between height Gini and income Gini



Notes. Data for income Ginis are derived from van Zanden et al. (2014). Labels refer to the country-decade combination. For country codes see Appendix Table B.13.

Figure B.2: Distribution of our joint inequality index by the percentage of world population covered



Notes. The joint inequality index is constructed by using income, health and land inequality data.

Figure B.3: Comparison of different inequality measures

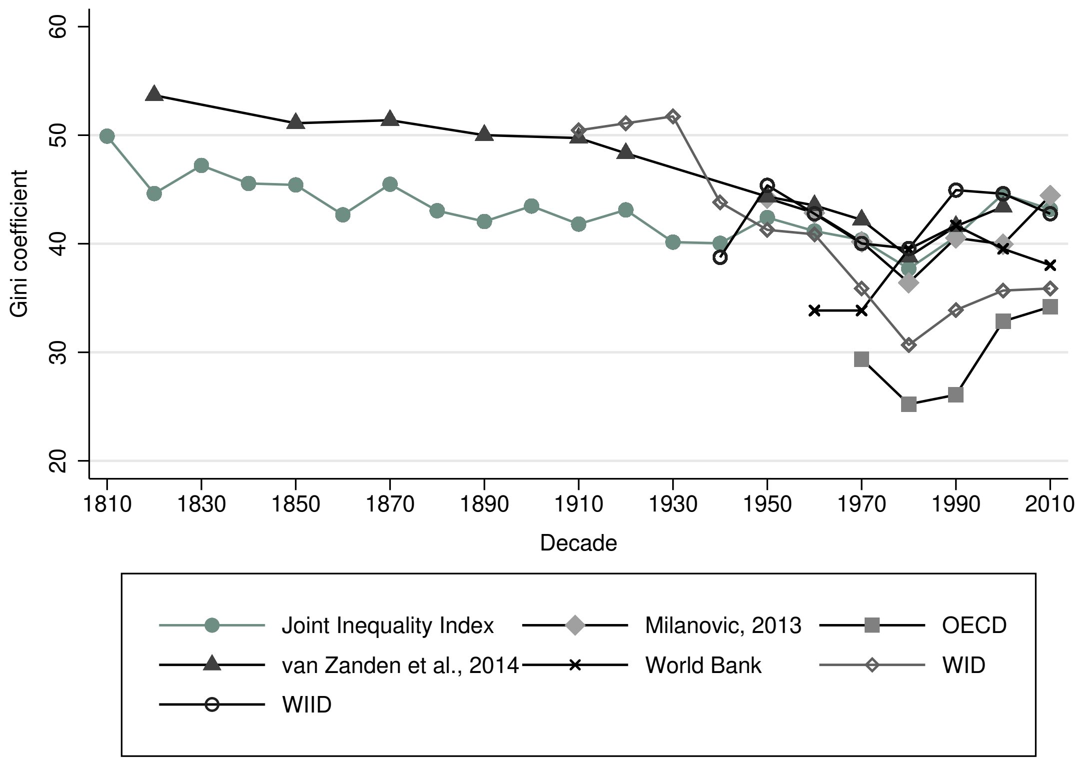
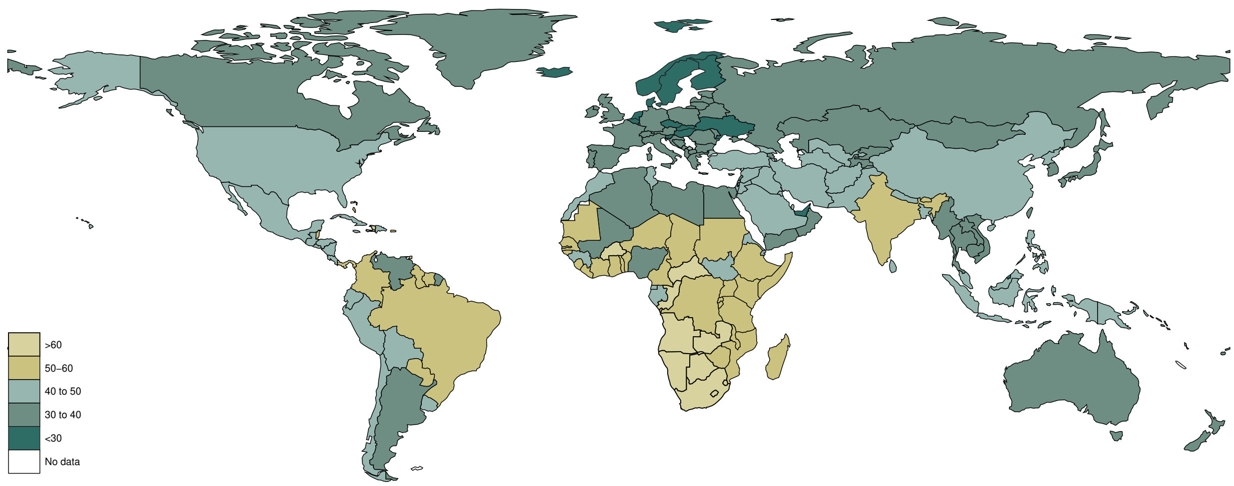


Figure B.4: Joint inequality index, most recent data available



Notes. The joint inequality index is constructed by using income, health and land inequality data. Data refer to the most recent data available which is the decade of 2010 in most cases, and the decade of 2000 for four countries, namely Dominica, Grenada, Kiribati and Puerto Rico.

Figure B.5: Inequality by world region

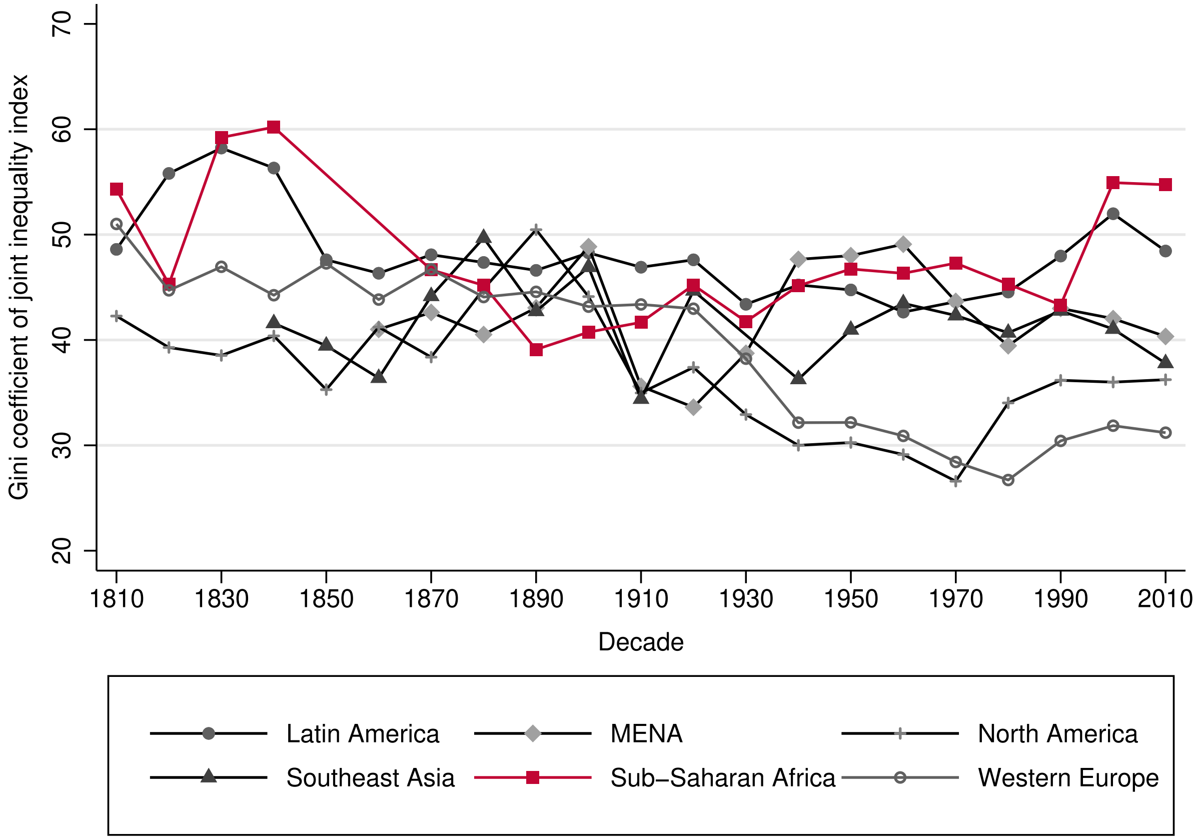
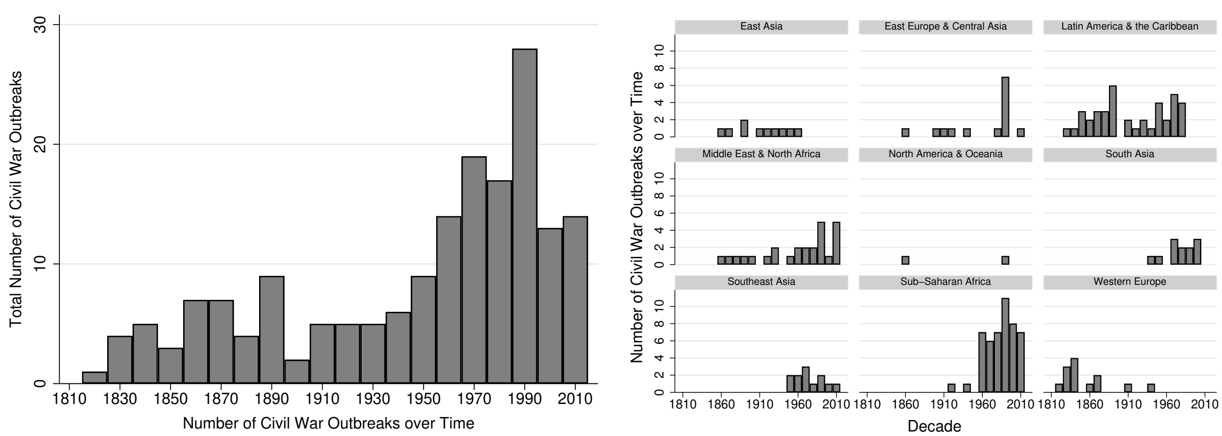
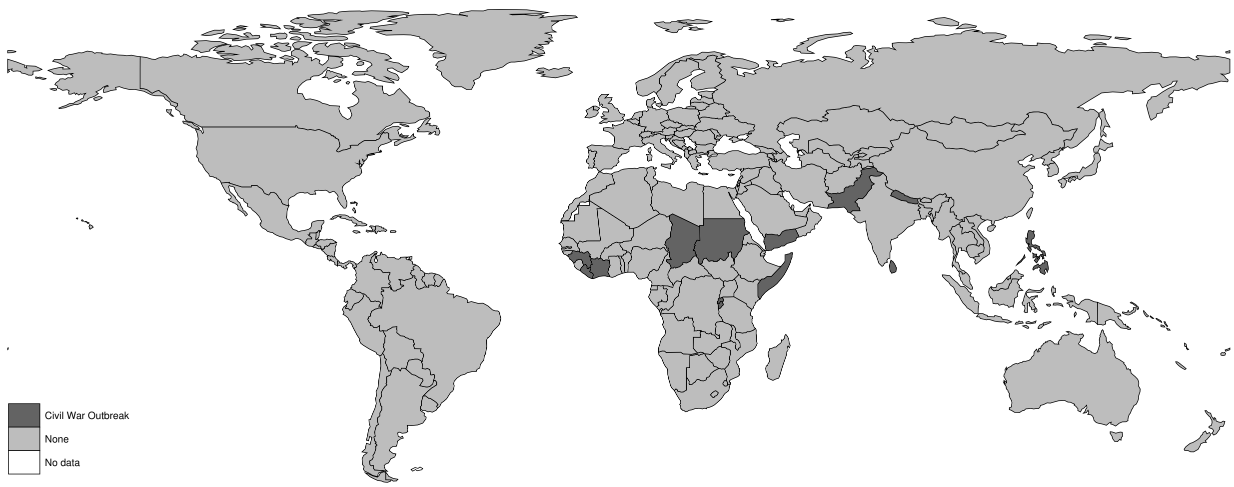


Figure B.6: Number of civil war outbreaks over time and by world region



Notes. Data on civil wars is derived from the Correlates of War (COW) project. We consider one civil war outbreak per decade and country. Data for the 2010 decade only includes civil war outbreaks occurring until 2014.

Figure B.7: Civil war outbreaks worldwide for the decade 2000 (2000–2009)



Notes. Data on civil wars is derived from the Correlates of War (COW) project. We consider one civil war outbreak per decade and country.

Figure B.8: Civil war outbreaks by type

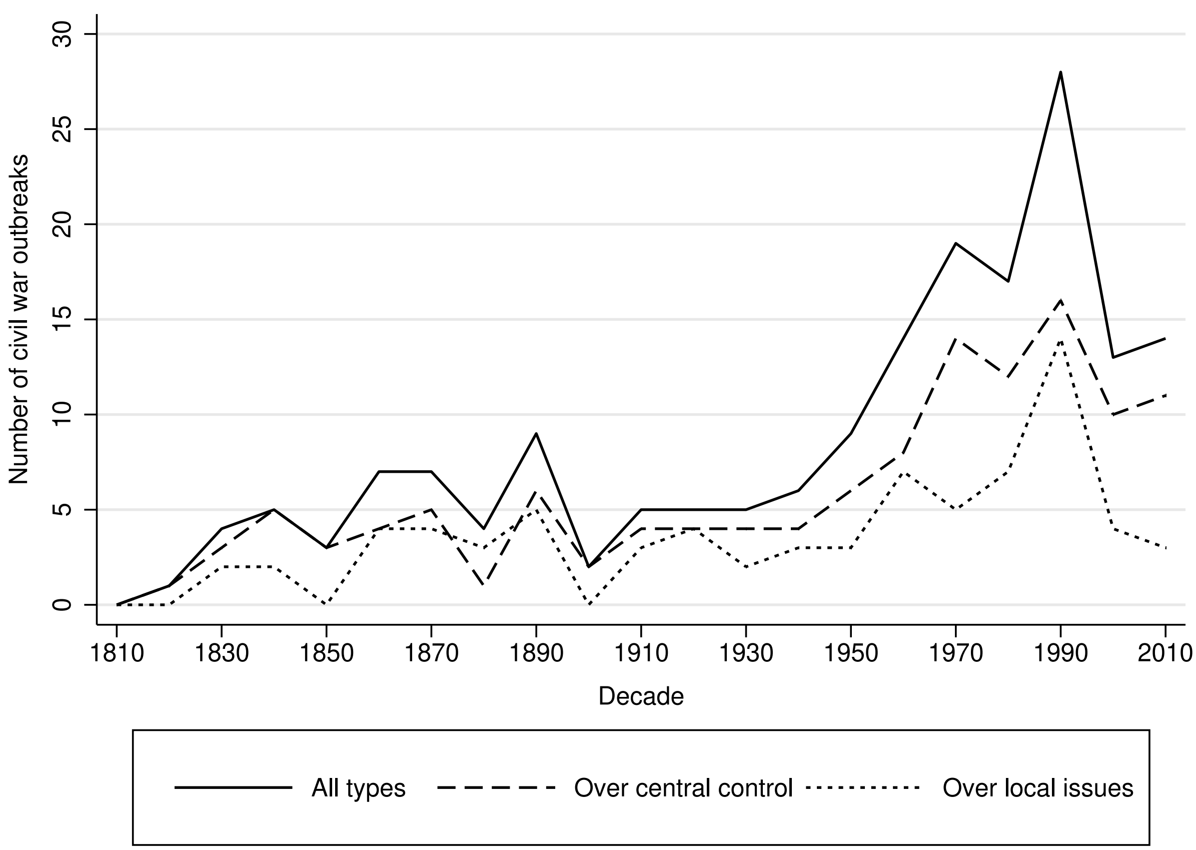
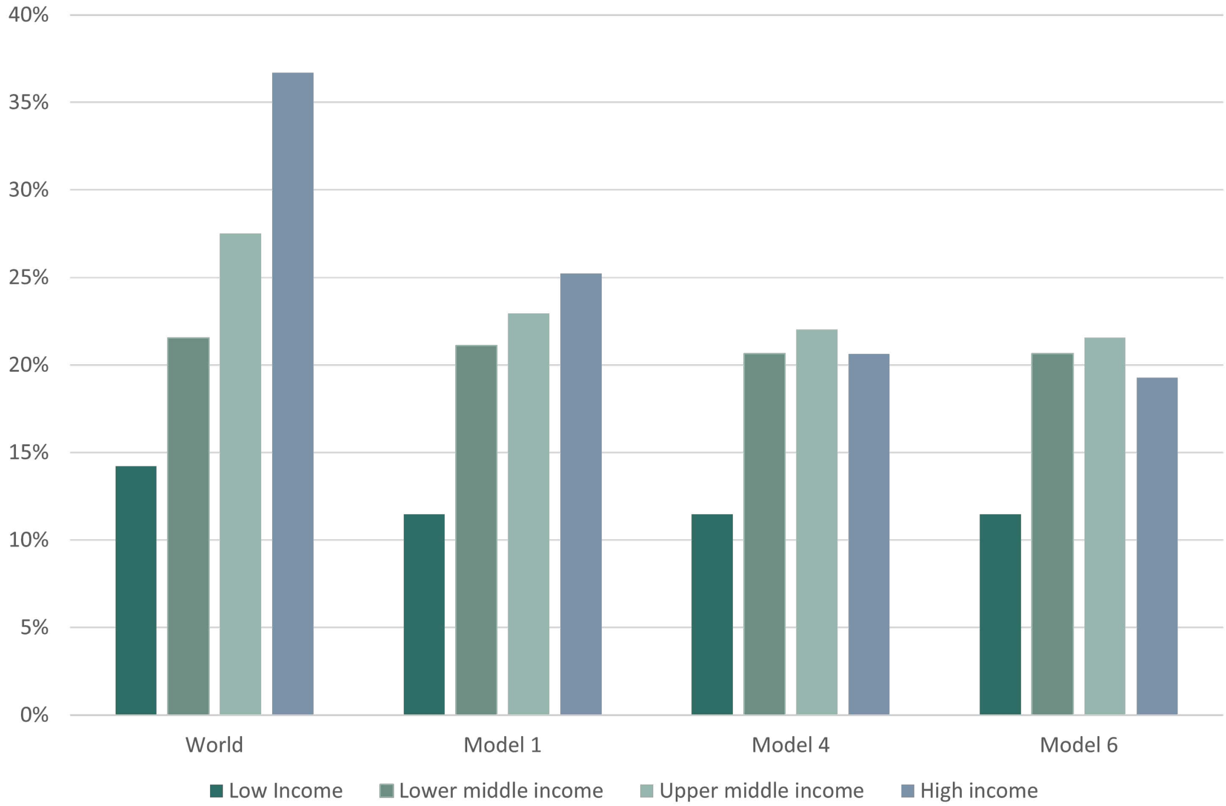


Figure B.9: Data selection



Notes. The graph shows the data selection in the models of Table B.6 by income group as defined by the World Bank, and is derived from the World Development Indicators. We refer to the following classification of income groups: low-income $1045 or less, lower middle-income $1046–$4,095, upper middle income: $4,096–$12,695, and high-income $12,695 or higher. We compare our models to the world's real income distribution.

Table B.1: Comparison between income Gini and height Gini: regressions based on selected different datasets

	(1) van Zanden et al. 2014b	(2) WID	(3) Top 10% WID
Height Gini	0.23***	0.16*	0.22***
	(0.082)	(0.084)	(0.067)
Constant	41.81***	49.79***	39.36***
	(5.703)	(5.843)	(4.528)
Observations	221	184	164
Number of countries	79	69	69
Adj. R-squared	0.1763	0.2513	0.1359
Time Fixed Effects	Y	Y	Y

Notes. Income Ginis are derived from different databases and sources, namely the World Inequality Database (WID) and van Zanden et al., (2014b). As income Ginis are just available for a limited number of countries, the top 10% income share from the WID is included and multiplied by 100. Time-fixed effects are included in all models.

Table B.2: Correlation between income Ginis and top 1% income share from WID on a yearly basis

	Income Gini
Top 1% income share	1.46***
	(0.092)
Constant	0.21***
	(1.125)
Observations	1,498
Number of countries	39
R-squared	0.460
Time Fixed Effects	N
Region Fixed Effects	N

Notes. Random effect model. Standard error in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. Data for the top 1% income share and the Gini coefficients for post-tax income are from the World Inequality Database (WID).

Table B.3: Overview of our joint inequality index – time and region coverage as a percentage of the world region's population

World Region	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
East Asia	89	89	90	90	90	89	88	98	99	99	99	97	14	99	100	100	100	100	100	100	100
East Europe & Central Asia	15	35	35	51	52	57	65	70	58	47	51	49	64	65	60	66	63	88	98	98	100
Latin America and the Caribbeans	19	31	32	60	64	63	71	75	73	72	71	74	27	40	89	91	89	96	97	100	99
Middle East & North Africa						24	40	77	23	28	23	7	53	19	59	79	76	78	100	100	100
North America	92	100	100	100	100	100	91	92	93	93	93	100	100	100	100	100	100	100	100	100	100
Oceania					32	79	54	75	62	64	65	86	94	85	83	85	84	80	96	99	100
South Asia				83	98	86	97	97	92	10	83	82	90	89	89	100	100	100	100	100	100
Southeast Asia				50	59	70	65	66	58	67	54	64	47	1	36	52	96	97	100	100	100
Sub-Saharan Africa	36	28	2	2			13	8	31	19	48	58	39	47	80	95	95	95	100	100	100
Western Europe	79	85	71	79	71	71	79	91	80	70	69	64	83	84	82	100	100	100	100	100	100

Notes. Years refer to the beginning of a birth decade (1810 for 1810–1819). For the birth decade of 1870–79, for example, we have evidence for 53% of Africa's population.

Table B.4: Correlation between our joint inequality index and income Ginis from different sources

	Joint inequality index
Income Gini Milanovic, 2013	0.66***
Income Gini OECD	0.97***
Income Gini van Zanden et al., 2014b	0.57***
Income Gini World Inequality Database (WID)	0.63***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: Summary statistics

	Obs	Mean	Std. Dev.	Min	Max
Civil war	1190	0.11	0.32	0	1
Inequality (lag)	1190	0.42	0.09	.1723938	.8133366
Population (log)	1183	15.83	1.56	11.30382	20.95647
Democracy	874	0.01	0.07	-.1	.1
Democracy^2	874	0.47	0.36	0	1
Diamond	1190	0.17	0.38	0	1
History of wars	1190	0.22	0.74	0	6
Colony	1190	0.13	0.34	0	1
Ethn. frac.	1141	0.43	0.27	0	.9302
Height growth	657	0.00	0.02	-.084287	.1259581
GDP p.c.	934	7618.98	9558.36	485.735	77798.96

Notes. All variables are measured using country-decade units. Civil war indicates a civil war outbreak per country and decade.

Table B.6: Regression of civil war onset

	(1) PLOG	(2) XTLOG	(3) XTLOG	(4) XTLOG	(5) XTLOG	(6) XTLOG	(7) XTLOG	(8) XTLOG	(9) RELOG
Inequality	0.68***	0.54***	0.51***	0.51***	0.50***	0.50***	0.89***	0.44**	5.58***
	(0.095)	(0.111)	(0.163)	(0.163)	(0.158)	(0.159)	(0.307)	(0.176)	(1.218)
Population (log)	0.04***	0.04***	0.05***	0.05***	0.04***	0.05***	0.06***	0.04***	0.34***
	(0.005)	(0.008)	(0.011)	(0.012)	(0.011)	(0.012)	(0.020)	(0.014)	(0.076)
Democracy			−0.45*	−0.45*	−0.38	−0.38	−0.52	−0.23	−6.10***
			(0.262)	(0.262)	(0.264)	(0.263)	(0.506)	(0.269)	(2.025)
Democracy^2			−0.27***	−0.27***	−0.26***	−0.26***	−0.31***	−0.23***	−2.19***
			(0.046)	(0.046)	(0.045)	(0.046)	(0.079)	(0.052)	(0.388)
Diamond				0.01		−0.00		0.03	
				(0.049)		(0.046)		(0.050)	
History of wars					0.01	0.01	0.02		0.39***
					(0.011)	(0.011)	(0.112)		(0.105)
Colony						0.01	0.01		0.21
						(0.058)	(0.073)		(0.483)
Ethn. frac.					0.09	0.09		0.11	
					(0.099)	(0.098)		(0.101)	
Height growth							−0.90		
							(1.597)		
GDP p.c.								−0.20***	
								(0.066)	
Constant									−9.04***
									(1.559)
Observations	1183	1183	850	850	827	827	447	753	873
Time Fixed Effects	N	Y	Y	Y	Y	Y	Y	Y	N
Region Fixed Effects	N	N	Y	Y	Y	Y	Y	Y	N

Notes. We used Pooled Logit (PLOG), Panel Logit (XTLOG) and Rare Events Logit Models (RELOG). Clustered standard errors in parentheses, ***,**,* significant on the 1, 5, and 10%-level, respectively. Marginal effects reported, except for model (9). Inequality is composed of income Gini, height Gini and land Ginis which are all lagged by 1 decade. For expository purposes, democracy and democracy squared are divided by 100 before running the regressions. Diamond and colonial history are dummy variables. Fractionalisation measures are time-invariant. GDP per capita is divided by 10,000 before running the regression.

Table B.7: OLS regression of civil war onset

	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) Pooled OLS	(5) Random effect model
Inequality	0.69***	0.55***	0.45**	0.45**	0.85***
	(0.128)	(0.140)	(0.181)	(0.175)	(0.305)
Population (log)	0.04***	0.03***	0.04***	0.03***	0.05***
	(0.008)	(0.009)	(0.010)	(0.011)	(0.019)
Democracy		−0.13	0.11	0.04	0.45
		(0.207)	(0.216)	(0.228)	(0.381)
Democracy^2		−0.17***	−0.15***	−0.14***	−0.16***
		(0.037)	(0.041)	(0.046)	(0.061)
Diamond		0.02		−0.01	
		(0.040)		(0.050)	
History of wars		0.09***		0.09***	
		(0.017)		(0.018)	
Colony		−0.02	−0.03		−0.01
		(0.066)	(0.063)		(0.073)
Ethn. frac.			0.10	0.09	
			(0.091)	(0.078)	
GDP p.c.				0.02	
				(0.013)	
Height growth					−1.03
					(1.998)
Constant	−0.85***	−0.58***	−0.83***	−0.66***	−0.99**
	(0.147)	(0.177)	(0.203)	(0.240)	(0.452)
Observations	1183	874	851	790	460
R-squared	0.091	0.173	0.159	0.200	0.181
Time Fixed Effects	Y	Y	Y	Y	Y

Notes. Heteroscedasticity-robust clustered standard errors in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. Inequality is composed of income Gini, height Gini and land Ginis which are all lagged by 1 decade. For expository purposes, democracy and democracy squared are divided by 100 before running the regressions. Diamond and colonial history are dummy variables. Fractionalisation measures are time-invariant. GDP per capita is divided by 10,000 before running the regression.

Table B.8: OLS regression with region and country fixed effects

	(1) Pooled OLS	(2) Pooled OLS
Inequality	0.52***	0.50***
	(0.141)	(0.157)
Population (log)	0.05***	0.01
	(0.008)	(0.045)
Constant	−0.92***	−0.19
	(0.155)	(0.707)
Observations	1183	1183
R-squared	0.112	0.319
Time Fixed Effects	Y	Y
Region Fixed Effects	Y	N
Country Fixed Effects	N	Y

Notes. Heteroscedasticity-robust clustered standard errors in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. Inequality is composed of income Ginis, height Ginis and land Ginis and lagged by one decade.

Table B.9: OLS regression by different civil war types and measures

	(1) Civil war over central control	(2) Civil war over local issues	(3) Civil war severity
Inequality	0.23**	0.29**	14.16**
	(0.115)	(0.148)	(5.988)
Population (log)	0.02**	0.03***	1.69**
	(0.007)	(0.008)	(0.674)
Democracy	−0.02	−0.03	−2.16
	(0.186)	(0.142)	(10.535)
Democracy^2	−0.13***	0.00	−2.73
	(0.034)	(0.025)	(2.201)
History of wars (by type)	0.10***	0.06**	2.56***
	(0.024)	(0.026)	(0.941)
Constant	−0.33**	−0.56***	−35.59**
	(0.135)	(0.166)	(14.018)
Observations	874	874	874
R-squared	0.140	0.158	0.086
Time Fixed Effects	Y	Y	Y
Region Fixed Effects	Y	Y	Y

Notes. Heteroscedasticity-robust clustered standard errors in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. Inequality is composed of income Ginis, height Ginis and land Ginis, which are all lagged by one decade. For expository purposes, democracy and democracy squared are divided by 100 before running the regressions.

Table B.10: Robustness check: regression for low- and high-income countries

	(1) XTLOG	(2) XTLOG
Omitted	*GDP p.c. < 1,036 USD*	*GDP p.c. > 12,535 USD*
Inequality	0.44***	0.71***
	(0.153)	(0.199)
Population (log)	0.04***	0.04***
	(0.012)	(0.014)
Democracy	−0.46**	−0.37
	(0.218)	(0.313)
Democracy^2	−0.18***	−0.25***
	(0.051)	(0.060)
Diamond	−0.04	−0.02
	(0.043)	(0.060)
Colony	−0.03	−0.01
	(0.095)	(0.059)
Observations	684	659
Time Fixed Effects	Y	Y
Region Fixed Effects	Y	Y

Notes. We used Panel Logit (XTLOG) Models. Clustered standard error by country in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. Marginal effects reported. Inequality is composed of income Ginis, height Ginis and land Ginis, which are all lagged by one decade. For expository purposes, democracy and democracy squared are divided by 100 before running the regressions. Diamond and colonial history are dummy variables.

Table B.11: Instrumental variable regression

	(1) 2SLS	(2) 2SLS	(3) LIML	(4) LIML	(5) 2SLS
First stage					
WheatRiceSugar	0.12***	0.08***	0.08***	0.08***	0.10***
	(0.010)	(0.023)	(0.023)	(0.021)	(0.012)
Second stage					
Inequality	3.31***	5.35***	5.30***	5.36***	2.68***
	(0.248)	(1.662)	(1.623)	(1.676)	(0.334)
Population (log)	0.07***	0.06***	0.06***	0.06***	0.08***
	(0.014)	(0.021)	(0.021)	(0.022)	(0.018)
Democracy	0.64**	0.97*	0.96**	0.97*	
	(0.249)	(0.498)	(0.476)	(0.497)	
Democracy^2	−0.06	−0.14*	−0.14*	−0.14*	
	(0.051)	(0.084)	(0.081)	(0.084)	
Diamond	−0.03	0.01		0.01	
	(0.061)	(0.082)		(0.082)	
History of wars	0.04**				
	(0.019)				
Colony	−0.05			0.02	
	(0.073)			(0.090)	
Exports					−0.06
					(0.136)
Constant	−2.45***	−2.68***	−2.67***	−2.69***	−2.21***
	(0.247)	(0.749)	(0.742)	(0.760)	(0.276)
Observations	844	844	844	844	328
Adj. R-squared	0.218	0.347	0.347	0.346	0.089
Time Fixed Effects	Y	Y	Y	Y	Y
Region Fixed Effects	N	Y	Y	Y	N
F-statistic	153.58	14.02	14.84	13.82	82.33
Kleinbergen-Paap rk LM statistic	Exactly identified				
Hansen J statistic	Exactly identified				

Notes. Heteroscedasticity- and country-clustered robust standard errors in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. The dependent variable in the first stage is inequality. The dependent variable in the second stage is civil war onset. Inequality is composed of income Ginis, height Ginis and land Ginis, which are all lagged by one decade. For expository purposes, democracy and democracy squared are divided by 100 before running the regressions. Diamond and colonial history are dummy variables.

B.7 Appendix

B.7.1 Variable definitions

Civil War. We use the outbreak of civil war as dependent variable. Civil war is coded as a dummy variable, which takes the value one if a civil war outbreak occurred in the regarded country and decade. Our data is derived from the Correlates of War Project (COW). The COW defines a threshold of 1,000 conflict-related deaths per conflict and year to be classified as civil war. The COW further differentiates three types of intra-state conflicts based on the conflict sides involved. Civil wars and regional internal wars both include the government and a non-state entity, whereas the government on the regional level is included in the latter. Civil wars are further split into two types: conflict for control over the central government, as well as fights over local issues. As an alternative measure for civil wars, we include the average number of civil war deaths for the respective decade and country as a measure of the severity of a conflict. Source: Sarkees (2010).

Colony. We control for colonial history, where one indicates that a country was a former colony and zero if not. Source: Correlates of War Project. Colonial Contiguity Data, 1816—2016. Version 3.1.

Democracy. The quality of institutions is measured by the polity2 index. This variable ranges from -10, indicating a fully autocratic regime, to +10, which reflects a highly consolidated democracy. We use democracy and democracy squared in our regressions. Source: Polity5 Project, Marshall and Gurr (2020).

Diamond. Coded as a dummy variable that takes the value one if a diamond deposit is/was present in a country and zero otherwise.

Ethnic Fractionalization. Composed of the index of racial and linguistic characteristics to measure the ethnic fractionalization within a country. Source: Alesina et al. (2003).

Exports. Share of raw materials and mining products divided by the number of total exports. Source: World Bank Data (1999) (CD-Rom).

GDP p.c. GDP per Capita. Source: Bolt and Van Zanden (2020).

Height growth. Indicates the growth of height from one period to the period $t + 1$. Source: Baten and Blum (2015), available via Clio Infra.

Inequality. Composed by income Gini, height Gini, and land Gini where data is available.

Population (log). The natural logarithm of a country's population at the beginning of a decade. Source: Fink-Jensen (2015).

History of Wars. Indicates whether a civil war occurred in the previous period. It counts the number of decades. Source: Sarkees (2010).

B.7.2 The construction of multidimensional inequality

B.7.3 Health inequality

Our dataset on height Ginis is partly based on the data collection of heights by Baten and Blum (2011). We substantially extend this dataset with data from different surveys and individual height studies. In Table B.12 we provide an overview of the updated data for the countries and time periods covered, and sources used. For the data used from the dataset of Baten and Blum (2011), an overview of the data sources can be found on the website of Clio Infra.[26] We did remove the following data from our figures and analysis as they represent outliers in our dataset of height Ginis: the values for Albania in the birth decade of 1900 and for Chile in the birth decade of 1910. The study on Albania is based only on a few measurements from a narrowly defined mountain region, published by Coon (1950). We can therefore omit this extreme value for the decade 1900 (1900–1909) with good reason. Baten and Llorca-Jaña (2021) provide important new estimates for Chile, which we include in our dataset. However, as they themselves note, the estimates of the size of inequality for Chile in the 1910 birth decade differ considerably from other inequality estimates. We therefore remove this value from our sample.

B.7.4 Land inequality

Our data on land inequality is based on data from Frankema (2005, 2010), which is then derived from the country-specific data collection from the Food and Agriculture Organization (FAO). FAO provides data on the total agricultural population and divides it by the total number of land holdings (FAO, 2019). We update this dataset with the data from the latest 2019 FAO report.

[26]Clio Infra, https://clio-infra.eu/Indicators/HeightGini.html.

We calculate the Gini coefficients for 91 new observations based on the formula from Frankema (2005):

$$\textit{Gini coefficient} = \frac{(\sum_{(j=1)} \sum_{(k=1)} n_j n_k |y_j - y_k|)}{(2n^2 \cdot \frac{1}{n})},$$

where $n = \textit{amount of decile shares} = 10$.

Since land inequality does not change significantly over time, Baten and Juif (2014) have suggested some adjustments to interpolate between two given data points. They calculate the impact of a successfully implemented land reform on the value of the land Gini. They estimate a decrease in land inequality after a land reform by -5.57 Gini points. However, if a land reform is unsuccessfully carried out, then land inequality remains unchanged. Following the method of Baten and Juif (2014), we replicate the effect of land reform and obtain an estimated average effect of -4.47 Gini points (Table B.14). To do so, we extend the collection of land reforms from Baten and Juif (2014) with data for land reforms from WCARRD (1988). The reduced average effect of a land reform in comparison with that of Baten and Juif (2014) may be explained by our higher country and time coverage, including recent years from 2000–2010, where land reforms may not have the same effect on land inequality as it had during the 1900s. Based on the estimated average impact of land reform on land inequality, we adjust our dataset by subtracting 4.47 Gini points for the following period from the land Gini coefficient if a land reform was successful. In this manner, we can gain several observations for our analysis.

B.7.5 Joint inequality index

We construct our inequality variable by combing our three indicators, income, health and land inequality into one joint inequality index (Table B.13 report our sources for income Ginis). We used the common method of assigning equal weight to each dimension. We hereby follow a normative approach. This is also done to construct the Human Development Index (HDI), which takes the geometric mean of three dimensions to construct a single indicator, the Gender-related Development Index (UNDP, 2013), and this approach is also used by Alkire and Foster (2010) on the construction of a multidimensional poverty index. In Table B.15 we include sensitivity analyses for our joint inequality index due to different construction decisions and weighting. Our results remain robust. In model (1) we construct a joint inequality index for those birth decades and countries, where we do have data for all three indicators height Gini, income Gini and land Gini without any weighting. In model (2) we do not weight height Gini by the degree of urbanisation. In model (3)–(4) we control for different weighting of the three components rather than following a normative approach, which is $\alpha_1 = 0.5$, $\alpha_2 = 0.3$ and $\alpha_3 = 0.2$ in the case of model (3) and $\alpha_1 = 0.2$, $\alpha_2 = 0.4$ and $\alpha_3 = 0.4$ in case of model (4). Model (5) excludes the observations for income Ginis, which we calculated based on the top 1% income share from the WID. Finally, model (6) displays the regression results if we use health inequality as the only indicator for inequality.

B.7.6 References

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic growth*, *8*, 155–194.

Alkire, S., & Foster, J. E. (2010). Designing the inequality-adjusted human development index.

Baten, J., & Blum, M. (2015). Height gini. http://hdl.handle.net/10622/HKZPR9

Baten, J., & Blum, M. (2011). Anthropometric within-country inequality and the estimation of skill premia with anthropometric indicators. *Review of Economics*, *62*(2), 107–138.

Baten, J., & Juif, D. (2014). A story of large landowners and math skills: Inequality and human capital formation in long-run development, 1820–2000. *Journal of Comparative Economics*, *42*(2), 375–401.

Baten, J., & Llorca-Jaña, M. (2021). Inequality, low-intensity immigration and human capital formation in the regions of chile, 1820-1939. *Economics & Human Biology*, *43*, 101030.

Bolt, J., & Van Zanden, J. L. (2020). Maddison style estimates of the evolution of the world economy. a new 2020 update. *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands*.

Coon, C. S. (1950). *The mountains of giants: A racial and cultural study of the north albanian mountain ghegs*. Peabody Museum Press. https://iiif.lib.harvard.edu/manifests/view/drs:53772366%5C$1i.

Fink-Jensen, J. (2015). Total population. http://hdl.handle.net/10622/SNETZV

Food and Agriculture Organization. (2019). *Main results and metadata by country (2006–2015). world programme for the census of agriculture 2010.*

Frankema, E. H. (2005). *The colonial origins of inequality: Exploring the causes and consequences of land distribution* (tech. rep.). IAI Discussion Papers.

Frankema, E. H. (2010). The colonial roots of land inequality: Geography, factor endowments, or institutions? *The Economic History Review*, *63*(2), 418–451.

López-Alonso, M., & Condey, R. P. (2003). The ups and downs of mexican economic growth: The biological standard of living and inequality, 1870–1950. *Economics & Human Biology*, *1*(2), 169–186.

Marshall, M. G., & Gurr, T. R. (2020). Polity5: Political regime characteristics and transitions, 1800-2018. *Center for Systemic Peace*.

Sarkees, M. R. (2010). The cow typology of war: Defining and categorizing wars (version 4 of the data). *Note with version 4 of the Correlates of War Data*.

World Conference on Agrarian Reform and Rural Development. (1988). Guidelines on socio-economic indicators for monitoring and evaluating agrarian reform and rural development.

B.7.7 Tables and figures

Table B.12: Additional sources of height Gini

Notes. For the other sources of height Ginis please see Baten and Blum (2011) and clio-infra.eu.

Country	ccode	Birth Decade	Source
Algeria	dz	1950–1990	STEPS
Austria	at	1970–1980	ESS 2014, Round 7
Belgium	be	1960–1970; 1990	ESS 2014, Round 7
Benin	bj	1960–2000	Demographic and Health Survey
Bolivia	bo	1880–1920	Peres-Cajias et al., 2020
		1950–1990	Demographic and Health Survey
Botswana	bw	1950–1990	STEPS
Brazil	br	1950–1970	Demographic and Health Survey
Burkina Faso	bf	1960–1990	Demographic and Health Survey
Burundi	bi	1960–1990	Demographic and Health Survey
Cameroon	cm	1960–1990	Demographic and Health Survey
Cape Verde	cv	1950–1980	STEPS
Chad	td	1960–1990	Demographic and Health Survey
Chile	cl	1820–1900; 1930–1980	Baten and Llorca-Jaña, 2021
China	cn	1940	China Health and Nutrition Surveys, Wave of 1989
		1960–1980	EASS, 2010
Comoros	km	1960–1990	Demographic and Health Survey
Cyprus	cy	1860–1890	Buxton, 1920
Czech Republic	cz	1960–1990	ESS 2014, Round 7
Democratic Republic of the Congo	cd	1960–1990	Demographic and Health Survey
Denmark	dk	1980	ESS 2014, Round 7
Dominican Republic	do	1940–1990	Demographic and Health Survey
Estonia	ee	1960; 1980	ESS 2014, Round 7
Ethiopia	et	1950–1990	Demographic and Health Survey
Finland	fi	1970–1980	ESS 2014, Round 7
France	fr	1960	Pineau, 1993
		1970	Olivier, 1991
		1980	ESS 2014, Round 7
Gabon	ga	1960–1990	Demographic and Health Survey

Table B.12: (continued)

Country	ccode	Birth Decade	Source
Gambia	gm	1960–1990	Demographic and Health Survey
Germany	de	1960–1990	ESS 2014, Round 7
Ghana	gh	1960–1990	Demographic and Health Survey
Greece	gr	1880; 1930; 1950–1960	Capocasa et al., 2019
Guatemala	gt	1960–1990	Demographic and Health Survey
Guinea	gn	1960–2000	Demographic and Health Survey
Guyana	gy	1960–1990	Demographic and Health Survey
Haiti	ht	1960–1990	Demographic and Health Survey
Honduras	hn	1960–1990	Demographic and Health Survey
Hungary	hu	1960	Gyenis and Joubert, 2004
		1970–1980	ESS 2014, Round 7
Iraq	iq	1960–1990	STEPS
Ireland	ie	1960–1980	ESS 2014, Round 7
Israel	il	1960–1990	ESS 2014, Round 7
Ivory Coast	ci	1960–1990	Demographic and Health Survey
Kenya	ke	1960–1990	Demographic and Health Survey
Lesotho	ls	1960–1990	Demographic and Health Survey
Liberia	lr	1970–1990	Demographic and Health Survey
Lithuania	lt	1960–1990	ESS 2014, Round 7
Malawi	mw	1960–1990	Demographic and Health Survey
Mali	ml	1960–2000	Demographic and Health Surveys
Mexico	mx	1900–1920	López-Alonso and Condey, 2003
Mozambique	mz	1960–1990	Demographic and Health Survey
Myanmar	mm	1960–1980	STEPS
Namibia	na	1950–1990	Demographic and Health Survey
Nepal	np	1960–1990	STEPS
Netherlands	nl	1810–1920	Kees Mandemakers, HSN dataset Heights and Life Courses, 2018_02
		1960–1970	ESS 2014, Round 7
Nicaragua	ni	1950–1980	Demographic and Health Survey
Niger	ne	1960–1990	Demographic and Health Survey
Nigeria	ng	1960–2000	Demographic and Health Survey
Norway	no	1960–1970	ESS 2014, Round 7
Palestine	ps	1940–1970	Abdeen et al., 2000
Peru	pe	1820–1880	Clio Infra

Table B.12: (continued)

Country	ccode	Birth Decade	Source
		1960–1990	Demographic and Health Survey
Poland	pl	1970–1990	ESS 2014, Round 7
Puerto Rico	pr	1890–1910	Godoy, 2007 EHB
		1920–1930	Thieme, 1959
		1980	Hossain et al., 2005
Republic of the Congo	cg	1960–1990	Demographic and Health Survey
Russia	ru	1850–1880	Mironov and Freeze, 2012
Rwanda	rw	1960–1990	Demographic and Health Survey
Senegal	sn	1960–1990	Demographic and Health Survey
Sierra Leone	sl	1960–1990	Demographic and Health Survey
South Africa	za	1960–1990	Demographic and Health Survey
South Korea	kr	1890–1910	Choi, 2020
		1960–1980	EASS 2010
Spain	es	1960–1990	ESS 2014, Round 7
Sudan	sd	1960–1990	STEPS
Sweden	se	1970–1990	ESS 2014, Round 7
Switzerland	ch	1940–1990	Koepke et al., 2018
Taiwan	tw	1960–1980	EASS 2010
Tanzania	tz	1960–1990	Demographic and Health Survey
Togo	tg	1960–1990	Demographic and Health Survey
Uganda	ug	1960–1990	Demographic and Health Survey
United Kingdom	uk	1960–1980	ESS 2014
United States	us	1970	BRFSS Annual Survey Data 1995
Vietnam	vn	1950–1990	STEPS
Zambia	zm	1960–1990	Demographic and Health Survey
Zimbabwe	zw	1960–1990	Demographic and Health Survey

Table B.13: Sources for income Gini coefficients

Country	ccode	Birth Decade	Source
Afghanistan	af	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Albania	al	1990–2010	WIID, 2021
Algeria	dz	1980–2010	WIID, 2021
		2000	WID, 2021, calc. top 1% income share
		2010	WIID, 2021
Andorra	ad	2000–2010	WIID, 2021
Angola	ao	1990	WID, 2021, calc. top 1% income share
Angola	ao	2000–2010	WIID, 2021
Argentina	ar	1930–1940	WID, 2021, calc. top 1% income share
		1950–2010	WIID, 2021
Armenia	am	1990–2010	WIID, 2021
Australia	au	1910–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
Austria	at	1980–2010	WIID, 2021
Azerbaijan	az	1990–2010	WIID, 2021
Bahamas	bs	1970	WIID, 2021
		2000–2010	WIID, 2021
Bahrain	bh	1990–2010	WID, 2021, calc. top 1% income share
Bangladesh	bd	1960–2010	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Barbados	bb	1950	WIID, 2021
		1970	WIID, 2021
		2010	WIID, 2021
Belarus	by	1980–2010	WIID, 2021
Belgium	be	1970–2010	WIID, 2021
Belize	bz	1990	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Benin	bj	1950	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Bhutan	bt	1990	WID, 2021, calc. top 1% income share
Bhutan		2000–2010	WIID, 2021
Bolivia	bo	1960	WIID, 2021
		1980–2010	WIID, 2021

Table B.13: (continued)

Country	ccode	Decade	Source
Bosnia and Herzegovina	ba	1980–1990	WID, 2021
		2000–2010	WIID, 2021
Botswana	bw	1980–2010	WIID, 2021
Brazil	br	1970–2010	WIID, 2021
Brunei	bn	2000–2010	WIID, 2021
Bulgaria	bg	1960–2010	WIID, 2021
Burkina Faso	bf	1990–2010	WIID, 2021
Burundi	bi	1990–2010	WIID, 2021
Cambodia	kh	1990–2010	WIID, 2021
Cameroon	cm	1990–2010	WIID, 2021
Canada	ca	1920–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
Cape Verde	cv	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Central African Republic	cf	1990–2000	WIID, 2021
		2010	WID, 2021, calc. top 1% income share
Chad	td	1950	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Chile	cl	1960	WIID, 2021
		1980–2010	WIID, 2021
China	cn	1950–2010	WIID, 2021
Colombia	co	1960	WIID, 2021
		1990–2010	WIID, 2021
Comoros	km	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Costa Rica	cr	1960	WIID, 2021
		1980–2010	WIID, 2021
Croatia	hr	1980–2010	WIID, 2021
Cuba	cu	1950	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Cyprus	cy	1990	WID, 2021
		2000–2010	WIID, 2021
Czech Republic	cz	1950–2010	WIID, 2021
Democratic Republic of the Congo	cd	1990	WID, 2021, calc. top 1% income share

Table B.13: (continued)

Country	ccode	Decade	Source
		2000–2010	WIID, 2021
Denmark	dk	1870	WID, 2021, calc. top 1% income share
		1900–1960	WID, 2021, calc. top 1% income share
		1970–2010	WIID, 2021
Djibouti	dj	1990–2010	WIID, 2021
Dominica	dm	2000	WIID, 2021
Dominican Republic	do	1960	WIID, 2021
		1980–2010	WIID, 2021
East Timor	tl	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Ecuador	ec	1960	WIID, 2021
		1980–2010	WIID, 2021
Egypt	eg	1960–2010	WIID, 2021
		1990–2010	WIID, 2021
El Salvador	sv	1960–2010	WIID, 2021
		1990–2010	WIID, 2021
Equatorial Guinea	gq	1990	WID, 2021, calc. top 1% income share
		2000	WIID, 2021
		2010	WID, 2021, calc. top 1% income share
Eritrea	er	1990	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Estonia	ee	1980	WID, 2021
		1990–2010	WIID, 2021
Ethiopia	et	1980	WID, 2021, calc. top 1% income share
		1990–2010	WIID, 2021
Fiji	fj	1960–2010	WIID, 2021
		1990–2010	WIID, 2021
Finland	fi	1920–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
France	fr	1900–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
Gabon	ga	1970	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Gambia	gm	1990–2010	WIID, 2021
Georgia	ge	1980	WID, 2021, calc. top 1% income share

Table B.13: (continued)

Country	ccode	Decade	Source
		1990–2010	WIID, 2021
Germany	de	1870–1960	WID, 2021, calc. top 1% income share
		1970–2010	WIID, 2021
Ghana	gh	1980–2010	WIID, 2021
Greece	gr	1950	WIID, 2021
		1960–1970	WID, 2021, calc. top 1% income share
		1980–2010	WIID, 2021
Greenland	gl	2000–2010	WIID, 2021
Grenada	gd	2000	WIID, 2021
Guatemala	gt	1970–2010	WIID, 2021
		2000–2010	WIID, 2021
Guinea	gn	1990–2010	WIID, 2021
Guinea-Bissau	gw	1990–2010	WIID, 2021
Guyana	gy	1990	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Haiti	ht	2000–2010	WIID, 2021
Honduras	hn	1960	WIID, 2021
		1980–2010	WIID, 2021
Hong Kong	hk	1960–2010	WIID, 2021
Hungary	hu	1920–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
Iceland	is	1990	WID, 2021
		2000–2010	WIID, 2021
India	in	1920–1940	WID, 2021, calc. top 1% income share
		1950–2010	WIID, 2021
Indonesia	id	1920–1930	WID, 2021, calc. top 1% income share
		1970–2010	WIID, 2021
Iran	ir	1980–2010	WIID, 2021
Iraq	iq	1950	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Ireland	ie	1930–1940	WID, 2021, calc. top 1% income share
		1970–2010	WIID, 2021
Israel	il	1980–2010	WIID, 2021
Italy	it	1940	WIID, 2021
		1960–2010	WIID, 2021

Table B.13: (continued)

Country	ccode	Decade	Source
Ivory Coast	ci	1950	WIID, 2021
		1980–2010	WIID, 2021
Jamaica	jm	1950	WIID, 2021
		1970–2010	WIID, 2021
Japan	jp	1880–1990	WID, 2021, calc. top 1% income share
		1950–2010	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Jordan	jo	1980–2010	WIID, 2021
Kazakhstan	kz	1990–2010	WIID, 2021
Kenya	ke	1970	WIID, 2021
		1990–2010	WIID, 2021
Kiribati	ki	2000	WIID, 2021
Kosovo	xk	2000–2010	WIID, 2021
Kuwait	kw	1970–2000	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000	WIID, 2021
		2010	WID, 2021, calc. top 1% income share
Kyrgyzstan	kg	1990–2010	WIID, 2021
Laos	la	1990–2010	WIID, 2021
Latvia	lv	1980	WID, 2021
		1990–2010	WIID, 2021
Lebanon	lb	1960	WIID, 2021
		1990–2000	WID, 2021, calc. top 1% income share
		2010	WIID, 2021
Lesotho	ls	1980–2010	WIID, 2021
Liberia	lr	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Libya	ly	1990–2010	WID, 2021, calc. top 1% income share
Lithuania	lt	1980	WID, 2021
		1990–2010	WIID, 2021
Luxembourg	lu	1980–2010	WIID, 2021
Macau	mo	1990–2010	WID, 2021, calc. top 1% income share
Macedonia	mk	1980	WID, 2021
		1990–2010	WIID, 2021
Madagascar	mg	1960	WIID, 2021

Table B.13: (continued)

Country	ccode	Decade	Source
		1980–2010	WIID, 2021
Malawi	mw	1960–2010	WIID, 2021
Malaysia	my	1960 -2010	WIID, 2021
Maldives	mv	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Mali	ml	1980–2010	WIID, 2021
Malta	mt	2000 -2010	WIID, 2021
Mauritania	mr	1980–1990	WIID, 2021
Mauritius	mu	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Mexico	mx	1950–2010	WIID, 2021
Micronesia	fm	2000–2010	WIID, 2021
Moldova	md	1980	WID, 2021
		1990–2010	WIID, 2021
Mongolia	mn	1990–2010	WIID, 2021
Montenegro	me	1980–1990	WID, 2021
		2000–2010	WIID, 2021
Morocco	ma	1960	WIID, 2021
		1980–2010	WIID, 2021
Mozambique	mz	1990–2010	WIID, 2021
Myanmar	mm	1950	WIID, 2021
		1990–2000	WID, 2021, calc. top 1% income share
		2010	WIID, 2021
Namibia	na	1990–2010	WIID, 2021
Nauru	nr	2010	WIID, 2021
Nepal	np	1970	WIID, 2021
		1990–2010	WIID, 2021
Netherlands	nl	1910–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
New Zealand	nz	1920–1960	WID, 2021, calc. top 1% income share
		1970–2010	WIID, 2021
Nicaragua	ni	1990–2010	WIID, 2021
Niger	ne	1960	WIID, 2021
		1990–2010	WIID, 2021
Nigeria	ng	1950	WIID, 2021
		1980–2010	WIID, 2021

Table B.13: (continued)

Country	ccode	Decade	Source
North Korea	kp	1990–2010	WID, 2021, calc. top 1% income share
Norway	no	1870–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
Oman	om	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Pakistan	pk	1960–2010	WIID, 2021
Palau	pw	2010	WIID, 2021
Palestine	ps	1990–2010	WIID, 2021
Panama	pa	1960–2010	WIID, 2021
Papua New Guinea	pg	1990	WIID, 2021
		2000	WID, 2021, calc. top 1% income share
		2010	WIID, 2021
Paraguay	py	1980–2010	WIID, 2021
Peru	pe	1970	WIID, 2021
		1990–2010	WIID, 2021
Philippines	ph	1950–2010	WIID, 2021
Poland	pl	1980–2010	WIID, 2021
Portugal	pt	1970–2010	WIID, 2021
Puerto Rico	pr	1950–2000	WIID, 2021
Qatar	qa	1990–2010	WID, 2021, calc. top 1% income share
Republic of the Congo	cg	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Romania	ro	1980–2010	WIID, 2021
Russia	ru	1900	WID, 2021, calc. top 1% income share
		1920–1970	WID, 2021, calc. top 1% income share
		1950–1970	WID, 2021, calc. top 1% income share
		1980–2010	WIID, 2021
Rwanda	rw	1980	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Saint Lucia	lc	1990	WIID, 2021
		2010	WIID, 2021
Samoa	ws	2000–2010	WIID, 2021
Sao Tome and Principe	st	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Saudi Arabia	sa	1990–2010	WID, 2021, calc. top 1% income share

Table B.13: (continued)

Country	ccode	Decade	Source
Senegal	sn	1960	WIID, 2021
		1990–2010	WIID, 2021
Serbia	rs	1980	WID, 2021
		1990–2010	WIID, 2021
Seychelles	sc	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Sierra Leone	sl	1960	WIID, 2021
		1980	WID, 2021, calc. top 1% income share
		1990–2010	WIID, 2021
Singapore	sg	1940–1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Slovakia	sk	1980–2010	WIID, 2021
Slovenia	si	1980–2010	WIID, 2021
Solomon Islands	sb	2000–2010	WIID, 2021
Somalia	so	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
South Africa	za	1910–1980	WID, 2021, calc. top 1% income share
		1990–2010	WIID, 2021
South Korea	kr	1930–1980	WID, 2021, calc. top 1% income share
		1970–1980	WID, 2021, calc. top 1% income share
		1990–2010	WIID, 2021
South Sudan	ss	2000	WIID, 2021
		2010	WID, 2021, calc. top 1% income share
Spain	es	1960–2010	WIID, 2021
Sri Lanka	lk	1950–2010	WIID, 2021
Sudan	sd	1960	WIID, 2021
		1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Suriname	sr	1960	WIID, 2021
		1990	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Swaziland	sz	1990–2010	WIID, 2021
Sweden	se	1900–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
Switzerland	ch	1930–1970	WID, 2021, calc. top 1% income share
		1980–2010	WIID, 2021

Table B.13: (continued)

Country	ccode	Decade	Source
Syria	sy	1990–2000	WIID, 2021
		2010	WID, 2021, calc. top 1% income share
Taiwan	tw	1950–2010	WIID, 2021
Tajikistan	tj	1990–2010	WIID, 2021
Tanzania	tz	1960	WIID, 2021
		1990–2010	WIID, 2021
Thailand	th	1960–2010	WIID, 2021
Togo	tg	1990	WID, 2021, calc. top 1% income share
		2000–2010	WIID, 2021
Tonga	to	2000–2010	WIID, 2021
Trinidad and Tobago	tt	1950	WIID, 2021
		1970–1990	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Tunisia	tn	1960	WIID, 2021
		1980–2010	WIID, 2021
Turkey	tr	1960–2010	WIID, 2021
Turkmenistan	tm	1990	WIID, 2021
		2000–2010	WID, 2021, calc. top 1% income share
Tuvalu	tv	2010	WIID, 2021
Uganda	ug	1980–2010	WIID, 2021
Ukraine	ua	1980–2010	WIID, 2021
United Arab Emirates	ae	1990–2000	WID, 2021, calc. top 1% income share
		2010	WIID, 2021
United Kingdom	uk	1910	WID, 2021, calc. top 1% income share
		1930–1950	WID, 2021, calc. top 1% income share
		1960–2010	WIID, 2021
United States	us	1910–1930	WID, 2021
		1940–2010	WIID, 2021
Uruguay	uy	1960–2010	WIID, 2021
Uzbekistan	uz	1980–2000	WIID, 2021
		2010	WID, 2021, calc. top 1% income share
Vanuatu	vu	2010	WIID, 2021
Venezuela	ve	1980–2010	WIID, 2021
Vietnam	vn	1990–2010	WIID, 2021
Yemen	ye	1990–2010	WIID, 2021
Zambia	zm	1950	WIID, 2021

Table B.13: (continued)

Country	ccode	Decade	Source
Zimbabwe	zw	1970	WIID, 2021
		1990–2010	WIID, 2021
		1990	WIID, 2021
		2000	WID, 2021, calc. top 1% income share
		2010	WIID, 2021

Notes. Income Ginis are derived from the World Income Inequality Database (WIID) and the World Inequality Database (WID). In cases where income Ginis were missing, we calculated the income Ginis from the top 1% income shares provided by the WID.

Table B.14: The average effect of a land reform

	LSDV
Land reform	-4.47^{*}
	(2.561)
GDP p.c. 25,000	-11.74^{**}
	(5.349)
Constant	0.367
	(0.273)
Observations	138
R-squared	0.523
Time Fixed Effects	Y

Notes. Robust standard errors in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. The dependent variable is land inequality measured as Gini coefficient.

Table B.15: Sensitivity test for the construction of the joint inequality index

	(1) Civil War	(2) Civil War	(3) Civil War	(4) Civil War	(5) Civil War	(6) Civil War
Altern. Joint Index I	0.36*					
	(0.191)					
Altern. Joint Index II		0.61***				
		(0.130)				
Altern. Joint Index III			0.57***			
			(0.129)			
Altern. Joint Index IV				0.52***		
				(0.124)		
Altern. Joint Index V					0.62***	
					(0.133)	
Height Gini						0.72***
						(0.157)
Constant	−0.12	−0.24***	−0.22***	−0.20***	−0.25***	−0.29***
	(0.104)	(0.077)	(0.079)	(0.077)	(0.081)	(0.090)
Observations	228	1190	1190	1190	1118	924
R-squared	0.045	0.052	0.050	0.049	0.054	0.052
Time Fixed Effects	Y	Y	Y	Y	Y	Y

Notes. Heteroskedasticity-robust clustered standard errors in parentheses, ***, **, * significant on the 1, 5, and 10%-level, respectively. Inequality is composed of income Ginis, height Ginis and land Ginis which are all lagged by 1 decade. Alternative calculations for joint inequality index is used in models (1)–(5).

C Social and Economic Disparities and the 1918 Influenza Pandemic: Lessons for Today

---

Abstract

Over the course of the ongoing COVID-19 pandemic, it has become evident that its impact on society is disparate, both between and within countries. Hence, any pre-existing social and economic inequalities might be exacerbated by this event. We aim to provide important insights into the relationship between such pandemics and the social and economic disparities by investigating the 1918 influenza pandemic, including the underlying conditions and pre-existing inequalities, and empirically analysing its relationship with income and health inequality. We summarise different pathways through which pandemics might affect inequality, including asymmetric health risks, the labour market and the shock in aggregated demand. Our empirical findings suggest a positive but statistically insignificant relationship between income and health inequality and the 1918 influenza pandemic.

C.1 Introduction

As the COVID-19 pandemic has progressed, it has become clear that its impact on society is disparate. There has been a great divergence between the number of infected people and the mortality rates, both between and within countries. In addition, the social and economic implications of the pandemic itself and the effects of the countermeasures imposed differ widely. The latter reflects both divergent public health approaches and very different capabilities to deal with this pandemic medically, socially and economically. As a result, any pre-existing social and economic inequalities might be exacerbated by this event. This may be a justified concern, especially in countries with unequal access to health services between different social groups.

The implications of COVID-19 on inequality have been vividly discussed in public and in academia. While some economists find that the pandemic has put a constraint on an increase in income inequality (Deaton, 2021; Milanovic, 2020) others argue for opposing effects (Stiglitz, 2020). Particularly to mitigate negative effects, dealing with pandemics requires a thorough understanding of the global dynamics, taking into account country- and region-specific circumstances. One approach to gain this understanding is to look at similar events in the past. Consequently, recent research has analysed the relationship between past pandemics and their socioeconomic consequences. This includes the impact on economic growth (e.g., Barro et al., 2020; Eichenbaum et al., 2021; Jordà et al., 2022), the labour market (e.g., Coibion et al., 2020), individual economic outcomes such as unemployment (Guimbeau et al., 2022; Nelson, 2010), education (Beach et al., 2022; Helgertz & Bengtsson, 2019; Percoco, 2016), health (e.g. Almond, 2006; Guimbeau et al., 2022 and, finally, inequalities (e.g., Alfani, 2015; Alfani and Murphy, 2017; Furceri et al., 2022; Galletta and Giommoni, 2022). For example, Furceri et al. (2022) analyse the difference in income inequality in response to five epidemics over the last two decades for countries worldwide. They note an increase in income inequality and a rise in the top income shares (Furceri et al., 2022). Looking at the 1918 influenza pandemic in Italy, Galletta and Giommoni (2022) draw the same conclusion: a higher severity of the pandemic is associated with higher income inequality in the affected communities, due to increased incomes of the richest 20%. Past pandemics can provide important lessons. However, the relationship with inequality has mainly been studied in specific (developed) countries, and the empirical evidence is still unclear, as one of the main limitations is the availability of data for earlier periods and developing countries.

We aim to contribute to this discussion by empirically investigating the relationship between such a pandemic and inequality in countries worldwide. More specifically, we look at inequalities in income and in health. We focus on the following aspects in our study: Can the 1918 influenza pandemic explain variations in inequality? Did high-income and low-income countries differ in their response to the pandemic? How do the underlying factors that existed before the pandemic interact with the health and economic shock caused by the pandemic?

In 1918, the first cases of a heavily spreading and deathly influenza were reported, most possibly in the United States (US) (Olson et al., 2005; Patterson & Pyle, 1991), quickly growing into a pandemic. The disease was characterized by high mortality rates and high dispersion across individuals, regions, and countries with great spatial disparities. Over the course of two years, the influenza pandemic caused an immediate global health shock, killing up to 17.4 million people, which corresponded to almost one per cent of the world's population (Spreeuwenberg et al., 2018).[27] The countries with the highest number of victims in 1918 were among the poorest countries in the world at that time, with India recording the highest death rate[28], along with Cameroon, Kenya and other East African countries (Guntupalli & Baten, 2006). Within Latin America, Guatemala and Mexico had the highest numbers of victims, both of which had fallen back in their income development over the previous decades and had enormous health inequalities (Baten & Blum, 2014).

A pandemic, such as the 1918 influenza pandemic, could lead to an increase in various dimensions of inequality, such as income, wealth, health, education, and gender (Blumenshine et al., 2008). Interestingly, we actually observe an increase in inequality from the 1910s to the 1920s in most regions of the world. In this study, we therefore analyse whether the rise in inequality in the 1920s can be linked to the pandemic. We first qualitatively summarize different possible pathways based on the literature. This includes the asymmetric health risk between poorer and

[27]Patterson and Pyle (1991) estimated between 24.7–39.3 million. Johnson and Mueller (2002) estimated a number of 50 million up to 100 million deaths. More recent studies arrive at 40.1 million deaths (Barro et al., 2020).

[28]The estimation of India's mortality rate is discussed in Section C.4.

richer individuals and populations, the effect on the labour market, and the shock in aggregate demand and economic growth. We then assess the quantitative relationship between pandemics and inequalities within countries by empirically examining the 1918 influenza pandemic, including the underlying conditions and pre-existing disparities. We focus on income inequalities and health inequalities as outcomes. The income Gini measure is derived from databases that use tax and survey data on income distribution. Health inequalities refer to the health status of individuals (Evans et al., 2009) as reflected in different life expectancy, underlying health conditions or in the height of an adult. The latter in particular mirrors social and economic circumstances in early childhood and adolescence. In addition, it is a reflection of the state of health and the access to medical care in a country, as well as the nutritional status of its population (Baten & Blum, 2011). We use the distribution of heights within the country as a proxy for health inequalities. With the usage of anthropometric data, we address the problem of data availability at the beginning of the 20^{th} century, and we are able to include a larger share of developing countries in our analysis than previous studies. For our empirical analysis, we perform several cross-country regression analyses. We do not find that the increase in inequality from the birth decade 1910 to 1920 can be linked to the pandemic's mortality rates. As we aim to provide critical insights into the ongoing COVID-19 pandemic with regard to the social and economic inequalities of one of the largest pandemics in history, we discuss policy implications based on our qualitative and quantitative research findings, such as social transfers which target the poorest households in a society.

This study is structured as follows: In Section C.2 we first summarise different pathways through which pandemics might affect inequality. Section C.3 provides an overview of the empirical evidence on the impact of pandemics. Next, we describe our data and empirical methodology, and Section C.5 presents the descriptive statistics and regression results. Section C.6 includes a discussion of today's implications and Section C.7 concludes.

C.2 Mechanism

Based on the literature, we explore different pathways on how a pandemic might affect inequalities within a country, in terms of income and health. We consider socioeconomic conditions which lead to asymmetric health risks in a society, the dynamics of the labour market and the economic shock on aggregate demand.

Asymmetric health risk. A disease, which is prevalent worldwide and could theoretically affect everyone equally, could nevertheless disproportionately hit poor people, people from specific socioeconomic groups and ethnic minorities or people in low-income countries. As a result, it could increase inequalities, as poorer people already suffer more from pre-existing health conditions that make them particularly vulnerable (Blundell et al., 2020). In addition, certain groups may be exposed differently to the virus itself because of more crowded households, workplaces, and/or the need to take public transportation, as well as the occupation itself. This may prevent low-income workers in particular from having the flexibility to work from home (Blumenshine et al., 2008). In fact, studies of past epidemics and pandemics have shown that people living in poverty or on low incomes are at higher risk than richer individuals, both in terms of infections and deaths (e.g., Alfani, 2022). Lower skills are associated with a higher risk of mortality (Furceri et al., 2022), worse livings conditions (Mamelund, 2006), as well as having a specific occupation (Benitez et al., 2020; Brandily et al., 2021), and the belonging to an ethnic or religious minority in a country (Blundell et al., 2020). In Bombay and Bahia, for example, the highest mortality rates during the 1918 influenza pandemic occurred among those social groups that were already most disadvantaged before the pandemic (Mills, 1986; Souza, 2005).

Also, for the COVID-19 pandemic, we observe an unequal spread of the disease: various studies show that death rates are higher for black people than for white people, controlling for age and socioeconomic status (Gross et al., 2020; McLaren, 2021), and that women are more affected by redundancy and loss of income than men (Dang & Nguyen, 2021). Inferential, evidence suggests, that underlying conditions such as pre-existing health issues and living conditions are essential in defining the impact of the pandemic within a country, hitting those harder with poor underlying conditions, resulting in larger mortality rates and higher economic losses and in turn, might lead to higher inequalities in terms of income and health.

A long-term effect of a pandemic on inequality could be explained by in-utero exposure to the disease, affecting individual health in adulthood. This hypothesis is linked to the so-called Baker hypothesis or 'fetal origins hypothesis', which states that environmental conditions, such as a health shock, experienced in the womb, can lead to worse health conditions as an adult, such as heart disease or diabetes (Barker, 1998). Studying the influenza pandemic, multiple studies have shown that children born in the year 1919 had poorer health and socioeconomic outcomes than surrounding birth cohorts. This includes outcomes such as a higher likelihood of disability, and lower educational attainment, and it is assumed to have inhibited child growth (see Almond, 2006; Almond and Currie, 2011; Almond and Mazumder, 2005; Lin and Liu, 2014).

Yet a pandemic's severity depends not only on the individual level but also on a country's ability to cope with it health-wise and economically. An important factor is the status of the health care system, which is neither similar between countries, nor is it provided with the same quality across regions (Case & Deaton, 2017). If low-income countries face lower health of their population and higher contagion risk in the event of a pandemic, even more, appropriate resources will be needed to combat it. This includes hospitals, doctors, nurses and medicines, while at the same time, the efficiency and quality of the use and allocation of these resources are lower compared to richer countries (Weil, 2014). In fact, it is shown that there is a negative association between a country's income level and the severity of the 1918 influenza pandemic (Murray et al., 2006). Rising health inequality could therefore not only be linked to the disparity in the health risk and variations in case fatality rates within a country (see e.g., Alfani, 2022; Furceri et al., 2022), but also between different income groups, as well as between high- and low-income countries. Since low-income countries do not have the resources to recover quickly from a pandemic, the COVID-19 pandemic might end up exacerbating inequalities between countries (Stiglitz, 2020).

Labour market dynamics. A health shock caused by a pandemic might quickly be followed by an economic downturn, which might result in job losses and a decline in aggregate demand. As in previous recessions, including the 2008/09 financial crisis, people in low-skilled jobs may be the first to be laid off (Ferreira, 2021). Due to the economic shock caused by the COVID-19 pandemic, less-educated and low-income workers more easily lose their jobs and subsequently find themselves without income and—in some countries—without any health insurance (Rota & Weisdorf, 2020).[29] In early 2020, at the beginning of the COVID-19 pandemic Coibion et al. (2020) carried out a household survey in the US and find that there were already over 20 million job losses until April 2020. This number is even higher than the US labour market experienced during the financial crisis of 2008/09. A similar trend was observed in other European countries for low-income workers and minority ethnic groups (Crossley et al., 2021; Galasso & Foucault, 2020).

Working remotely is one way to limit the risk of exposure. But this requires a job that allows it, which is mainly possible for higher-skilled workers (Avdiu & Nayyar, 2020; Blundell et al., 2020). Blundell et al. (2020) visualize the possibilities for remote work by sector and social groups. They state that education, and higher earnings, are clearly correlated with the possibility of working from home. Lower-skilled workers are most likely employed in sectors and occupations that either pose higher health risks or have been closed during a lockdown, because home-based work is barely possible. Beach et al. (2022) thus conclude that COVID-19 may have interacted with underlying inequalities and predict higher labour market inequalities.

Shock on aggregated demand. In addition to the direct health effects of the virus, there may also be economic implications, which in turn affect the poorest in society. In the case of the 1918 influenza pandemic, it is assumed that there was a negative shock to aggregate demand (Brainerd & Siegler, 2003). This might be due to the shock in the labour market, but also to lower consumption and higher amount of savings in times of uncertainty. Therefore, higher mortality rates are associated with a higher fall in income per capita in the 1920s (Brainerd & Siegler, 2003). The COVID-19 pandemic could also act as a negative demand shock for consumer goods, and simultaneously, for labour supply (Eichenbaum et al., 2021). In addition, economic growth may be hampered by Non-pharmaceutical interventions (NPIs) in the form of sector closures or lock-downs, as this imposes direct costs on affected firms and households (Correia et al., 2022).

The social and economic implications of COVID-19 itself and the effects of the countermea-

[29] The latter might not be true for countries with general health care systems such as Germany. But it might be the reality for developing countries, or in high-income countries with no universal health care system such as the US.

sures imposed differ widely. The latter reflects both divergent public health approaches and very different capacities to deal with this pandemic medically, socially, and economically. As a result, any pre-existing social and economic inequalities amongst and within countries might actually be exacerbated by the current pandemic. However, similar to the case of a recession (Atkinson, 2008), the disproportionate impact on the most vulnerable, socially and economically, could be cushioned or even prevented by appropriate public policies. We will discuss this possibility and policy interventions in the final section.

C.3 Literature review: empirical evidence from previous research

Pandemics, much like wars or the impacts of climate change, are regarded as a threat to global public goods (McGregor et al., 2019). Dealing with pandemics therefore requires a thorough understanding of the global dynamics, taking the underlying social and economic aspects of different countries and regions into account. The recent COVID-19 pandemic renewed the interest in its socioeconomic consequences. Previous research looked at the relationship between pandemics and their short- and long-term impacts on health, inequality, employment, economic growth and education on the individual level as well as on a country level.

The economic effects of pandemics have been studied extensively, with several authors looking at economic outcomes such as GDP and GDP growth, income, real wages, unemployment and capital returns. For example, Barro et al. (2020) concentrated on the world's population mortality rate of the influenza pandemic of 1918 and its negative correlation with economic growth and applied it to today's population. They find that the influenza pandemic is linked to lower economic growth and therefore predict a six per cent decrease in real GDP per capita for the COVID-19 pandemic (Barro et al., 2020). For Italy after 1920, Carillo and Jappelli (2022) also suggest a lower GDP for regions with higher mortality rates during the influenza pandemic. However, they find that this link is not persistent. Similar findings are made by Dahl et al. (2022) for the development of income in Denmark. They look at the relationship between the influenza pandemic and income during and in the years after. In line with Beach et al. (2022) for economic growth, they find a slight dip in income growth in the short-run combined with a fast and full recovery in the following years until 1925, and no correlation in the long run. In addition, they suggest that the variation in employment follows a similar pattern (Dahl et al., 2022). Brainerd and Siegler (2003) show that US states with low death rates during the influenza pandemic did not grow as much as other states over the following ten years, but they do not find an association with per capita income. Correia et al. (2022) account for the severity of the pandemic in terms of deaths, and for the imposed NPIs in 1918 and 1919 across major US cities. They point out that cities that reacted fast and shut down schools, theatres, etc. developed better in the short- and medium-run. Using a Difference-in-Difference approach, Karlsson et al. (2014) find a statistically significant negative long-run impact on capital returns and an increase in poverty rates in Sweden due to the pandemic 1918. This is in contrast to Garrett et al. (2009), who do not confirm an effect on real earnings in the US. To predict the short-term economic impacts of the COVID-19 pandemic, Eichenbaum et al. (2021) apply a new Keynesian model and show that the pandemic acts as a negative demand shock for consumption goods and for labour supply. The imposed measures in some OECD countries had a direct effect on the labour market for the lower-income group, as observed by Galasso and Foucault (2020).

Much less research has been done on the relationship between pandemics and inequality, and it remains controversial: it is argued that the character of the event is analogous to a financial crisis or recession by causing an exogenous shock (Peckham, 2013), resulting in higher income inequality. This positive correlation with income inequality is confirmed by de Haan and Sturm (2017), who looked at banking crises at the country level. In contrast, using annual data from 1960 to 2016 for 43 countries, Camacho and Palmieri (2019) do not find a significant impact of economic recessions on income inequality. Furceri et al. (2022) analyse past pandemics and financial crises and recessions during the 2000s and 2010s and perform a cross-sectional analysis. They find that past pandemics are associated with higher income inequality, but not in the case of financial crises. Their findings are in line with Galletta and Giommoni (2022): by looking at Italian municipalities they show that the 1918 influenza pandemic lowered the income of the lower-income groups, leading to higher income inequality. In contrast, Alfani (2015) and Alfani and Murphy (2017) find that the Black Death is linked to lower inequality in the long run due to an increase in real wages. However, this is not the case when looking at past epidemics that have occurred more recently: Alfani (2022) states that the significant correlation with previous events might be driven by specific circumstances at that time, such as the quality of institutions.

A third strand of literature deals with the long-term consequences of pandemics, more specifically the socioeconomic and health status. Different determinants of health are studied, mainly on the individual level, like infant mortality, life expectancy and human capital formation. As long-term data for the COVID-19 pandemic are not yet available, the following studies mainly analyse the relationship with the 1918 influenza pandemic. Guimbeau et al. (2022) conducted an individual-level study for Brazil and found negative associations of all of these aforementioned socioeconomic and health outcomes. Nelson (2010) confirms those findings by comparing individuals born during the pandemic with individuals born in the years before and after. He finds that the latter is more likely to have more years of schooling and higher wages. In-utero exposure during a pandemic is associated with worse health conditions as an adult. This theory is reflected by the 'fetal origins hypothesis' (Barker, 1998), which was first studied in an economic context by Almond (2006) and has experienced a surge in popularity since. As the 1918 influenza pandemic functions as a natural experience in economic terms, Almond (2006) used birth cohort data for the US to compare individuals born during the pandemic to birth cohorts in the surrounding years. His findings confirm the hypothesis that individuals who were exposed to the pandemic in-utero were worse off, both in terms of health and socioeconomic status. In addition, individual studies examine the effect on specific countries confirming this theory. For Taiwan, Lin and Liu (2014) support the theory that the birth cohorts of 1918 and 1919 are more likely to have lower educational attainment, are substantially less healthy and report lower heights. Those results are similar to Hong and Yun (2017) for Korea; and for Japan, where the authors find that children's growth was hindered in the respective birth cohort (Ogasawara, 2017). Percoco (2016) examines a negative correlation with human capital accumulation in Italy, which is also observed in the US (Beach et al., 2022), Sweden (Helgertz & Bengtsson, 2019) and Switzerland (Neelsen & Stratmann, 2012), even though much lower in scale. Vollmer and Wójcik (2017) contradict with previous findings on the effect of in-utero exposure on health and economic outcomes as adults. Conducting a global study they cannot confirm that adults, who experienced the influenza pandemic in the womb, are statistically significantly different in education, employment, or disability outcomes. They conclude that the previous positive findings might be driven by a publication bias (Vollmer & Wójcik, 2017).

We contribute to these strands of literature by providing a cross-country study of the 1918 pandemic in regard to income inequalities and health inequalities. Previous research mainly focused on economic outcomes such as GDP, or on individual health outcomes. The limited number of empirical cross-country studies examining the link between pandemics and inequality is mainly due to the scarce availability of data in the early 20$^{\text{th}}$ century. Thanks to the collections of global historical income data and health data, namely the data collection on height by Baten and Blum, 2015 and the recent update done by Radatz and Baten, n.d., we are able to provide empirical evidence on whether and to which extent inequality varies with the severity of the 1918 influenza pandemic.

C.4 Data and method

C.4.1 Measuring the severity of a pandemic

Our main independent variable is the excess mortality caused by the 1918 influenza pandemic. It is expressed relative to a country's total population, of 1,000 inhabitants. Our main data source is Barro et al. (2020) from whom we gained data for 48 countries. We combine this data with values from Patterson and Pyle (1991) and extended it with manually collected data from different national statistical reports and demographic censuses.

As in Barro et al. (2020), the mortality rate is calculated from the death tolls from the flu and flu-related deaths such as pneumonia. If those numbers were not available, we took the total deaths per year and calculated the excess mortality rate for the pandemic years 1918–1920. This was done following the method of Murray et al. (2006):

$$\textit{Excess Flu Mortality Rate} = \sum_{t=1918}^{t=1920} (M_t - (\frac{M_{1915} + M_{1916} + M_{1917} + M_{1921} + M_{1922} + M_{1923}}{6}))$$

The average mortality rate for three years before and three years after the pandemic is calculated for each country and compared to the annual mortality rate from 1918–1920 (except for Chile, which was also hit in 1921). The value exceeding the average value is our excess flu mortality rate (*flu rate*). All years that were not affected by this pandemic, are set to zero. For the regression analysis, we use the cumulative excess flu rate for all pandemic years. For all countries included, mortality during the pandemic years exceeded the mortality rates in the years before and after.

The variation in the severity of the pandemic is graphically shown in the world map in Figure C.1. We display the excess flu mortality rate by country, where darker shades indicate higher rates and lower rates are shown in lighter shades. We show 52 countries for which we have data. Southeast Asia as well as Eastern Europe and Central Asia stand out as regions with the highest mortality rates (in line with Patterson and Pyle, 1991).

However, some of the data might be regarded with caution. First, missing data on mortality during this period could potentially be a problem. This could lead to a bias when looking at the world regional level, as data is much more available for Latin American countries than for African or Asian countries. Moreover, within countries, existing statistical reports from the early 20$^{\text{th}}$ century might lead to an under-reporting of real death rates (Johnson & Mueller, 2002). In India, where no official statistical reporting was conducted, and mortality rates were estimated across the whole population for all causes of deaths (Chandra & Kassens-Noor, 2014), excess mortality caused by the influenza pandemic could be over- or underestimated and the impact of the entire region could be magnified. High death rates in Southeast Asia might therefore be mainly driven by India.

C.4.2 Measuring inequality in income and health

For our dependent variable inequality, we compare disparities between the cohort of people born during the 1910s with the cohort born in the 1920s. Even though the pandemic might have lasted from 1918 to 1920, the most severe mortality impact was experienced in 1918 by most countries.[30] We consider income inequality as well as health inequality as our outcome variable. The usage

[30] There were still some reported cases in 1920, however, Johnson and Mueller (2002) conclude that those cases can be attributed to local (epidemic) outbreaks. We therefore assume that the birth decade 1920 (ranging from 1920–1929) should not be primarily influenced by the influenza pandemic.

of these two inequality measures might tell us more about the impact of the pandemic than one measure alone could do, especially regarding the early 20^{th} century, where inequalities in health are mainly driven by the surrounding circumstances and data on inequalities for this period is still scarce.

For income inequality, we rely on the dataset from van Zanden et al. (2014) and the World Inequality Database (WID), as these databases provide us with data for years before 1930. Income inequality is calculated as the Gini coefficient, ranging from zero to one, where the latter indicates the highest inequality level. Our measure of health inequality is based on the collection of height data of Baten and Blum (2014). Height Gini is calculated per birth cohort and country and is already widely used in empirical studies as a proxy for health inequality (see literature review in Blum, 2014). We follow this approach. Adult height is defined by the experience an individual made during early childhood and adolescence. Besides the individual genes, sufficient intake of high-quality nutrition, as well as access to medical supplies, matters to determine the final height of an adult. In times of war, famine and (health) crisis, these social circumstances might severely affect a specific country or region. Therefore, the distribution of height data in a country reflects the level of inequality in that country.

To ensure comparability, we restrict our sample to those countries for which we have data on both measures of inequality, health and income, as well as data on mortality that can be attributed to the 1918 influenza pandemic. We can include up to 29 countries worldwide in our analysis, including 16 developing countries across all world regions (see Appendix for included countries).[31]

In Figure C.2 we display the change in inequality between the birth decade 1910 and 1920 with income inequality on the left and health inequality on the right side, respectively. In addition, we compare high- and low-income countries for each inequality measure. Countries are classified based on their national GDP per capita in 1910, available from the Maddison Project Database. A positive change in inequality is shown by both measures of inequality and the two income groups, implying an increase in inequality between 1910 and 1920. In the high-income group, the changes in income and health inequality vary around zero, with a mean of one income Gini point, whereas the health Gini coefficient is slightly lower with a mean of 0.28. We observe an exceptionally high change in

[31] Unfortunately, we had to remove Chile from our dataset due to data concerns for the estimates for the decade 1910 (see Baten and Llorca-Jaña, 2021).

income inequality in Canada, with a sharp increase between 1910 and 1920. In comparison, low-income countries show a substantially higher mean for both measures, with Brazil, South Africa and Peru being outliers in the case of income inequality.

C.4.3 Methodology

To assess the relationship between the influenza pandemic and inequality, we perform a first difference estimation. Our specification is described in the form

$$Inequality_{it+1} - Inequality_{it} = \beta_1 \triangle Flurate_{it} + \beta_2 \triangle X_{it} + \triangle \epsilon_{it},$$

where the outcome *Inequality* describes the differences in inequality in terms of either income inequality or health inequality between the periods $t+1$=1920s and t=1910s. With our main independent variable *flu rate*, we capture the severity of the 1918 influenza pandemic. We include a set of controls in our analysis, described by $\triangle X_{it}$. $\triangle \epsilon_{it}$ is the first-difference error term. As the influenza pandemic appeared at the same time as the end of the First World War, we include the number of combat deaths related to WW1 for each country. We further control for demographic factors such as the degree of urbanisation, population density and the educational level within a country, measured as years of schooling. As we are not able to account for government expenditures (which might be especially relevant for health expenditures) due to data limitations, we use GDP per capita as a proxy for a country's resources on government spending, pensions, or social transfers. Although one might be concerned that GDP per capita might influence *inequality* through our main independent variable *flu rate*, as mortality rates due to the 1918 influenza pandemic might be higher in low-income countries compared to high-income countries, we still see the need to include it to control for a country's economic development. To address the concern of bad controls, we also display our regression results without the inclusion of additional control variables.

We further control for the degree of democratisation in a country, as a more democratic country is associated with a larger share of redistribution towards the poor (Gradstein & Milanovic, 2002). Besides these political factors, we include climate as a geographical characteristic. An overview of all variables and sources is given in the Appendix. The inclusion of these controls might provide us with a more precise model and higher explanatory power. However, it might also be the case that some control act as bad controls. For example, GDP per capita might be correlated with our main independent variable *flu rate*, as mortality rates due to the 1918 influenza pandemic could be higher in low-income countries compared to high-income countries, and therefore have an influence on our dependent variable *inequality*.

C.5 Descriptive statistics and regression results

C.5.1 Descriptive statistics

In Table C.1 we provide our summary statistics. Our variables are measured on a country level per (birth) decade and show the differences between the birth decade 1910 and 1920. We include countries for which mortality and inequality data are available. Unfortunately, the scarce availability of both, mortality data and inequality data at the beginning of the 20^{th} century severely limits our sample. Nonetheless, we are able to construct a dataset for 29 countries worldwide, including 16 low-income countries. For our analysis, we take the natural logarithm of *flu rate*. The change in income inequality shows 4.19 Gini points on average, indicating a rise in income inequality, and is slightly lower for health inequality (2.45 Gini points).

We display the correlation between the pandemic and the differences in income inequality on the left and differences in health inequality on the right of Figure C.3. In addition, we mark whether a country belongs to the group of high-income or low-income countries, which, according to our observation, tend to be more negatively correlated with the pandemic. A negative value for the difference in inequality indicates a decrease in the Gini coefficient between the (birth) decade 1910 and 1920. France (fr), for example, shows a slight reduction in income inequality despite having high flu mortality rates. Overall, however, we observe a positive relationship between the severity of the pandemic and the differences in income inequality. Brazil (br), Peru (pe) and Canada (ca) are exceptions: inequality has increased in these countries, but they have not been as severely affected by the pandemic as other countries. Denmark (dk) and Sweden (se) on the other hand, both show a negative development of the income Gini of between -16 and -24 Gini points. This discrepancy might be explained by the fact that Sweden and Denmark already had high inequality levels during the 1910s, much higher than the levels of Canada for example. With the introduction of the welfare state in Scandinavian countries in the 1910s, inequality was systematically reduced by the late 1940s (Gärtner & Prado, 2016), which may have diminished any negative effects of the pandemic.

Similarly, we observe a positive relationship between health inequality and excess flu mortality. One exception is Italy (it) showing a sharp fall in Gini points between the birth cohorts 1910 and 1920 of almost -10 Gini points while having relatively high mortality rates. This might be due to the introduction of Italian insurance for occupational illness in 1925 (Reich & Goldman, 1984).

C.5.2 Regression results

The results from our first-difference estimation are displayed in Table C.2. We include up to 29 countries in our analysis. Our results suggest a positive relationship between the pandemic's mortality rate and the differences in income inequality in models (1) and (2). The coefficient for *flu rate* is statistically significant in these two models but becomes insignificant in all other models once we include further controls. Also, we do not find that WW1 is significantly correlated with a change in income inequality. In the case of health inequality, we observe similar results: the coefficients for *flu rate* are positive but statistically insignificant. In a step-wise fashion, in Table C.3, we include political, social and economic controls in models (1)–(6) with *income inequality* as the outcome, and for the outcome variable *health inequality* in models (7) to (12), respectively. The coefficients for *flu rate* remain positive but statistically insignificant. Therefore, we do not observe that a change in inequality between the birth decades 1910 and 1920 can be linked to the 1918 influenza pandemic.

However, it might be the case that different income groups reacted differently to the pandemic, as developing countries might not be able to deal with the medical, social or economic implications of the pandemic in the same way as high-income countries. In Table C.4 we therefore check for possible heterogeneous effects by controlling for different income groups. In models (1)–(2), and (5)–(6) we omit countries belonging to the higher income group, based on the world average GDP per capita in 1910. Compared to the regression models for higher-income countries (3)–(4) and (7)–(8), we observe modestly larger coefficients. We observe insignificant coefficients for *flu rate* on health inequality, except for model (5). But the significance disappears as soon as we control for other factors (model 6).

To test the robustness of our results, we exclude India from our sample, as it might have driven the results. India shows the highest flu rate in our sample, but the estimates need to be regarded with caution based on the controversial discussion about the reliability of mortality estimates of India, as discussed in Section C.4. Therefore, in Table C.5 we re-run out regression analysis without the inclusion of India. The coefficient of flu rate actually turns significant on the 5% level for our outcome income inequality in models (1) and (2). However, as soon as we include additional controls, flu rate and income inequality seem not to be significantly correlated. In models (3)–(4) we can confirm our main results that the relationship between health inequality and the severity of the 1918 influenza pandemic is not statistically significant.

C.6 Discussion and implications for today

Even though our empirical findings are mixed, we can learn from the dynamics of one of the major past pandemics. Reacting with adequate public policy intervention could prevent possible effects of the recent COVID-19 pandemic. Already in the first months of the pandemic, there was some evidence that an increase in inequality could be hindered with an appropriate fiscal policy like public transfers (Aspachs et al., 2022; Hacıoğlu-Hoke et al., 2021) or benefit transfers to the poorer population (Balasubramanian et al., 2021).

There are strong similarities between the influenza pandemic of 1918 and the current COVID-19 pandemic, but it also has its limitations. There was a high level of infection in each pandemic (Wheelock, 2020), affecting individuals worldwide. But even though both consist of an influenza disease attacking the respiratory system, the characteristics of the virus itself might differ. Also, the underlying conditions 100 years ago are different from today's: the influenza pandemic in 1918 encountered an extremely poor population, where large parts of the population were malnourished, people lived in poor health conditions, and crowded households were common. Although poverty is still prevalent, the share of the world population living in extreme poverty dropped from 60% in the 1920s (Bourguignon & Morrison, 2002) to approximately around 9% in 2020 (World Bank, 2018). It is seen that higher population density and low-income groups, as well as minorities, are more exposed to the virus and therefore have a higher risk of infection and death (Garrett et al., 2009). Targeting the poorest households in society through social or cash transfers might therefore reduce inequalities. Also, low-income countries could introduce social policies such as unconditional cash transfers or utility tariff reliefs to cushion the consequences of COVID-19, as already done in Ghana, Mozambique and Zambia (Lastunen et al., 2021).

Although underlying health conditions, such as life expectancy, have changed tremendously in recent decades, health status and the capacity of the health systems to deal with this pandemic medically, still vary between countries and different regions. Thus, Universal Health Coverage (UHC) remains an important determinant in reducing health inequalities. Not only in 1918 hospitals were overwhelmed by a large number of infected people and deaths (see Crosby, 2003); but also, in 2020 the health care system broke down or nearly collapsed (e.g. in Italy). Therefore, government expenditures are needed in order to increase the quality of the health sector as well as the number of hospitals and medical services.

Finally, in response to the pandemic, countries imposed economic and social measures to reduce the spread of the virus, such as lockdowns or restrictions. For the 1918 influenza pandemic, we do have evidence of the introduction and the effectiveness of NPIs in the US, including school closures, ban of public gatherings, and reduced working hours of businesses (see Bootsma and Ferguson, 2007; Correia et al., 2022; Markel et al., 2007). However, today's NPIs have had a very different scale and economic impact, as complete sector closures during COVID-19 led to layoffs, especially of low-income workers. But governments also have reacted with social policies. Clark et al. (2021) observed that policy interventions during COVID-19, mainly targeting the poorest households, led to a fall in absolute income inequality in the short run. In the long run, it even might be the case that within-country inequality remains unchanged because of the massive redistribution to low-income households and firms (Mahler et al., 2022), which might mitigate the effect of the lockdown measures and the decrease in consumption behaviour. This includes tax cuts, subventions to firms and social transfers to the most vulnerable in society, including unemployed people and people working in specific sectors who are exposed to health problems or are affected by imposed restrictions such as lockdowns (Ferreira, 2021). To ensure that the COVID-19 pandemic does not have long-lasting effects on inequality, these measures however need to be more widely distributed between countries, but especially within countries to mitigate possible effects on the individual in terms of economic outcomes as well as health.

C.7 Conclusion

At the beginning of the COVID-19 pandemic, in early 2020, some economists, as well as international organizations like the United Nations, warned that COVID-19 might increase income inequality between countries (e.g., Deaton, 2021) and within-countries (Ferreira, 2021). Opponents argued that the pandemic might function as an equalizer (Milanovic, 2020).

By looking at one major past pandemic 100 years ago, we investigate the association between the 1918 influenza pandemic and income and health disparities for countries worldwide. We summarize different pathways including the asymmetric health risk between individuals, the dynamics on the labour market, and the shock in aggregate demand. We empirically test the relation between the severity of the pandemic on the difference in inequality from the birth decades 1910 to 1920 by applying a first-difference estimation. Our results suggest no statistically significant relationship between the 1918 influenza pandemic and income inequality or health inequality. However, we could—and should—learn from the consequences of one of the biggest pandemics of the past: appropriate public intervention could prevent potential short- and long-term impacts of the recent COVID-19 pandemic. The implementation of social policies to cushion the impact of job losses and economic downturns could include, for example, (unconditional) cash transfers to the poorest in society, redistribution to businesses, tax cuts and subsidies. Universal Health Coverage remains an important factor in reducing health inequalities in a country and should be further encouraged. Similarly, states should increase government spending to improve the quantity but also the quality of medical infrastructure and services.

C.8 References

Alfani, G. (2015). Economic inequality in northwestern italy: A long-term view (fourteenth to eighteenth centuries). *The Journal of Economic History*, *75*(4), 1058–1096.

Alfani, G. (2022). Epidemics, inequality, and poverty in preindustrial and early industrial times. *Journal of Economic Literature*, *60*(1), 3–40.

Alfani, G., & Murphy, T. E. (2017). Plague and lethal epidemics in the pre-industrial world. *the Journal of economic History*, *77*(1), 314–343.

Almond, D. (2006). Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. *Journal of political Economy*, *114*(4), 672–712.

Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives*, *25*(3), 153–172.

Almond, D., & Mazumder, B. (2005). The 1918 influenza pandemic and subsequent health outcomes: An analysis of sipp data. *American Economic Review*, *95*(2), 258–262.

Aspachs, O., Durante, R., Graziano, A., Mestres, J., Montalvo, J. G., & Reynal-Querol, M. (2022). Real-time inequality and the welfare state in motion: Evidence from covid-19 in spain. *Economic Policy*, *37*(109), 165–199.

Atkinson, A. B. (2008). Unequal growth, unequal recession? *Organisation for Economic Cooperation and Development. The OECD Observer*, (270/271), 7.

Avdiu, B., & Nayyar, G. (2020). When face-to-face interactions become an occupational hazard: Jobs in the time of covid-19. *Economics Letters*, *197*, 109648.

Balasubramanian, S., Bounader, L., Bricco, M. J., & Vasilyev, D. (2021). *Is there a one-size-fits-all approach to inclusive growth? a case study analysis*. International Monetary Fund.

Barker, D. J. (1998). In utero programming of chronic disease. *Clinical science (London, England: 1979)*, *95*(2), 115–128.

Barro, R. J., Ursúa, J. F., & Weng, J. (2020). *The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity* (tech. rep.). National Bureau of Economic Research.

Baten, J., & Blum, M. (2015). Height gini. http://hdl.handle.net/10622/HKZPR9

Baten, J., & Blum, M. (2011). Anthropometric within-country inequality and the estimation of skill premia with anthropometric indicators. *Review of Economics*, *62*(2), 107–138.

Baten, J., & Blum, M. (2014). Why are you tall while others are short? agricultural production and other proximate determinants of global heights. *European Review of Economic History*, *18*(2), 144–165.

Baten, J., & Llorca-Jaña, M. (2021). Inequality, low-intensity immigration and human capital formation in the regions of chile, 1820-1939. *Economics & Human Biology*, *43*, 101030.

Beach, B., Clay, K., & Saavedra, M. (2022). The 1918 influenza pandemic and its lessons for covid-19. *Journal of Economic Literature*, *60*(1), 41–84.

Benitez, J., Courtemanche, C., & Yelowitz, A. (2020). Racial and ethnic disparities in covid-19: Evidence from six large cities. *Journal of Economics, Race, and Policy*, *3*(4), 243–261.

Blum, M. (2014). Estimating male and female height inequality. *Economics & Human Biology*, *14*, 103–108.

Blumenshine, P., Reingold, A., Egerter, S., Mockenhaupt, R., Braveman, P., & Marks, J. (2008). Pandemic influenza planning in the united states from a health disparities perspective. *Emerging infectious diseases*, *14*(5), 709.

Blundell, R., Costa Dias, M., Joyce, R., & Xu, X. (2020). Covid-19 and inequalities. *Fiscal studies*, *41*(2), 291–319.

Bootsma, M. C., & Ferguson, N. M. (2007). The effect of public health measures on the 1918 influenza pandemic in us cities. *Proceedings of the National Academy of Sciences*, *104*(18), 7588–7593.

Bourguignon, F., & Morrison, C. (2002). Income distribution among world citizens: 1820–1990. *American Economic Review*, *92*(September), 1113–1132.

Brainerd, E., & Siegler, M. V. (2003). The economic effects of the 1918 influenza epidemic. *Available at SSRN 394606*.

Brandily, P., Brébion, C., Briole, S., & Khoury, L. (2021). A poorly understood disease? the impact of covid-19 on the income gradient in mortality over the course of the pandemic. *European economic review*, *140*, 103923.

Camacho, M., & Palmieri, G. (2019). Do economic recessions cause inequality to rise? *Journal of Applied Economics*, *22*(1), 304–320.

Carillo, M. F., & Jappelli, T. (2022). Pandemics and regional economic growth: Evidence from the great influenza in italy. *European Review of Economic History*, *26*(1), 78–106.

Case, A., & Deaton, A. (2017). Mortality and morbidity in the 21st century. *Brookings papers on economic activity*, *2017*, 397.

Chandra, S., & Kassens-Noor, E. (2014). The evolution of pandemic influenza: Evidence from india, 1918–19. *BMC infectious diseases*, *14*(1), 1–10.

Clark, A. E., d'Ambrosio, C., & Lepinteur, A. (2021). The fall in income inequality during covid-19 in four european countries. *The Journal of Economic Inequality*, *19*, 489–507.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). *Labor markets during the covid-19 crisis: A preliminary view* (tech. rep.). National Bureau of economic research.

Correia, S., Luck, S., & Verner, E. (2022). Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. *The Journal of Economic History*, *82*(4), 917–957.

Crosby, A. W. (2003). *America's forgotten pandemic: The influenza of 1918*. Cambridge University Press.

Crossley, T. F., Fisher, P., & Low, H. (2021). The heterogeneous and regressive consequences of covid-19: Evidence from high quality panel data. *Journal of public economics*, *193*, 104334.

Dahl, C. M., Hansen, C. W., & Jensen, P. S. (2022). The 1918 epidemic and a v-shaped recession: Evidence from historical tax records. *The Scandinavian Journal of Economics*, *124*(1), 139–163.

Dang, H.-A. H., & Nguyen, C. V. (2021). Gender inequality during the covid-19 pandemic: Income, expenditure, savings, and job loss. *World Development*, *140*, 105296.

de Haan, J., & Sturm, J.-E. (2017). Finance and income inequality: A review and new evidence. *European Journal of Political Economy*, *50*, 171–195.

Deaton, A. (2021). *Covid-19 and global income inequality* (tech. rep.). National Bureau of Economic Research.

Eichenbaum, M. S., Rebelo, S., & Trabandt, M. (2021). The macroeconomics of epidemics. *The Review of Financial Studies*, *34*(11), 5149–5187.

Evans, T., Whitehead, M., Diderichsen, F., Bhuiya, A., & Wirth, M. (2009, September). *Challenging inequities in health: From ethics to action.* https://doi.org/10.1093/acprof:oso/9780195137408.001.0001

Ferreira, F. H. (2021). Inequality: In the time of covid-19 all metrics are not equal when it comes to assessing the pandemic's unequal effect. *Finance and Development*, 20–23.

Furceri, D., Loungani, P., Ostry, J. D., & Pizzuto, P. (2022). Will covid-19 have long-lasting effects on inequality? evidence from past pandemics. *The Journal of Economic Inequality*, 1–29.

Galasso, V., & Foucault, M. (2020). Working during covid-19: Cross-country evidence from real-time survey data.

Galletta, S., & Giommoni, T. (2022). The effect of the 1918 influenza pandemic on income inequality: Evidence from italy. *Review of Economics and Statistics*, *104*(1), 187–203.

Garrett, L., Chowdhury, A. M. R., & Pablos-Méndez, A. (2009). All for universal health coverage. *The Lancet*, *374*(9697), 1294–1299.

Gärtner, S., & Prado, S. (2016). Unlocking the social trap: Inequality, trust and the scandinavian welfare state. *Social Science History*, *40*(1), 33–62.

Gradstein, M., & Milanovic, B. (2002). Does liberte= egalite? a survey of the empirical links between democracy and inequality. *A Survey of the Empirical Links between Democracy and Inequality (August 2002).*

Gross, C. P., Essien, U. R., Pasha, S., Gross, J. R., Wang, S.-y., & Nunez-Smith, M. (2020). Racial and ethnic disparities in population-level covid-19 mortality. *Journal of general internal medicine*, *35*, 3097–3099.

Guimbeau, A., Menon, N., & Musacchio, A. (2022). Short-and medium-run health and literacy impacts of the 1918 spanish flu pandemic in brazil. *The Economic History Review.*

Guntupalli, A. M., & Baten, J. (2006). The development and inequality of heights in north, west, and east india 1915–1944. *Explorations in Economic History*, *43*(4), 578–608.

Hacıoğlu-Hoke, S., Känzig, D. R., & Surico, P. (2021). The distributional impact of the pandemic. *European Economic Review*, *134*, 103680.

Helgertz, J., & Bengtsson, T. (2019). The long-lasting influenza: The impact of fetal stress during the 1918 influenza pandemic on socioeconomic attainment and health in sweden, 1968–2012. *Demography*, *56*(4), 1389–1425.

Hong, S. C., & Yun, Y. (2017). Fetal exposure to the 1918 influenza pandemic in colonial korea and human capital development. *Seoul Journal of Economics*, *30*(4).

Johnson, N. P., & Mueller, J. (2002). Updating the accounts: Global mortality of the 1918-1920" spanish" influenza pandemic. *Bulletin of the History of Medicine*, 105–115.

Jordà, Ò., Singh, S. R., & Taylor, A. M. (2022). Longer-run economic consequences of pandemics. *Review of Economics and Statistics*, *104*(1), 166–175.

Karlsson, M., Nilsson, T., & Pichler, S. (2014). The impact of the 1918 spanish flu epidemic on economic performance in sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of health economics*, *36*, 1–19.

Lastunen, J., Rattenhuber, P., Adu-Ababio, K., Gasior, K., Jara, H. X., Jouste, M., McLennan, D., Nichelatti, E., Oliveira, R. C., Pirttilä, J., et al. (2021). *The mitigating role of tax and benefit rescue packages for poverty and inequality in africa amid the covid-19 pandemic* (tech. rep.). WIDER Working Paper.

Lin, M.-J., & Liu, E. M. (2014). Does in utero exposure to illness matter? the 1918 influenza epidemic in taiwan as a natural experiment. *Journal of health economics*, *37*, 152–163.

Mahler, D. G., Yonzan, N., & Lakner, C. (2022). The impact of covid-19 on global inequality and poverty.

Mamelund, S.-E. (2006). A socially neutral disease? individual social class, household wealth and mortality from spanish influenza in two socially contrasting parishes in kristiania 1918–19. *Social science & medicine*, *62*(4), 923–940.

Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M., & Cetron, M. S. (2007). Nonpharmaceutical interventions implemented by us cities during the 1918-1919 influenza pandemic. *Jama*, *298*(6), 644–654.

McGregor, T., Smith, B., & Wills, S. (2019). Measuring inequality. *Oxford Review of Economic Policy*, *35*(3), 368–395.

McLaren, J. (2021). Racial disparity in covid-19 deaths: Seeking economic roots with census data. *The BE Journal of Economic Analysis & Policy*, *21*(3), 897–919.

Milanovic, B. (2020). Inequalities and covid-19 [Accessed: 2021-04-22, http://glineq.blogspot.com/search?q=Inequalities+and+Covid-19.].

Mills, I. D. (1986). The 1918-1919 influenza pandemic—the indian experience. *The Indian Economic & Social History Review*, *23*(1), 1–40.

Murray, C. J., Lopez, A. D., Chin, B., Feehan, D., & Hill, K. H. (2006). Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: A quantitative analysis. *The Lancet*, *368*(9554), 2211–2218.

Neelsen, S., & Stratmann, T. (2012). Long-run effects of fetal influenza exposure: Evidence from switzerland. *Social Science & Medicine*, *74*(1), 58–66.

Nelson, R. E. (2010). Testing the fetal origins hypothesis in a developing country: Evidence from the 1918 influenza pandemic. *Health economics*, *19*(10), 1181–1192.

Ogasawara, K. (2017). Persistence of pandemic influenza on the development of children: Evidence from industrializing japan. *Social Science & Medicine*, *181*, 43–53.

Olson, D. R., Simonsen, L., Edelson, P. J., & Morse, S. S. (2005). Epidemiological evidence of an early wave of the 1918 influenza pandemic in new york city. *Proceedings of the National Academy of Sciences*, *102*(31), 11059–11063.

Patterson, K. D., & Pyle, G. F. (1991). The geography and mortality of the 1918 influenza pandemic. *Bulletin of the History of Medicine*, *65*(1), 4–21.

Peckham, R. (2013). Economies of contagion: Financial crisis and pandemic. *Economy and Society*, *42*(2), 226–248.

Percoco, M. (2016). Health shocks and human capital accumulation: The case of spanish flu in italian regions. *Regional Studies*, *50*(9), 1496–1508.

Radatz, L., & Baten, J. (n.d.). Measuring multidimensional inequality and conflict in africa and in a global comparison.

Reich, M. R., & Goldman, R. H. (1984). Italian occupational health: Concepts, conflicts, implications. *American journal of public health*, *74*(9), 1031–1041.

Rota, M., & Weisdorf, J. (2020). Italy and the little divergence in wages and prices: New data, new results. *The Journal of Economic History*, *80*(4), 931–960.

Souza, C. M. C. d. (2005). The spanish flu in salvador, 1918: City of alleys and tenements. *História, Ciências, Saúde-Manguinhos*, *12*, 71–99.

Spreeuwenberg, P., Kroneman, M., & Paget, J. (2018). Reassessing the global mortality burden of the 1918 influenza pandemic. *American journal of epidemiology*, *187*(12), 2561–2567.

Stiglitz, J. (2020). Conquering the great divide. *Finance & Development*, *57*(3), 17–19.

van Zanden, J. L., Baten, J., Foldvari, P., & van Leeuwen, B. (2014). The changing shape of global inequality 1820–2000; exploring a new dataset. *Review of income and wealth*, *60*(2), 279–297.

Vollmer, S., & Wójcik, J. (2017). The long-term consequences of the global 1918 influenza pandemic: A systematic analysis of 117 ipums international census data sets. *Available at SSRN 3083584*.

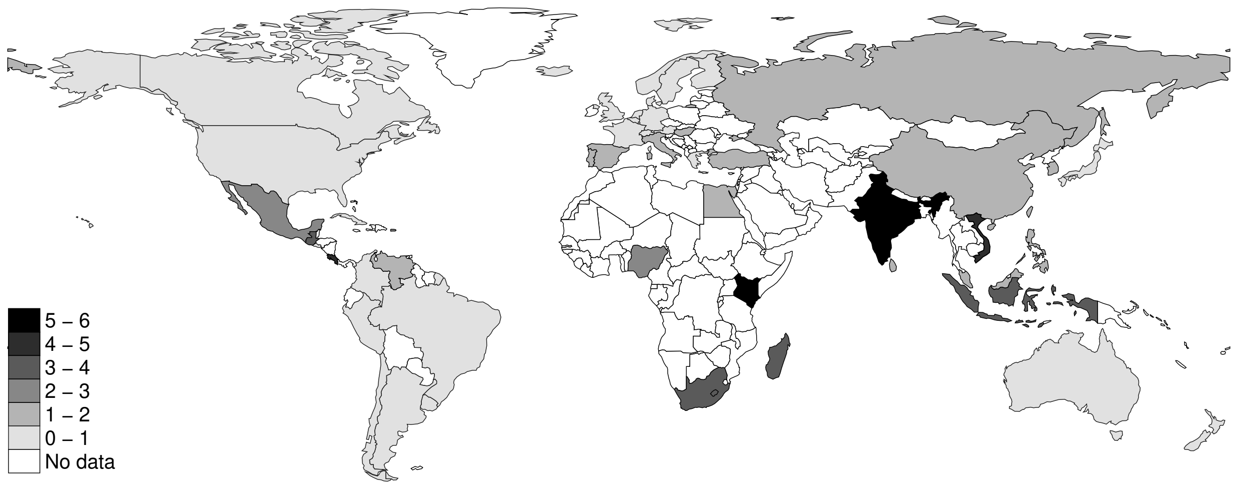
Weil, D. N. (2014). Health and economic growth. In *Handbook of economic growth* (pp. 623–682, Vol. 2). Elsevier.

Wheelock, D. C. (2020). What can we learn from the spanish flu pandemic of 1918-19 for covid-19? *Available at SSRN 3605782*.

World Bank. (2018). Poverty and shared prosperity 2018: Piecing together the poverty puzzle.

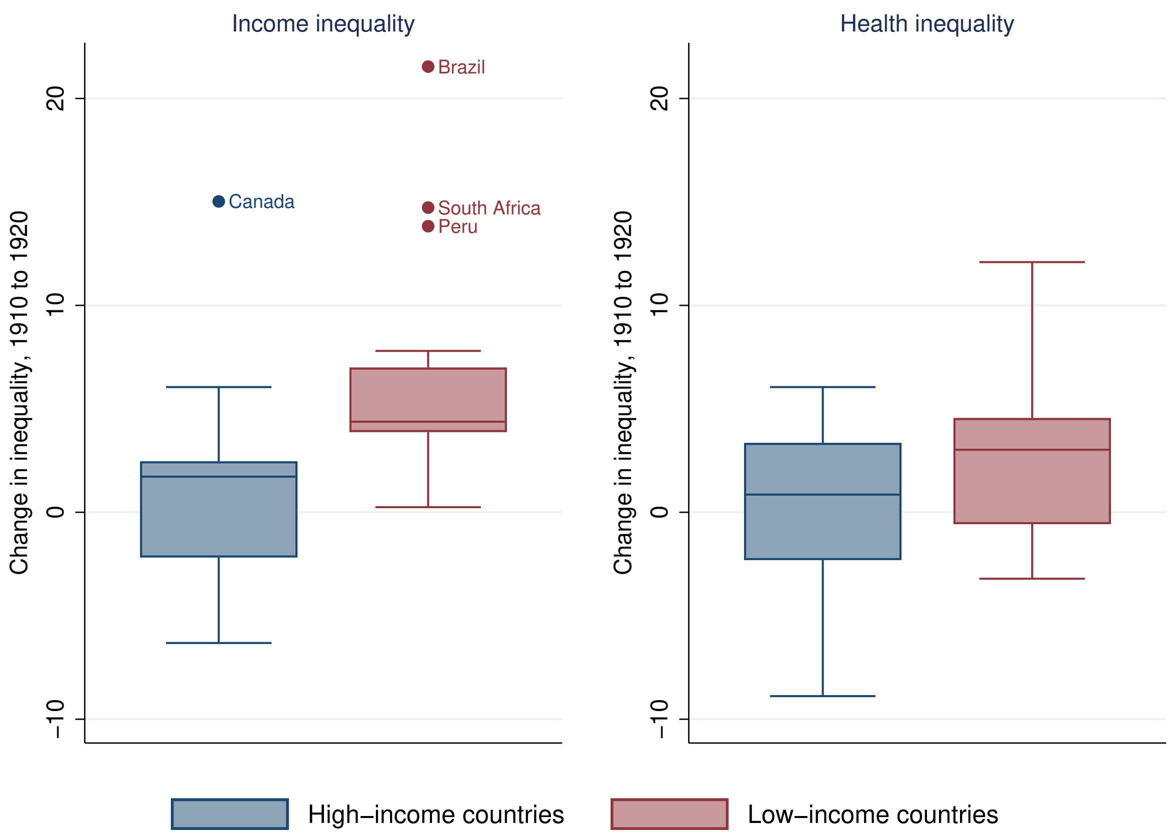
C.9 Tables and figures

Figure C.1: Severity of the influenza pandemic 1918



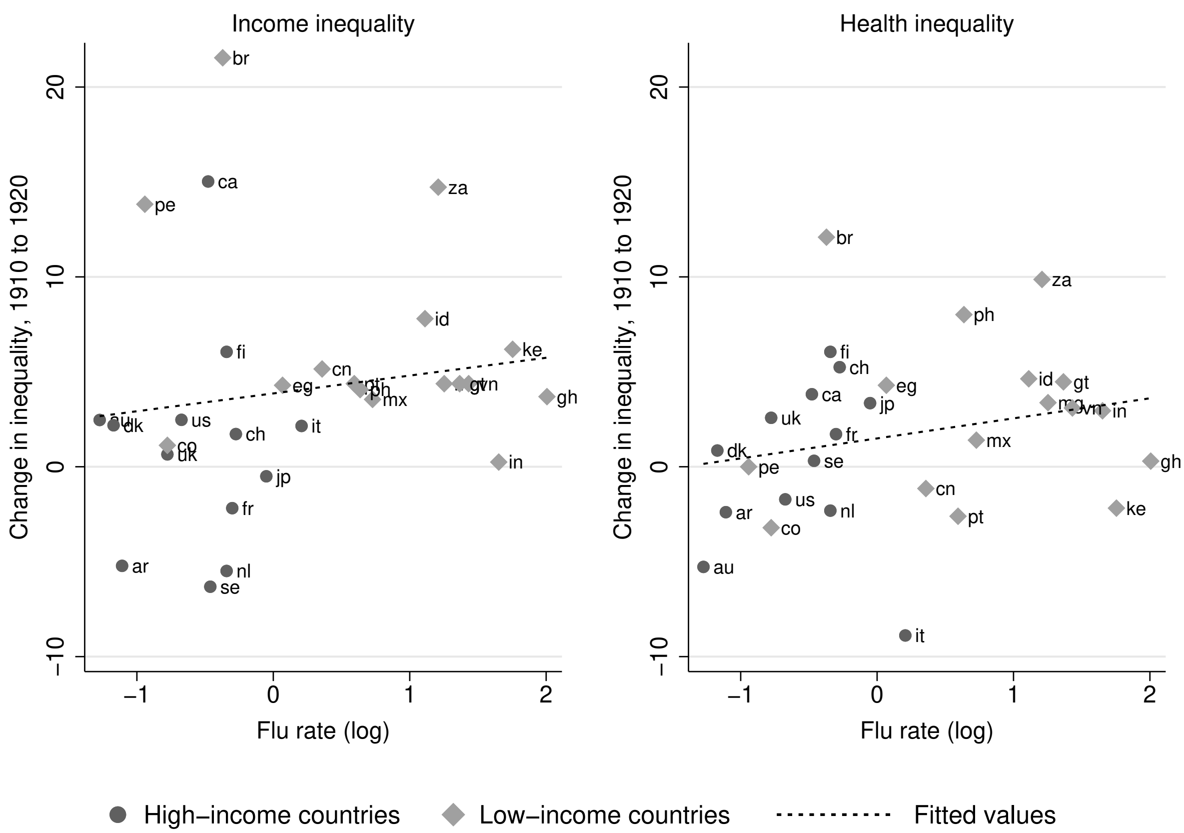
Notes. The severity of the 1918 influenza pandemic is based on the flu mortality rate and is calculated per 1,000 inhabitants. The flu mortality rate is divided into six categories, where darker shades indicate a higher severity of the 1918 influenza pandemic, and lighter grey indicates a lower severity. Sources for flu mortality rates include Barro et al. (2020), Patterson and Pyle (1991), and manually collected data from different national statistical reports.

Figure C.2: Differences in inequality between the (birth) decade 1910 and 1920 by income group



Notes. We show the difference in income inequality (left) and health inequality (right) between the (birth) decade 1910 and 1920. Inequality is measured by the Gini coefficient. A positive value indicates an increase in inequality between the 1910s and the 1920s. Countries are further divided into income groups based on the World Bank classification and refer to national GDP in 1910.

Figure C.3: Scatterplot between inequality and excess flu mortality rate



Notes. We display the correlation between the flu mortality rate of the 1918 influenza pandemic and inequality in income (left) and health (right). Inequality is measured by the change in the Gini coefficient between the (birth) decade 1910 and 1920. A positive value indicates an increase in inequality between the 1910s and the 1920s. Countries are further divided into income groups based on the World Bank classification and refer to national GDP in 1910.

Table C.1: Summary statistics

	Obs	Mean	Std. Dev.	Min	Max
Income inequality	29	4.02	6.14	-6.318542	21.5399
Health inequality	29	1.68	4.48	-8.88604	12.09173
Flu rate (log)	29	0.17	0.97	-1.272966	2.005974
WW1 deaths	29	0.21	0.44	0	1.35
Population density	29	49.46	61.01	.7478	215.3694
Urbanisation	27	0.02	0.01	-.0117543	.0407516
Years of schooling	27	0.48	0.31	.035	1.065
Drought	23	0.05	0.20	-.3500004	.7399998
GDP per capita	23	566.09	482.77	-364.9	1852
Democracy	21	0.41	4.01	-10.4	9.9
Democracy^2	21	12.20	29.97	-69.3	75.8

Notes. All variables are measured as differences between the (birth) decade 1910 and 1910. Positive values indicate an increase from the 1910s to the 1920s. We only include observations for which our main variables, inequality and flu rate, are available.

Table C.2: First-difference estimation

	Income inequality		Health inequality	
	(1)	(2)	(3)	(4)
Flu rate	0.94	0.80	1.06	0.68
	(1.080)	(1.248)	(0.803)	(0.935)
WW1 deaths		−0.90		−2.51
		(3.056)		(2.711)
Constant	3.86***	4.08***	1.50*	2.10**
	(1.307)	(1.553)	(0.787)	(0.940)
Observations	29	29	29	29
Adj. R-squared	-0.014	-0.049	0.017	0.037

Notes. Bootstrap country-clustered standard errors in parentheses. ***, **, *, significant on the 1, 5, and 10%-level, respectively. Flu rate is calculated as the excess mortality rate for the 1918 influenza pandemic. All variables are calculated as the first difference between the (birth) decade 1910 and 1920.

Table C.3: First-difference estimation including controls

	Income inequality						Health inequality					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Flu rate	2.09*	1.49*	0.36	0.92	1.08	2.26	1.20	0.95	1.28	0.88	2.40	3.70
	(1.144)	(0.850)	(2.104)	(5.212)	(14.533)	(51.434)	(0.799)	(0.768)	(1.496)	(2.965)	(4.619)	(17.030)
WW1 Deaths	3.99	1.50	0.81	0.53	−1.99	−3.42	−1.01	−2.25	−2.21	−2.08	−5.18	−6.75
	(4.437)	(3.542)	(3.099)	(5.052)	(15.099)	(40.907)	(2.509)	(2.326)	(3.108)	(5.685)	(10.807)	(7.889)
Population density	−0.01	−0.02	−0.02	−0.02	−0.05	−0.06	0.01	0.00	0.00	−0.00	−0.01	−0.02
	(0.015)	(0.017)	(0.020)	(0.042)	(0.130)	(0.234)	(0.009)	(0.010)	(0.017)	(0.025)	(0.039)	(0.227)
Urbanisation		2.37**	2.54**	2.24	3.31	2.42		1.08*	1.21*	1.52	1.52	0.55
		(0.953)	(0.999)	(1.480)	(3.003)	(12.376)		(0.641)	(0.726)	(1.469)	(1.537)	(5.903)
GDP per capita			−1.22	−1.41	−3.87	−4.68			−0.58	−1.88	−1.35	−2.25
			(2.390)	(5.081)	(8.875)	(18.504)			(2.535)	(4.437)	(4.517)	(14.466)
Democracy				−0.43	1.04	0.93				0.51	1.97	1.84
				(1.889)	(7.726)	(5.832)				(2.400)	(2.732)	(6.925)
Democracy^2				0.05	0.27	0.27				0.01	0.23	0.24
				(0.180)	(0.595)	(2.304)				(0.214)	(0.357)	(0.508)
Drought					15.15	15.80					8.53	9.25
					(49.067)	(92.994)					(17.699)	(66.946)
Years of schooling						5.66						6.25
						(19.632)						(16.157)
Observations	29	27	21	19	17	17	29	27	21	19	17	17
Adj. R-squared	-0.002	0.311	0.199	0.106	0.294	0.250	-0.014	0.104	0.017	-0.110	0.123	0.151

Notes. Bootstrap country-clustered standard errors in parentheses. ***, **, *, significant on the 1, 5, and 10%-level, respectively. Flu rate is calculated as the excess mortality rate during the 1918 influenza pandemic. All variables are calculated as the changes between the birth decade 1910 (1910–1919) and 1920 (1920–1929). For interpretation, *urbanisation* is multiplied by 100, and GDP p.c. is divided by 1,000.

Table C.4: First-difference estimation by income group

	Income inequality				Health inequality			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low-income countries		*High-income countries*		*Low-income countries*		*High-income countries*	
Flu rate	1.75	1.06	0.24	0.85	1.58*	1.18	−0.53	0.45
	(1.444)	(1.471)	(2.843)	(1.925)	(0.832)	(1.335)	(2.453)	(2.102)
WW1 deaths	82.08	52.87**	2.41	2.39	16.12	3.95	−1.18	−2.37
	(64.270)	(20.315)	(4.193)	(2.199)	(87.295)	(44.756)	(2.756)	(2.689)
Population density		−0.04*		−0.02*		−0.01		0.00
		(0.021)		(0.011)		(0.033)		(0.011)
Urbanisation		2.99*		1.33		1.11		0.98*
		(1.423)		(1.077)		(0.871)		(0.503)
Observations	16	14	13	13	16	14	13	13
Adj. R-squared	0.082	0.355	-0.074	0.022	0.034	-0.023	-0.143	-0.171

Notes. Bootstrap country-clustered standard errors in parentheses. ***, **, *, significant on the 1, 5, and 10%-level, respectively. Flu rate is calculated as the excess mortality rate during the 1918 influenza pandemic. All variables are calculated as the changes between the birth decade 1910 (1910–1919) and 1920 (1920–1929). For interpretation, *urbanisation* is multiplied by 100, and GDP p.c. is divided by 1,000. Income groups are based on the world average and a country's GDP p.c. for the reference year of 1910, based on Maddison Project Database, version 2020.

Table C.5: Robustness checks

	Income inequality		Health inequality	
	(1)	(2)	(3)	(4)
Omitted = observations for India				
Flu rate	2.23*	2.26	1.24	3.70
	(1.182)	(14.894)	(0.800)	(9.042)
WW1 deaths	3.51	−3.42	−0.32	−6.75
	(2.146)	(33.125)	(1.930)	(11.303)
Population density		−0.06		−0.02
		(0.573)		(0.081)
Urbanisation		2.42		0.55
		(15.701)		(4.455)
GDP per capita		−4.68		−2.25
		(31.818)		(12.245)
Democracy		0.93		1.84
		(20.004)		(190.727)
$Democracy^2$		0.27		0.24
		(1.485)		(9.504)
Drought		15.80		9.25
		(94.620)		(88.126)
Years of Schooling		5.66		6.25
		(74.671)		(31.765)
Observations	28	17	28	17
Adj. R-squared	0.036	0.250	-0.008	0.151

Notes. Bootstrap country-clustered standard errors in parentheses. ***, **, *, significant on the 1, 5, and 10%-level, respectively. Flu rate is calculated as the excess mortality rate during the 1918 influenza pandemic. We exclude India from all models. All variables are calculated as the changes between the birth decade 1910 (1910–1919) and 1920 (1920–1929).

C.10 Appendix

C.10.1 Variable description and sources

The following list provides an overview of the variables and sources used in this study:

Income inequality. Measured as Gini coefficient from different sources. Source: van Zanden et al. (2014) and Alvaredo et al. (2022).

Health inequality. Health inequality is proxied by height Gini, based on the distribution of heights within a country. Source: Baten and Blum (2015) available via the website of Clio Infra; and the extension done by Radatz and Baten (n.d.).

Flu rate. Our main independent variable is calculated as excess flu mortality. It is measured per 1,000 inhabitants. Sources: Barro et al. (2020), Patterson and Pyle (1991), and own calculation for Chile, Costa Rica, Cuba, Jamaica and Puerto Rico, based on Murray et al. (2006). The main sources for total deaths and death rates are national statistical yearbooks.

WW1 deaths. Estimated severity of the First World War for each country, for the war years 1914–1918, expressed as a percentage of the population. Source: Barro et al. (2020).

Population density. A country's population divided by total land area. Source: Gapminder (Version 6) https://www.gapminder.org/data/documentation/gd003/ and FAO (2022), derived from World Development Indicators, The World Bank (2022).

Urbanisation. Degree of the urban population to the total population in a country. Source: Fink-Jensen (2015), available via Clio Infra.

Years of schooling. To control for the educational level we include average years of schooling by country in our regression analysis. Source: Lee and Lee (2016), Barro and Lee (2013), and UNDP (2018); derived from Roser and Ortiz-Ospina (2016).

Drought. We control for climate conditions with our variable drought. Data on drought is based on monthly data from the CRU TS dataset, version 4. Source: Harris et al. (2020).

GDP p.c. Expressed as GDP per capita. Source: Bolt and Van Zanden (2020).

Democracy. The variable *democracy* accounts for the regime type in a country and is based on the polity2 index from the Polity5 project. On a scale from -10 to +10 it distinguishes between different levels of democracies and autocracies. Source: Marshall and Gurr (2020).

C.10.2 Country coverage

We include the following countries in our regression analysis. Countries are further divided into two income groups, high-income and low-income, based on the World Bank classification, referring to national GDP in 1910.

High-income countries: Argentina (ar), Australia (au), Canada (ca), Denmark (dk), Finland (fi), France (fr), Italy (it), Japan (jp), Netherlands (nl), Sweden (se), Switzerland (ch), United Kingdom (uk), United States (us).

Low-income countries: Brazil (br), China (cn), Colombia (co), Egypt (eg), Ghana (gh), Guatemala (gt), India (in), Indonesia (id), Kenya (ke), Madagascar (mg), Mexico (mx), Peru (pe), Philippines (ph), Portugal (pt), South Africa (za), Vietnam (vn).

C.10.3 References

Alvaredo, F., Atkinson, A. B., Piketty, T., & Saez, E. (2022). World inequality database. http://wid.world/data

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, *104*, 184–198.

Barro, R. J., Ursúa, J. F., & Weng, J. (2020). *The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity* (tech. rep.). National Bureau of Economic Research.

Baten, J., & Blum, M. (2015). Height gini. http://hdl.handle.net/10622/HKZPR9

Bolt, J., & Van Zanden, J. L. (2020). Maddison style estimates of the evolution of the world economy. a new 2020 update. *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands.*

Fink-Jensen, J. (2015). Urbanization ratio. http://hdl.handle.net/10622/LZ0Y36

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the cru ts monthly high-resolution gridded multivariate climate dataset. *Scientific data*, *7*(1), 109.

Lee, J.-W., & Lee, H. (2016). Human capital in the long run. *Journal of development economics*, *122*, 147–169.

Marshall, M. G., & Gurr, T. R. (2020). Polity5: Political regime characteristics and transitions, 1800-2018. *Center for Systemic Peace.*

Murray, C. J., Lopez, A. D., Chin, B., Feehan, D., & Hill, K. H. (2006). Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: A quantitative analysis. *The Lancet*, *368*(9554), 2211–2218.

Patterson, K. D., & Pyle, G. F. (1991). The geography and mortality of the 1918 influenza pandemic. *Bulletin of the History of Medicine*, *65*(1), 4–21.

Radatz, L., & Baten, J. (n.d.). Measuring multidimensional inequality and conflict in africa and in a global comparison.

Roser, M., & Ortiz-Ospina, E. (2016). Global education [https://ourworldindata.org/global-education]. *Our World in Data.*

United Nations Development Programme. (2018). Statistical update 2018. *UNDP (United Nations Development Programme).* http://report2017.archive.s3-website-us-east-1.amazonaws.com

van Zanden, J. L., Baten, J., Foldvari, P., & van Leeuwen, B. (2014). The changing shape of global inequality 1820–2000; exploring a new dataset. *Review of income and wealth*, *60*(2), 279–297.

D The Heights of Medical Care: Health Insurance and Inequality in Adult Stature[32]

Abstract

Reforms that help to facilitate access to necessary healthcare can have a positive impact on health outcomes. This is especially the case for the expansion of health insurance, which reduces information and financial barriers to health care and increases the likelihood that poorer people in particular will benefit from new health programmes. This paper explores the question by taking an anthropometric measure of health and well-being—namely human height—and examining whether expanded access to health insurance promotes body growth towards its height potential for those in poorer health in a society. We test the reduction of within-country height inequality following the introduction of health insurance programs. We draw our evidence using a panel of countries for which we could measure height inequality retrospectively for several decades. Controlling for a number of relevant control variables, we produce evidence suggesting that indeed within-country differences in height inequality declined after insurance expansion.

[32] This chapter was co-authored with Joerg Baten, Alberto Batinti and Joan Costa-Font.

D.1 Introduction

The expansion of health insurance to a larger proportion of the population and its impact on the use of healthcare services was crucial to human health in the nineteenth and twentieth centuries. Following Bismarck's expansion of the number of people being health insured in Germany to broaden the risk pool (Busse et al., 2017), the proliferation of state-funded health insurance in Denmark, Belgium, Norway, and the United Kingdom (UK) at the end of the 19th and the beginning of the 20th century (Gorsky, 2011) significantly reduced the costs of access to health technologies. It has also reduced the cost of access to information gatekeepers, such as physicians, who influence individuals' behaviour and health through preventive and curative care (Freeman et al., 2008). This positive influence is especially true for unskilled or low-income individuals who would otherwise be uninsured (Finkelstein et al., 2012).

Insurance expansions have been followed by improvements in several health outcomes: infant and child mortality, for example, have been shown to decrease after health insurance expansions; Currie and Gruber (1996) document evidence of such effects after the implementation of Medicaid in the United States (US). Goodman-Bacon (2021) compared birth cohorts before and after the introduction of Medicaid in different US states. He finds that if children are eligible for Medicaid in the early years, they are significantly healthier and economically better off as adults. This is because an individual's health in childhood is determined by parental actions guided by the recommendations of doctors as diet, vaccination, or treatments for specific conditions, and can be detrimental to a person's health later in life (Quinn & Woolley, 2001). Furthermore, health insurance gives people more financial stability, which reduces stress and improves well-being (Haushofer et al., 2020).

To date, however, the literature has documented often-ambiguous effects of insurance expansion whether through the state, the market, mutual or employment, on the health of adults and children. Of particular interest is the effect of insurance expansion on individuals that otherwise would have undergone healthcare difficulties. Hence, in addition to improving health for the whole population, health insurance expansion is arguably likely to influence health inequalities. Improving the health of those with the worst health conditions could, therefore, reduce health disparities in a population.

Examining the effect of health insurance introduction on a measure of inequality in health, namely the Gini coefficient of human heights, is the main purpose of this paper. The advantage of human height is that it allows for undertaking an unambiguous analysis of health over the past 200 years. Most insurance expansions across countries have taken place at a time when there was little access to health indicators to measure the effects of insurance expansions, which limited the analysis of insurance on some well-defined health measures such as mortality (Bauernschuster et al., 2020). One way to circumvent access to such data is the analysis of retrospective heights of individuals, a measure very sensitive to the improvement of standards of living, and early life health investments. In addition, adult height is fairly stable over a lifetime, once it has been reached and before one starts to shrink, which can happen around the age of fifty (Beard & Blaser, 2002; Stinson, 1985). The distribution of heights within a country, therefore, provides an estimate of health that does not suffer from omitted variable bias, which is often problematic with other health measures.

This paper draws on a rich set of data measuring health insurance expansion and height inequality, and we construct a large sample of countries where we have access to both data. Health insurance coverage reflects whether there is public or private healthcare insurance for all people living in the country, as defined by the World Health Organization (WHO, 2010). Our focus is the presence of Universal Health Coverage (UHC) in a country. Constructing a binary variable, our measure of health insurance indicates whether a country, first, has legislated for social insurance and, second, has achieved coverage for more than 90% of its population. We draw on instrumental variable (IV) models to assess whether there was a positive effect of the health insurance expansion on a reduction in height inequality. As a preview of our findings, we observe a positive and substantial reduction in height inequality. The results are robust to a number of robustness checks.

We contribute to the literature by examining retrospectively the effect of major health insurance expansions for which we did not have evidence back in time. Second, we focus on measuring inequalities in health, which have received limited attention in the literature. Finally, we draw on height inequality, which measures the dispersion in the use of human stature, a measure that is not affected by the traditional problems of self-report bias that health measures exhibit.

The paper is organised as follows. Section D.2 reports the background on health insurance's effect on health and discusses its potential mechanisms. We then report the data and empirical strategy. Section D.4 shows the descriptive statistics and results. Section D.5 presents the results from the instrumental variable estimation and Section D.6 the robustness checks. The final section provides the conclusion.

D.2 Health insurance, inequality and heights

Inequality and heights. Heights are commonly employed as a proxy for health and nutritional quality, and health determinants associated with living conditions. It correlates with a series of health measures (Fogel, 1994; Steckel, 1995). Like other measures of health, it predicts the economic performance of individuals, such as income or wages (Persico et al., 2004; Strauss & Thomas, 1998), which are highly dependent on cognitive abilities (Case et al., 2002). However, socioeconomic circumstances can influence individual heights, and therefore height disparities in a population (Kuh & Wadsworth, 1989; Li et al., 2004). For example, Candela-Martínez et al. (2022) document a reduction of height differences by educational attainment in Spain, in the period 1940 to 1994 when healthcare was universalised, and the welfare state exhibited significant development.

Insurance, health and inequality. Insurance minimises the risk of unexpected medical costs that individuals or households would instead have to bear. When costs are unaffordable, people forego healthcare, which can have negative consequences for their health. Thus, increasing health insurance coverage would lead to large utility gains for households through the reduction of the uncertainty and variable health-related expenditures (Einav & Finkelstein, 2018). The effect of health insurance on health is not clear from the literature: Finkelstein et al. (2012) document causal evidence that public health insurance expansion improves self-reported health and mental healthcare among low-income individuals in the US who randomly qualified for a Medicaid expansion in Oregon. The effect on other objective measures of health, however, is not significant. Costa-Font et al. (2021) find that the expansion of public insurance in Mexico has failed to reduce health inequality and mobility of individuals across the health distribution. Nevertheless, several related studies report evidence that is highly context-dependent, and incremental compared to earlier reforms. The effects of these reforms cannot be measured in such as straightforward manner with existing evidence. Evidence from China suggests that, while health insurance is linked to reduced health disparities, this effect appears to be mainly due to circumstances, which are not related to the health care system (Wang & Yu, 2016). The exception is Bauernschuster et al. (2020), who show that the introduction of insurance in 1884 in Bismarck's Germany accounts for a decrease in mortality between 24% and 45% across blue-collar occupations affected by the reform.

Income and access to healthcare. Insurance expansions can exert income effects, as the absence of insurance entails savings to pay for health care, which can be invested elsewhere in the presence of insurance. In market-based systems, increasing income and improvements in medical technology led to the expansion of the market (Thomasson, 2002). Besides, increasing individuals' income by reducing the need for preventive savings, insurance can reduce the barrier to accessing healthcare. Indeed, the introduction of health insurance with the provision of high-quality treatments can contribute to improved health. This is particularly important for children, as access to healthcare can benefit children's health directly through increased usage of high-value preventions (e.g., health growth check-ups) or vaccinations, and indirectly through behavioural changes (e.g., breastfeeding), as outlined in the following.

Information and preventive effects. Insurance expansion can be important in promoting the uptake of preventable behaviours. People are more likely to adopt highly valuable health behaviours when they receive such information from a source they trust, for example, from healthcare providers (Ellis & Manning, 2007). Insurance could reduce the impact of cognitive biases that lead to negative behavioural risks (Baicker et al., 2015), and to the underuse of health care. This would allow demand to be aligned with the marginal utility of an individual.

Hence, access to affordable insurance combined with the provision of high-value health information can give rise to improvement in human health, especially for those who were in poorer health before such insurance expansions. As a result, we expect health insurance to reduce inequalities in heights within a population. The rest of the paper will be devoted to documenting this question.

D.3 Data and methods

D.3.1 Estimating inequality using the Gini coefficient of height

A common measure of inequality of a continuous variable usually is the Gini coefficient, which is typically used to measure income inequality. However, for a large list of countries in the world, evidence of the period before the 1980s or 1990s does not exist and is often inconsistent for other countries. Recent studies have used the coefficient of variation of height as a proxy indicator, as well as the Gini coefficient in height, calculated from it. Below we discuss the idea, its origin, potential caveats, advantages, and the measurement procedure in detail (this discussion draws on Baten, 1999; Baten and Blum, 2011; Moradi and Baten, 2005, and Baten and Mumme, 2013).

We re-evaluate the indicator function of average height before estimating a measure of height inequality. The average adult height is commonly accepted as an indicator of biological well-being (Fogel et al., 1982; Komlos, 1985; Steckel, 1995). Human stature grows at the fastest rate during the first three years of life; hence we aggregate adult heights by birth cohorts (Baten, 1999; Eveleth & Tanner, 1976). This topic has been discussed in the anthropometric literature and will not be discussed further here (Baten, 2000a; Komlos, 1985; Steckel, 1995; on height inequality as in Baten, 1999). According to these studies, genetic factors have a distinct impact on height at the *individual level*, whereas *population averages* of height are influenced by health conditions and diet quality. Tall parents have tall children for genetic reasons, but genetics have less influence at the level of population averages because individual genetic differences average out. Consider the following examples of population-level differences in history: For example, during a period of severe protein deficiency in mid-nineteenth-century Holland, Dutch people were very short by European standards, whereas today they are frequently regarded as the tallest people on the planet (Baten & Blum, 2014). While earlier anthropologists attributed many size patterns (e.g., tall Tutsi and Massai) to genetics, these patterns were later identified to be the result of dietary quality and a healthy environment (Bogin, 1988).

Once we have established that stature is an indicator of average dietary quality and health, we now discuss stature inequality. We estimate inequalities in the standard of living of a population within a given birth decade as the Gini coefficient, using the coefficient of variation in height (hereafter CV). Baten (1999, 2000b) contends that the CV is a good indicator of income inequality within society (see also Moradi and Baten, 2005; van Zanden et al., 2014), because the two measures are correlated. To understand the influence of inequality on height, we compare outcomes of a notional situation, where a population is subjected to the alternative distribution of resources, (A) and (B), after birth (the following is based on Moradi and Baten, 2005):

A. Every individual gets the same amount and quality of resources (nutrition and health services). This setting constitutes a condition of perfect equality.

B. The resources are unequally distributed (yet independent of the genetics of an individual).

Situation (A) reflects a biological variance in (normally distributed) stature since the size distribution should only reflect the genetics. But what happens to the distribution of heights as inequality increases from case (A) to case (B)? Given the unequal distribution of resources for nutrition, health, and shelter, some people benefit and grow taller, while other individuals grow smaller as they endure poor dietary status. As a result, when compared to the scenario of perfect equality, the richer classes' individual height shifts to the right, while the poorer classes shift to the left. As a consequence, increasing inequality should result in greater inequality in height. If resource endowments differ greatly between groups, it may even result in a bimodal size distribution. Even though biological variance still accounts for a large proportion of total variance, most size distributions follow an (almost) normal distribution, albeit with a larger standard deviation than in theory (A). Yet, the biological variance is thought to increase with average stature, hence the standard deviation of stature is an ineffective measure of inequality (Schmitt & Harrison, 1988). This effect is accounted for by the CV, which is a reliable and consistent measure of inequality. For a ten-year birth decade t and country i, the CV is defined as follows:

$$CV_{it} = \frac{\sigma_{it}}{\mu_{it}} * 100. \qquad (1)$$

Therefore, the standard deviation σ_{it} is expressed as a percentage of the mean μ_{it}. Baten (1999, 2000a) used the CV to compare size differences between social groups in the early 19$^{\text{th}}$ century in the southern region of Germany, Bavaria, because an ideal dataset was available for this period and region, covering almost all men in the population. Furthermore, it even included the socioeconomic status of all parents. Height and socioeconomic status were found to be highly correlated. As such, high CVs adequately reflect social and economic circumstances without the need for classifications. Using the CV, Moradi and Baten (2005) developed a formula, adequately transforming the CV values in height Gini coefficient, which was already widely used as an inequality indicator in empirical studies (e.g. see review in Blum, 2014), and which we are also using in this study.

Most estimates of the relationship between the Gini coefficient of height and the income Gini have been done for the post-1950s when real income increased in many countries to unprecedented levels. This could result in a biased correlation between the height Gini and the income Gini coefficient downward since, in many world regions, the proportion of income needed for housing and food declined in the following years. Overall, the Gini coefficients of income and height appear to be quite strongly correlated (for a recent review, see Radatz and Baten, n.d.).

Our sample is therefore based on the data collection of height from Baten and Blum (2011) publicly available on the website of Clio Infra,[33] and the extension done by Radatz and Baten (n.d.). Data on height inequality is calculated as the country-level Gini coefficient. Sources include several national surveys for early decades and data from international household surveys such as the Demographic and Health Surveys (DHS), especially for developing countries. A more detailed overview of countries and respective sources can be found on the website of Clio Infra or in Radatz and Baten (n.d.). We use data on Gini-height for 134 countries with *height inequality* being our dependent variable. Our sample covers the birth decades from 1810 to 2000, where each decade includes the following 10 years, i.e., 2000 represents the year 2000 throughout the end of 2009.

D.3.2 Health insurance

The World Health Organization (WHO) distinguishes between Primary Health Care (PHC) as the first stage and UHC as the second stage of public health. UHC is defined as access to the national health system for all people living in the country. This health system may be publicly and/or privately funded (WHO, 2010).

[33]Data on height and height Gini is publicly available on the website of Clio infra: https://clio-infra.eu/. However, the updated version is not yet available (as of 04/04/2023).

However, there is no official list of countries fulfilling the WHO definition of UHC based on explicit criteria (see Stuckler et al., 2010). There does exist an official indicator from the WHO, namely the UHC service coverage index, for the percentage of the population covered (WHO, 2023), and an indicator for social health protection from the International Labour Organization (ILOSTAT, 2022). Both indicators provide us with important but limited data for the period 2000–2019 in the case of the UHC service coverage index, and a snapshot of the year 2020 for the ILO index. The OECD (2022) provides a dataset on social protection for OECD countries from 1960 onwards, which calculates the percentage of the population covered by both public and private health care systems within a country. The Varieties of Democracy (V-Dem) provides a huge dataset on different aspects of democracies. They include aspects of health as an important component of a democratic state. More specifically, the database includes indicators (e.g., *v2pehealth* and *v2peapsecon*) on the access to basic health care for the population from 1900 to 2022 (Coppedge et al., 2022). Given that sufficient observations are available, we use some of these measures to compare and check the robustness of our results (see Section D.6).

Historically, the first introduction of a health insurance system in a country was in Germany, beginning in 1883, quickly followed by other Western European countries in the 1880s, namely Austria, the Czech Republic, Hungary, and Slovakia. Many countries implemented health insurance systems in the decades that followed, primarily in and after the 1940s. However, differences between world regions remain significant, despite a process of harmonisation. Nonetheless, it is clear that the introduction of health insurance systems, particularly in developing world regions such as Africa, occurred much later, raising concerns about the stability and robustness of the most recent health systems. In South Sudan, for example, there is a government health service that is supposed to be free and accessible. However, only 32% of the population seems to be covered in 2019 (World Health Organization & World Bank [WHO and World Bank], 2021). The situation is similar in Nigeria, where the National Health Insurance Scheme (NHIS), introduced in 1999, was supposed to provide universal coverage, but due to shortcomings in recent years, free access and good quality are still quite limited (Makinde et al., 2018).

These examples show that the mere formal introduction of health insurance may not be sufficient to produce observable positive effects on population health. Therefore, the presence of UHC within a country is our primary variable of interest. We are interested in the timing of the legal introduction of health insurance in a country and whether this health care system actually covers the entire population. To construct our UHC variable, we manually collected data for the legal implementation of health insurance in a country from a variety of sources, with the main sources being Cutler and Johnson (2004), Kangas (2012), and the 'Social Security Programs Throughout the World' publication series published by the Social Security Administration (SSA) (see Appendix for more details). We also consider whether there is an actual implementation of a country's population by the UHC, mainly based on country reports and the OECD indicator on social protection. We include countries that have legally mandated UHC and have already completed the transition to (near) full population coverage. We define the threshold as population coverage of 90% or more with the reference year 2010. More specifically, we code our variable UHC as a dummy variable that takes the value one beginning with the year of the first legal implementation of health insurance, given this country achieved UHC in the meanwhile; and zero otherwise. As of 2010, we can find evidence of UHC in 43 countries around the world that meet our aforementioned criteria.

In the main regression analysis, we perform a pooled ordinary least squares (OLS) and repeated cross-sectional models in the form of equation 2:

$$Height\ inequality_{it} = \beta_0 + \beta_1 UHC_{it} + \gamma Z_{it} + \tau_t + \mu_j + \epsilon_{it} \tag{2}$$

where the outcome $Height\ Inequality_{it}$ is the Gini coefficient of height distribution per country i in region j and birth decade t. Our main independent variable is the dummy UHC_{it}. Z_{it} is a vector of control variables, capturing country characteristics. Here we control for a country's population and degree of urbanisation. In a larger country, public resources may be more unequally distributed across the population, and in more rural societies, the availability of health infrastructure may be more fragmented. We also include GDP per capita as a control variable to take account of a country's economic development. Finally, we control for the level of democratisation as measured by the polity2 index. A democratic state is expected to grant more (health) rights to the population and to provide higher social transfers (Ben-Bassat & Dahan, 2008; Rodrik, 1998). We further include time-fixed effects τ_t for the birth decades and μ_j world region-fixed effects. Including world region-fixed effects instead of country-fixed effects might capture cross-country differences and help to identify the relationship between social policies such as UHC and inequality more precisely (see Durlauf et al., 2005).

However, especially the inclusion of GDP per capita as a control variable should be regarded with caution, as it might be considered as bad control (Angrist & Pischke, 2009, 2014). We refer to the concern that there might exist a correlation between our main independent variable of interest, UHC, and the control for economic development. Countries with higher GDP per capita may be able to invest a higher proportion of their government expenditure in health services such as hospitals and medicines and be able to afford the provision of UHC to their population. In fact, the WHO (2010) notes that as a country's income rises, government spending on health tends to increase. However, there are still large differences in the share of government spending on health between low-income and high-income countries (WHO, 2010). Therefore, GDP per capita can provide important insights into variations in height inequality, as even poor people in high-income countries have adequate access to basic needs and health care compared to low-income countries. We include GDP per capita in our analysis to avoid potential omitted variable bias. However, to address the concern that GDP per capita, or any other control variable included in this model, is a bad control, we present our regression results as a sequence, including each control variable in turn.

To account for the possibility that the introduction of UHC and the implementation of reforms to achieve a wider coverage of the population is driven by high height inequality, we apply an IV approach.

D.4 Descriptive statistics and results

D.4.1 Descriptive statistics

Table D.1 reports the summary statistics of the sample used in our long-run analysis. In total, 134 countries are included in the dataset for the time period 1810 to 2000. It provides us with 1,190 observations for our main variables of interest, measured on a country-birth-decade unit.

In the world maps in Figure D.1 and Figure D.2, we compare height inequality within countries for the earliest and latest available observation in our sample. Figure D.1 displays the sample of country observations used to examine height inequality for the earliest observations available in each country, starting with the birth decade of 1810, with the 1900 birth decade providing the latest observations for this graph. Height inequality seems to be high in African and in Latin American countries for the 19th century and much lower for countries located in Europe and Asia. In Figure D.2 we show the most recent estimates for height Gini available per country, ranging from the birth decade 1960 to 2000. Again, European and Asian countries show the lowest levels of height inequality. The highest levels of inequality today can be seen mostly in the world regions of sub-Saharan Africa and the Middle East, with the highest inequality levels in countries such as the Democratic Republic of the Congo, Iraq, and Sudan.

In Figure D.3, we observe a rising gap in height inequality between different world regions despite having a similar starting point in the birth decades of around 1870/1880, before the Bismarckian social insurance system was introduced in Germany. The 19th century was also the period when, after an initial worsening before WW1, substantial achievements were obtained in improving health conditions across many societies and in reducing inequalities in health (Deaton, 2003, 2013). Especially after the birth decade of 1920, we observe this widening gap. For Europe, after the birth decade of 1920, inequality began to decrease. In contrast, in Africa and North America, we observe rising height disparities from the birth decade of 1930 to 1960. For our most recent estimates, the birth decade of 2000, inequality levels nowadays are the highest in South America and Africa.

One important factor for the difference in the development of height inequality might be the introduction of different welfare programs, especially UHC. Out of the 43 countries we find to have a 90% or higher coverage of health insurance for their population, as of 2010, we are able to include 37 countries in our sample where we do have additional data for height inequality. Over 67% of these countries are located in Western and Eastern Europe. In Figure D.4 we do observe that those countries having UHC show lower height inequality on average for the birth decade 2000 of 10.93 Gini points. For the whole sample period, starting in the 1810s, this difference decreases to 6.27 Gini points for each combination of country and decade of birth.

To explore this relationship in more detail, we divide inequality into five different sequences, ranging from low (<30), moderate (30–40), medium (40–50), high (50–60), and very high (>60) inequality. In Figure D.5 we display the percentage of the different inequality sequences. For countries and birth decades, where there was no UHC available for most of the population (=0), we observe that almost 25% of those observations (out of 965 country-birth decade combinations) show high to very high Gini coefficients. Moreover, it appears that there is a much higher prevalence of very high inequality in height compared to countries and birth decades where health insurance was already fully implemented.

D.4.2 Regression analysis

The regression results for our long-run analysis are presented in Table D.2. In all our models, we perform pooled OLS regressions with height inequality as the dependent variable, measured as Gini coefficient per country and birth decade. Our regression results show a consistently significant coefficient for UHC at a 1% significant level. As expected, the coefficients for UHC are negatively correlated with height inequality. In model (1), we run the regression analysis without time- or region-fixed effects but include them subsequently in the following models. In model (2), we include time-fixed effects, and in model (3), we include time- and world region-fixed effects, respectively. For comparison, in model (4), we include country-fixed effects. The statistically significant correlation between height inequality and UHC remains. In Table D.3 we study the relationship between health insurance and height inequality, including control variables on a step-by-step basis. We find that the coefficients or our main variable of interest remain statistically significant, even when the size of the coefficient and the significance level is slightly reduced with the inclusion of additional controls. Our results suggest a statistically significant correlation between GDP per capita and height inequality in all models. Higher levels of economic development are associated with lower levels of height inequality. For the degree of urbanisation, the results are less clear: the coefficient for urbanisation is statistically insignificant in model (3) but becomes significant once we include GDP per capita in the model, whereas population size and the level of democratisation do not seem to be significantly correlated with height inequality.

D.5 Instrumental variable estimates

Next, we need to consider endogeneity, as the results of the OLS regressions could be affected by reverse causality, measurement error or omitted variable issues. For example, apart from the direction of causation running from the introduction of health insurance to height inequality, one can also imagine that in the long run, regions with relatively 'good', i.e., low inequality values could reach a more easily consensus about introducing costly health insurance. Moreover, measurement errors or omitted variable issues are always an issue. Instrumental variable estimation allows us to circumvent these issues of endogeneity. We base our first stage of the two-stage-least-square (2SLS) estimate on the following equation 3:

$$UHC_{it} = \beta_1 + \beta_2 DistanceSovietUnion_i + \gamma' X_{it} + \epsilon_{it} \quad (3)$$

Where $DistanceSovietUnion_i$ is a cross-sectional spatial instrumental variable of the logged distance to the centre of the Soviet Union. The Soviet Union, as a socialist state, was perceived as a potential threat after its creation in 1922, and Western market economies introduced health insurance partly as a measure to keep workers from striving for socialism. X is a vector of other exogenous variables.

The results of the 2SLS regressions confirm that the IV fulfils the necessary requirements to be a good correlate of introducing social insurance in its proximity: first, it correlates negatively with the existence of UHC, as is documented by the 'first stage' section of Table D.4. The F-Test is clearly above 10 (see Stock and Yogo, 2005). We argue that the instrument influences the dependent variable only through the potentially endogenous variable, social insurance introduction. As a result, the significant impact of health insurance remains a consistent determinant of height inequality, assuming effective population coverage.

Communist states did emerge not only in Russia but also in other countries around the globe. However, we argue that the expansion of socialism in the Soviet Union was seen by neighbouring states as the greatest threat to the political system, compared to other, individual communist states. The advantage of using the distance to the Soviet Union as an instrument is its exogenous nature. This is because this socialist society was created in a country that was not even considered a potential candidate by Marxists around 1900. They believed that industrial countries like England or Belgium were prime candidates, not an agricultural Empire like the Russian one. It was by chance that the Russian Czar lost his strength after the disastrous WWI, and Lenin and his successors were convincing and brutal leaders personalities that initiated and kept socialism in their country (Malia, 2008). A violation of the exclusion restriction might result if health inequality is influenced by geographical characteristics. Geography is considered relevant in explaining health inequalities within countries, but only through the prevailing socioeconomic circumstances (see review in Smith and Easterlow, 2005). Our spatial instrument should therefore be valid if height inequality is not affected by the distance to Russia decades before the emergence of socialism and the expansion of the Soviet Union. In Table D.5 we display the correlation between height inequality and our IV, the distance to the Soviet Union, before 1880 and find no statistically significant correlation. We have chosen 1880 because all participants of the October Revolution 1917 were born after 1880. We, therefore, conclude that the geographic location itself has no direct effect on our dependent variable height inequality, but through the communist threat of the Soviet Union.

D.6 Robustness check

To test the robustness of our result, we exclude communist states from the regression analysis in model (2) of Table D.6. For comparison, our results from the main analysis are here again displayed in model (1). As argued in the previous section, the emergence of social parties was seen as a threat to political stability, especially after the Russian October Revolution of 1917 and the subsequent beginnings of the creation of the Soviet Union. It is therefore seen as a motivating force for social reforms in neighbouring countries, such as the introduction of health insurance (Bauernschuster et al., 2020). As a robustness check, we therefore exclude countries for the periods in which they are considered to be socialist states. Our results in model (2) are robust to the exclusion of communist countries. Interestingly, the coefficient on UHC is slightly higher in this case. The literature is ambiguous about this effect. Although the 'welfare regime theory' suggests that health inequalities are lower in socialist regimes, empirical evidence does not support this theory (see review in Brennenstuhl et al., 2012).

We also control for other definitions of UHC. We include the measurement of *v2pehealth* from the Varieties of Democracy (V-Dem). It controls for the access to public services by the population on a scale from zero to four, where zero indicates unequal access to public services and the highest number indicates equal access for everyone (Coppedge et al., 2022). We further include the UHC service coverage index (2.8.1) from the WHO in Table D.6, model (3), even though it substantially reduces the number of observations as we are just able to include the birth decade of 2000. In models (4) and (5), we include measures that only consider the legal implementation of health insurance, *UHC (de jure)*, or have included the right to health in their constitution, but do not control for the actual achievement of UHC. We find that 55 countries have implemented UHC, at least as part of their legal framework. Similar to our findings from the main regression analysis, we observe a negative and significant coefficient for UHC when we control for the actual achievement of UHC in models (3) and (4). However, the legal introduction of health insurance, without full implementation, does not seem to be significantly associated with height inequality, supporting our previous findings and the argument that achieving UHC is a prerequisite for significantly reducing health inequalities in a country.

D.7 Conclusion

This paper has examined the effects of the extension of health insurance on height inequality using a measure of human heights, which allows for examining the effects of institutional reform, namely health insurance, back in time. Our results suggest that reforms that reduce financial barriers to accessing health care can have an impact on the health of individuals. Hence, we would expect a statistically significant mitigating effect on inequalities in health. We use an anthropological measure of health and well-being, namely human heights. Although evidence on the effects of insurance on health inequality is contentious, we examine evidence of a period where there were large expansions of health insurance in several countries on a measure of height inequality, which is not sensitive to self-reporting bias. We study whether inequalities in heights decline with the expansion of healthcare systems, controlling for a number of relevant control variables. We drew on a unique dataset of countries where we can measure individuals' heights retrospectively for several decades. Our estimates of the cross-country comparison suggest evidence that within-country differences in height inequality are indeed explained by health insurance expansions.

D.8 References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Angrist, J. D., & Pischke, J.-S. (2014). *Mastering'metrics: The path from cause to effect.* Princeton university press.

Baicker, K., Mullainathan, S., & Schwartzstein, J. (2015). Behavioral hazard in health insurance. *The Quarterly Journal of Economics*, *130*(4), 1623–1667.

Baten, J. (1999). *Ernährung und wirtschaftliche entwicklung in bayern:(1730-1880).* Steiner Stuttgart.

Baten, J. (2000a). Economic development and the distribution of nutritional resources in bavaria, 1797–1839: An anthropometric study. *Journal of Income Distribution*, *9*(1), 89–106.

Baten, J. (2000b). Heights and real wages in the 18th and 19th centuries: An international overview. *Jahrbuch für Wirtschaftsgeschichte/Economic History Yearbook*, *41*(1), 61–76.

Baten, J., & Blum, M. (2011). Anthropometric within-country inequality and the estimation of skill premia with anthropometric indicators. *Review of Economics*, *62*(2), 107–138.

Baten, J., & Blum, M. (2014). Why are you tall while others are short? agricultural production and other proximate determinants of global heights. *European Review of Economic History*, *18*(2), 144–165.

Baten, J., & Mumme, C. (2013). Does inequality lead to civil wars? a global long-term study using anthropometric indicators (1816–1999). *European Journal of Political Economy*, *32*, 56–79.

Bauernschuster, S., Driva, A., & Hornung, E. (2020). Bismarck's health insurance and the mortality decline. *Journal of the European Economic Association*, *18*(5), 2561–2607.

Beard, A. S., & Blaser, M. J. (2002). The ecology of height: The effect of microbial transmission on human height. *Perspectives in Biology and Medicine*, *45*(4), 475–498.

Ben-Bassat, A., & Dahan, M. (2008). Social rights in the constitution and in practice. *Journal of Comparative Economics*, *36*(1), 103–119.

Blum, M. (2014). Estimating male and female height inequality. *Economics & Human Biology*, *14*, 103–108.

Bogin, B. (1988). Patterns of human growth. cambridge universitypress.[arjll, rbe](1990) the evolution of human childhood. *BioScience*, *40*, 1625.

Brennenstuhl, S., Quesnel-Vallée, A., & McDonough, P. (2012). Welfare regimes, population health and health inequalities: A research synthesis. *J Epidemiol Community Health*, *66*(5), 397–409.

Busse, R., Blümel, M., Knieps, F., & Bärnighausen, T. (2017). Statutory health insurance in germany: A health system shaped by 135 years of solidarity, self-governance, and competition. *The Lancet*, *390*(10097), 882–897.

Candela-Martínez, B., Cámara, A. D., López-Falcón, D., & Martínez-Carrión, J. M. (2022). Growing taller unequally? adult height and socioeconomic status in spain (cohorts 1940–1994). *SSM-Population Health*, *18*, 101126.

Case, A., Lubotsky, D., & Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review*, *92*(5), 1308–1334.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., et al. (2022). Vdem dataset v12.

Costa-Font, J., Cowell, F. A., & Saenz de Miera, B. (2021). Measuring pure health inequality and mobility during a health insurance expansion: Evidence from mexico. *Health Economics*, *30*(8), 1833–1848.

Currie, J., & Gruber, J. (1996). Health insurance eligibility, utilization of medical care, and child health. *The Quarterly Journal of Economics*, *111*(2), 431–466.

Cutler, D. M., & Johnson, R. (2004). The birth and growth of the social insurance state: Explaining old age and medical insurance across countries. *Public choice*, *120*(1-2), 87–121.

Deaton, A. (2003). Health, inequality, and economic development. *Journal of economic literature*, *41*(1), 113–158.

Deaton, A. (2013). What does the empirical evidence tell us about the injustice of health inequalities. *Inequalities in health: concepts measures, and ethics*, *263*, 281.

Durlauf, S. N., Johnson, P. A., & Temple, J. R. (2005). Growth econometrics. *Handbook of economic growth*, *1*, 555–677.

Einav, L., & Finkelstein, A. (2018). Moral hazard in health insurance: What we know and how we know it. *Journal of the European Economic Association*, *16*(4), 957–982.

Ellis, R. P., & Manning, W. G. (2007). Optimal health insurance for prevention and treatment. *Journal of Health Economics*, *26*(6), 1128–1150.

Eveleth, P. B., & Tanner, J. M. (1976). *Worldwide variation in human growth* (Vol. 8). CUP Archive.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., & Group, O. H. S. (2012). The oregon health insurance experiment: Evidence from the first year. *The Quarterly journal of economics*, *127*(3), 1057–1106.

Fogel, R. W. (1994). Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy1. *The American Economic Review*, *84*(3), 369–395.

Fogel, R. W., Engerman, S. L., & Trussell, J. (1982). Exploring the uses of data on height: The analysis of long-term trends in nutrition, labor welfare, and labor productivity. *Social Science History*, *6*(4), 401–421.

Freeman, J. D., Kadiyala, S., Bell, J. F., & Martin, D. P. (2008). The causal effect of health insurance on utilization and outcomes in adults: A systematic review of us studies. *Medical care*, 1023–1032.

Goodman-Bacon, A. (2021). The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes. *American Economic Review*, *111*(8), 2550–2593.

Gorsky, M. (2011). The Political Economy of Health Care in the Nineteenth and Twentieth Centuries. In *The Oxford Handbook of the History of Medicine*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199546497.013.0024

Haushofer, J., Chemin, M., Jang, C., & Abraham, J. (2020). Economic and psychological effects of health insurance and cash transfers: Evidence from a randomized experiment in kenya. *Journal of Development Economics*, *144*, 102416.

International Labour Organization. (2022). Statistics on social protection. https://ilostat.ilo.org/topics/social-protection/

Kangas, O. E. (2012). Testing old theories in new surroundings: The timing of first social security laws in africa. *International Social Security Review*, *65*(1), 73–97.

Komlos, J. (1985). Stature and nutrition in the habsburg monarchy: The standard of living and economic development in the eighteenth century. *The American Historical Review*, *90*(5), 1149–1161.

Kuh, D., & Wadsworth, M. (1989). Parental height: Childhood environment and subsequent adult height in a national birth cohort. *International Journal of Epidemiology*, *18*(3), 663–668.

Li, L., Manor, O., & Power, C. (2004). Are inequalities in height narrowing? comparing effects of social class on height in two generations. *Archives of disease in childhood*, *89*(11), 1018–1023.

Makinde, O. A., Sule, A., Ayankogbe, O., & Boone, D. (2018). Distribution of health facilities in nigeria: Implications and options for universal health coverage. *The International journal of health planning and management*, *33*(4), e1179–e1192.

Malia, M. (2008). *Soviet tragedy: A history of socialism in russia*. Simon; Schuster.

Moradi, A., & Baten, J. (2005). Inequality in sub-saharan africa: New data and new insights from anthropometric estimates. *World development*, *33*(8), 1233–1265.

Organisation for Economic Co-operation and Development. (2022). Social protection [OECD Health Statistics (database) accessed on 01 December 2022, https://doi.org/10.1787/data-00544-en].

Persico, N., Postlewaite, A., & Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of political Economy*, *112*(5), 1019–1053.

Quinn, D. P., & Woolley, J. T. (2001). Democracy and national economic performance: The preference for stability. *American journal of political science*, 634–657.

Radatz, L., & Baten, J. (n.d.). Measuring multidimensional inequality and conflict in africa and in a global comparison.

Rodrik, D. (1998). Why do more open economies have bigger governments? *Journal of political economy*, *106*(5), 997–1032.

Schmitt, L., & Harrison, G. (1988). Patterns in the within-population variability of stature and weight. *Annals of Human Biology*, *15*(5), 353–364.

Smith, S. J., & Easterlow, D. (2005). The strange geography of health inequalities. *Transactions of the Institute of British Geographers*, *30*(2), 173–190.

Steckel, R. H. (1995). Stature and the standard of living. *Journal of economic literature*, *33*(4), 1903–1940.

Stinson, S. (1985). Sex differences in environmental sensitivity during growth and development. *American Journal of Physical Anthropology*, *28*(S6), 123–147.

Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, *6*, 109–120.

Strauss, J., & Thomas, D. (1998). Health, nutrition, and economic development. *Journal of economic literature*, *36*(2), 766–817.

Stuckler, D., Basu, S., & McKee, M. (2010). Budget crises, health, and social welfare programmes. *Bmj*, *340*.

Thomasson, M. A. (2002). From sickness to health: The twentieth-century development of us health insurance. *Explorations in Economic History*, *39*(3), 233–253.

van Zanden, J. L., Baten, J., Foldvari, P., & van Leeuwen, B. (2014). The changing shape of global inequality 1820–2000; exploring a new dataset. *Review of income and wealth*, *60*(2), 279–297.

Wang, H., & Yu, Y. (2016). Increasing health inequality in china: An empirical study with ordinal data. *The Journal of Economic Inequality*, *14*(1), 41–61.

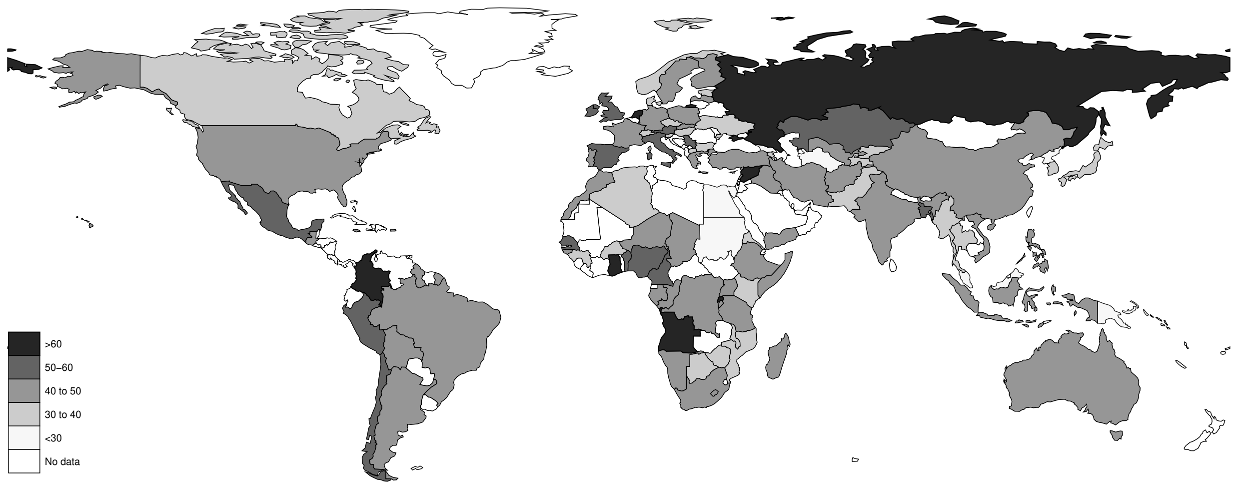
World Health Organization. (2010). Health systems financing: The path to universal coverage. *The world health report. Geneva*, 1–128.

World Health Organization. (2023). The global health observatory. uhc service coverage index (sdg 3.8.1). https://www.who.int/data/gho/data/indicators/indicator-details/GHO/uhc-index-of-service-coverage

World Health Organization & World Bank. (2021). Tracking universal health coverage: 2021 global monitoring report.

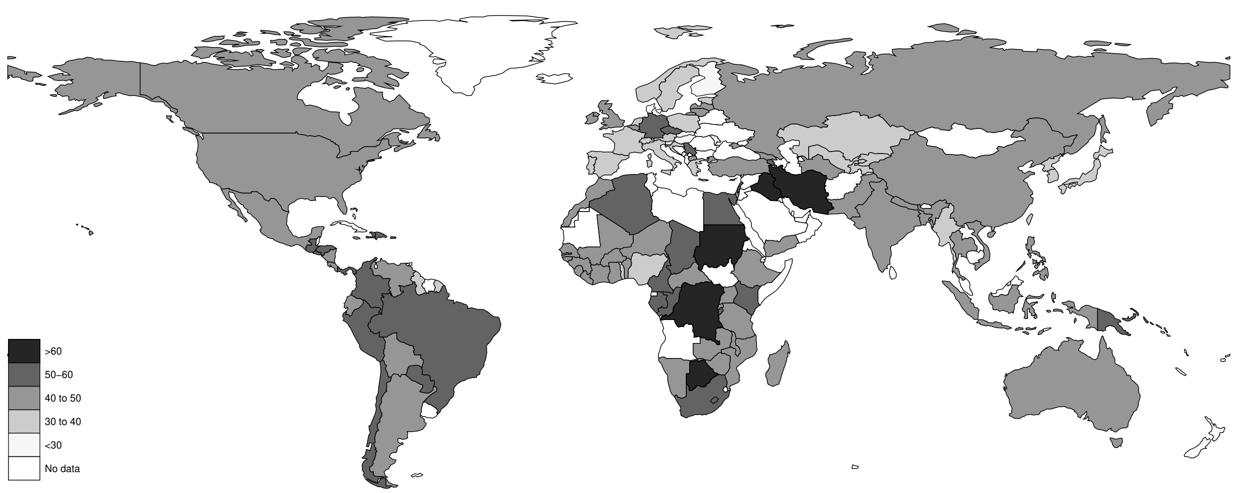
D.9 Tables and figures

Figure D.1: Height inequality worldwide: 1810–1900



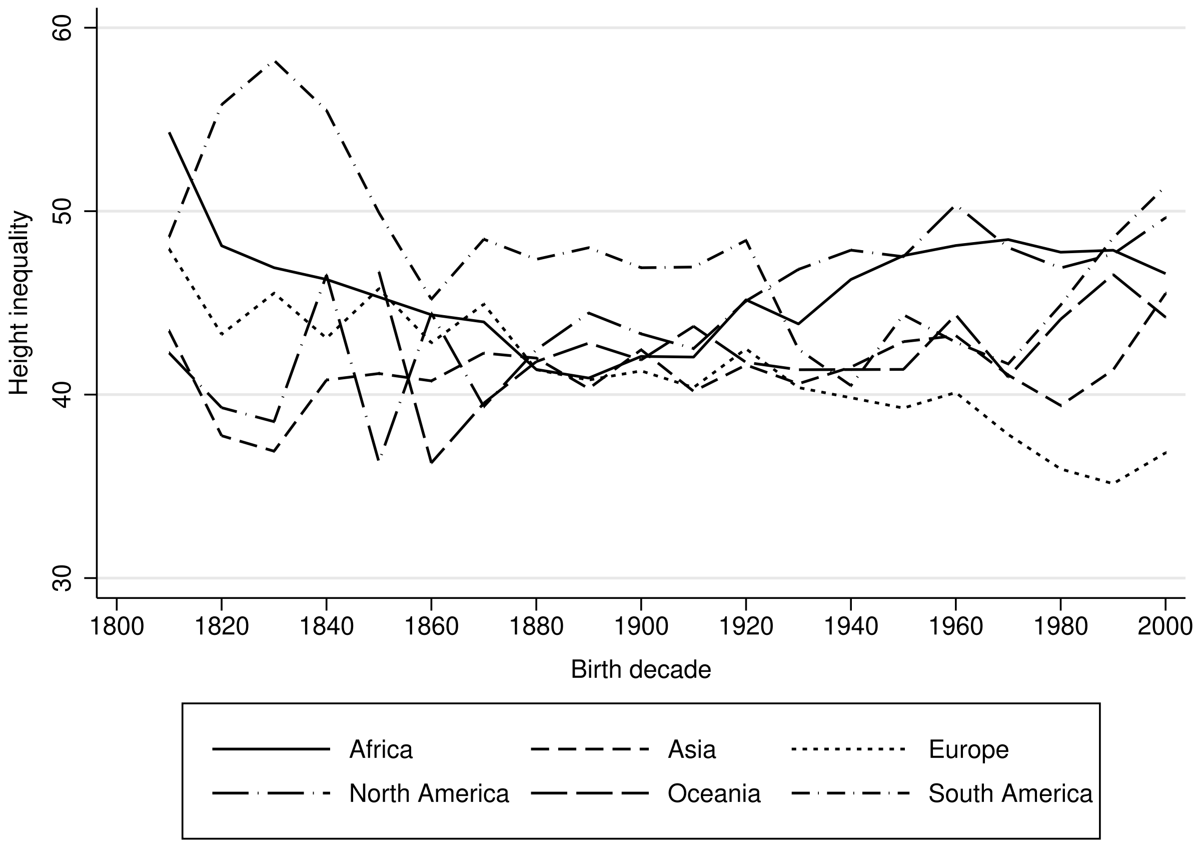
Notes. Darker shades indicate higher inequality levels, and lighter shades lower levels of height inequality. We display the earliest data available for each country beginning with birth decade 1810.

Figure D.2: Height inequality worldwide: 1960–2000



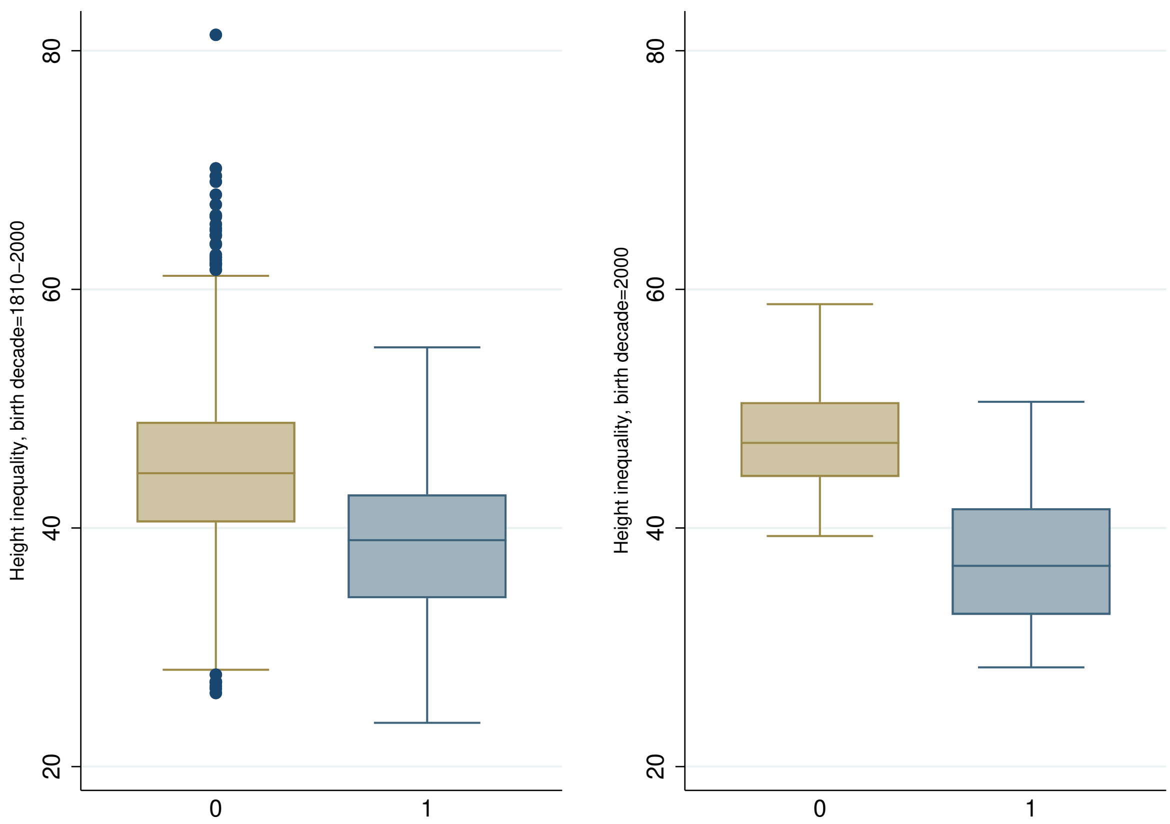
Notes. Darker shades indicate higher inequality levels, and lighter shades lower levels of height inequality. We display the most recent data available for each country beginning with the birth decade 1960 until 2000.

Figure D.3: Development of height inequality by world region



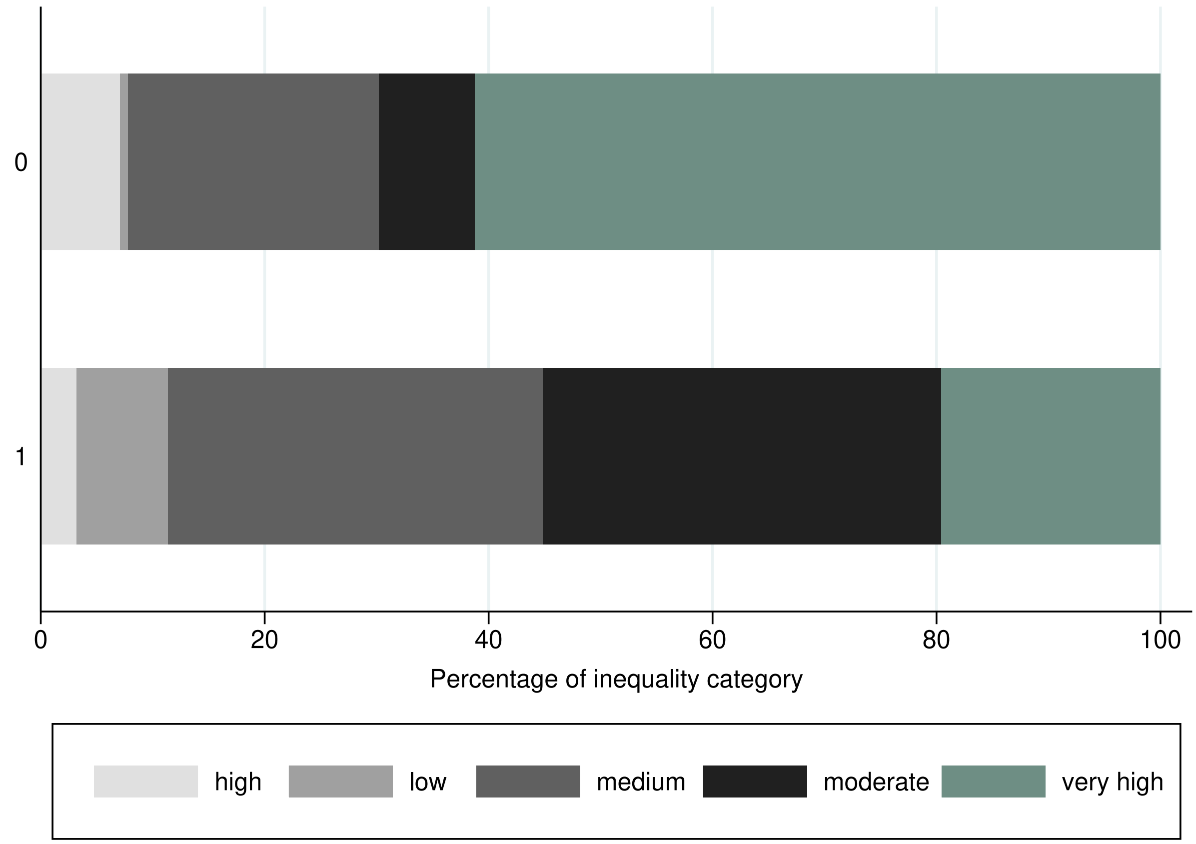
Notes. Height inequality over time and by world region. Height inequality is measured as Gini coefficient and calculated per birth decade and country.

Figure D.4: Differences in height inequality by Universal Health Coverage



Notes. We show the difference in height inequality by countries with (=1) and without (=0) UHC, for the whole time period 1810–2000 on the left, and for the birth decade 2000 on the right. Height inequality is measured as Gini coefficient.

Figure D.5: Percentage of different levels of height inequality by the (non-)presence of Universal Health Coverage



Notes. The percentage of country-decade combination are displayed depending on their inequality level. Height inequality is hereby divided into five sequences: low inequality (< 30 Gini points), moderate (30–39 Gini points), medium (40–49 Gini points), high (50–59 Gini points), and very high inequality (> 60 Gini points). We further distinguish between the presence and fulfilment of UHC (=1) compared to no UHC (=0).

Table D.1: Summary statistics

	Obs	Mean	Std. Dev.	Min	Max
Height inequality	1191	43.74	7.64	23.66735	81.33366
UHC	1191	0.19	0.39	0	1
Population (log)	1183	15.80	1.57	11.89412	20.95647
Urbanisation	1184	0.27	0.21	.0005245	.96377
GDP per capita (log)	864	8.20	1.04	6.262064	11.10131
Democracy	796	0.68	6.83	-10	10

Notes. All variables are measured on a country-decade unit. Height inequality is measured as Gini coefficient. UHC is coded as one if a country achieved UHC, indicating the years after the first implementation, zero otherwise.

Table D.2: Correlation between height inequality and Universal Health Coverage

	(1) OLS	(2) OLS	(3) OLS	(4) OLS
UHC	-6.27^{***}	-6.88^{***}	-4.73^{***}	-4.46^{***}
	(0.843)	(0.900)	(1.158)	(1.419)
Observations	1191	1191	1191	1191
R-squared	0.103	0.140	0.188	0.412
Time Fixed Effects	N	Y	Y	Y
Region Fixed Effects	N	N	Y	N
Country Fixed Effects	N	N	N	Y

Notes. Country-clustered robust standard errors in parentheses, ***, **, *, significant on the 1, 5, and 10%-level, respectively. The dependent variable is height inequality in every model. UHC is coded as one if a country achieved UHC, indicating the years after the first implementation, zero otherwise.

Table D.3: OLS regression including controls: correlates of height inequality

	(1) Height Gini	(2) Height Gini	(3) Height Gini	(4) Height Gini	(5) Height Gini	(6) Height Gini
UHC	−4.73***	−4.66***	−4.93***	−3.86**	−3.56**	−5.35***
	(1.158)	(1.159)	(1.296)	(1.502)	(1.556)	(1.394)
Population (log)		0.23	0.20	−0.07	−0.22	−0.59
		(0.287)	(0.281)	(0.352)	(0.432)	(0.427)
Urbanisation			4.50	9.86***	10.02**	9.84**
			(2.951)	(3.697)	(3.972)	(4.090)
GDP per capita (log)				−1.90**	−1.92**	−2.15**
				(0.746)	(0.865)	(0.955)
Democracy					−0.06	−0.03
					(0.089)	(0.092)
Democracy^2					−0.30	−0.74
					(1.298)	(1.284)
Constant	51.63***	48.23***	48.68***	67.74***	69.73***	74.90***
	(1.960)	(4.693)	(4.596)	(6.886)	(8.593)	(9.214)
Observations	1191	1183	1176	857	722	722
R-squared	0.188	0.190	0.188	0.227	0.226	0.192
Time Fixed Effects	Y	Y	Y	Y	Y	Y
Region Fixed Effects	Y	Y	Y	Y	Y	N

Notes. Country-clustered robust standard errors in parentheses, ***, **, *, significant on the 1, 5, and 10%-level, respectively. The dependent variable is height inequality in every model. UHC is coded as one if a country achieved UHC, indicating the years after the first implementation, zero otherwise. For interpretation, GDP per capita is divided by 1,000 before running the regression. Marginal effects reported.

Table D.4: Instrumental variable approach: determinants of height inequality

	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
First stage					
DistSovietUnion	−0.22***	−0.22***	−0.20***	−0.24***	−0.27***
	(0.021)	(0.023)	(0.021)	(0.027)	(0.029)
Second stage					
UHC	−24.11***	−26.58***	−28.63***	−25.49***	−22.37***
	(2.665)	(3.157)	(3.628)	(3.646)	(3.185)
Population (log)		0.47**	0.16	−0.19	−0.49**
		(0.220)	(0.192)	(0.204)	(0.214)
Urbanisation			24.82***	13.08***	13.72***
			(4.509)	(3.180)	(3.305)
GDP per capita (log)				2.28**	0.95
				(0.948)	(0.855)
Democracy					0.07
					(0.076)
Democracy^2					1.55
					(1.189)
Observations	1191	1183	1176	857	722
Adj. R-squared	0.178	0.176	0.369	0.472	0.509
Time Fixed Effects	Y	Y	Y	Y	Y
Region Fixed Effects	N	N	N	N	N
F-statistic	109.90	91.43	87.91	76.97	81.10
Kleinbergen-Paap rk LM statistic	Exactly identified				
Hansen J statistic	Exactly identified				

Notes. Robust standard errors in parentheses, ***, **, *, significant on the 1, 5, and 10%-level, respectively. The dependent variable in the first stage is UHC and height inequality in the second stage. UHC is coded as one if a country achieved UHC, indicating the years after the first implementation, zero otherwise. We take the natural logarithm for the variables *DistSovietUnion*, Population and GDP per capita. For interpretation, we divided *DistSovietUnion* by 1,000 before running the regression.

Table D.5: Correlation between height inequality and the distance to the Soviet Union

	(1) Height Gini	(2) Height Gini
	Omitted Birth decades >=1880	*Omitted Birth decades <1880*
DistSovietUnion	5.83	6.32***
	(4.322)	(1.444)
Constant	53.01***	47.44***
	(2.918)	(1.249)
Observations	185	1006
R-squared	0.134	0.218
Time Fixed Effects	Y	Y
Region Fixed Effects	Y	Y

Notes. Robust standard errors in parentheses, ***, **, *, significant on the 1, 5, and 10%-level, respectively. The dependent variable is height inequality. We take the natural logarithm of *DistSovietUnion* and divided it by 1,000 before running the regression.

Table D.6: Robustness check for height inequality and healthcare measurements

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
UHC	−4.73***					
	(1.158)					
UHC *(Communist states omitted)*		−4.83***				
		(1.185)				
Health coverage, V-DEM			−1.02***			
			(0.321)			
UHC Index, WHO				−0.16***		
				(0.042)		
UHC (de jure)					−1.49	
					(1.106)	
Right to health						−1.28
						(1.108)
Constant	51.63***	55.02***	41.67***	52.30***	52.75***	52.99***
	(1.960)	(2.226)	(1.004)	(2.177)	(1.921)	(1.934)
Observations	1191	1059	804	55	1191	1191
R-squared	0.188	0.208	0.223	0.158	0.159	0.157
Time Fixed Effects	Y	Y	Y	N	Y	Y
Region Fixed Effects	Y	Y	Y	N	Y	Y

Notes. Country-clustered robust standard errors in parentheses, ***, **, *, significant on the 1, 5, and 10%-level, respectively. The dependent variable is height inequality in every model. Our measurement of UHC is used in column (1) and (2), and communist countries are excluded in the latter. The variable health coverage is derived from the V-DEM project and refers to the v2pehealth indicator, ranging from zero to four, where four indicates the equal access of all citizens to the health system. The UHC index is provided by the WHO. UHC (de jure) indicates the legal implementation of health insurance, as well as the right to health.

D.10 Appendix

D.10.1 Height inequality

Height inequality is measured as the Gini coefficient using the coefficient of variation (CV) of height. Data is derived from Clio Infra and updated by the extension done by Radatz and Baten (n.d.). The compiled dataset is based on several sources such as household surveys, for example, the Demographic and Health Surveys (DHS) and individual authors. For more details see the collection of data from Baten and Blum (2015) and Radatz and Baten (n.d.). Data on height inequality is based on the birth cohort approach, therefore providing data by birth decade, starting from 1810 to the birth decade of 2000. In total, data is offered for 193 countries worldwide.

D.10.2 Universal Health Coverage: measurement and sources

Universal Health Coverage (UHC) is our main independent variable and is coded as a dummy variable. It takes on the value of one after the decade of the first legal implementation of health insurance, given that coverage for 90% of the population was achieved by 2010.

To construct our variable *UHC*, first, we collected data on the timing of the implementation from different sources. We mainly rely on Cutler and Johnson (2004) and the report series 'Social Security Programs Throughout the World' of the Social Security Administration (SSA), which is provided for the Americas, Asia and the Pacific, and Europe (see Social Security Administration [SSA], 2018, 2019, 2020). Moreover, we obtain data from Goudima and Rybalko (1996), Rosen et al. (2015), and WHO et al. (2013) for some individual countries. A detailed overview of the sources used for each country can be found in Table D.7. By rounding off the years, we refer to the respective decade for the year of the introduction of health insurance. For example, we refer to an implementation of health insurance for the decade 1910, if the legal implementation took place in a year between 1910 and 1915, as for example in Ireland in 1911 (SSA, 2018). If the introduction took place in the years 1916 to 1919, we add this observation to the decade 1920.

Second, we checked if a country achieved UHC. For 43 countries, we find a full achievement of UHC. For each country included we display the birth decade and our sources in Table D.7. However, our sample is limited to 37 countries, as we just include country-birth decades for which all our main variables are available. The coverage of the population is measured based on the indicators provided by the OECD (2022), for the percentage of the population covered by public or private health insurance, and the social protection indicator from ILOSTAT (2022) for non-OECD countries. The reference year is 2010.

D.10.3 Controls

We include the following control variables in our regression analysis:

Population (log). We control for the size of a country's population. Population size is measured by the natural logarithm of a country's population at the start of each decade. Source: Fink-Jensen (2015), available via Clio Infra.

Urbanisation. The variable urbanisation shows the ratio of the urban population to the population living in rural areas within a country and for a specific decade. Source: Fink-Jensen (2015), available via Clio Infra.

GDP per capita (log). Based on a country-birth decade unit we consider GDP per capita as a control variable, taking the natural logarithm. Source: Bolt and Van Zanden (2020).

Democracy. Our democracy variable is derived from the Polity5 project. It measures the degree of democratisation within a country. It ranges from -10 points for a full autocracy to +10 points for a fully consolidated democracy. Source: Marshall and Gurr (2020).

D.10.4 References

Baten, J., & Blum, M. (2015). Height gini. http://hdl.handle.net/10622/HKZPR9

Bolt, J., & Van Zanden, J. L. (2020). Maddison style estimates of the evolution of the world economy. a new 2020 update. *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands*.

Cutler, D. M., & Johnson, R. (2004). The birth and growth of the social insurance state: Explaining old age and medical insurance across countries. *Public choice*, *120*(1-2), 87–121.

Fink-Jensen, J. (2015). Urbanization ratio. http://hdl.handle.net/10622/LZ0Y36

Goudima, T., & Rybalko, L. (1996). Social insurance in russia: History and contemporaneity. *Nordisk försäkringstidskrift*, *1996*, 342–50.

International Labour Organization. (2022). Statistics on social protection. https://ilostat.ilo.org/topics/social-protection/

Marshall, M. G., & Gurr, T. R. (2020). Polity5: Political regime characteristics and transitions, 1800-2018. *Center for Systemic Peace*.

Organisation for Economic Co-operation and Development. (2022). Social protection [OECD Health Statistics (database) accessed on 01 December 2022, https://doi.org/10.1787/data-00544-en].

Radatz, L., & Baten, J. (n.d.). Measuring multidimensional inequality and conflict in africa and in a global comparison.

Rosen, B., Waitzberg, R., & Merkur, S. (2015). Health systems in transition. *Israel: Health system review*.

Social Security Administration. (2018). Social security programs throughout the world: Europe, 2018. https://www.ssa.gov/policy/docs/progdesc/ssptw/2018-2019/europe/ssptw18europe.pdf

Social Security Administration. (2019). Social security programs throughout the world: Asia and the pacific, 2018. https://www.ssa.gov/policy/docs/progdesc/ssptw/2018-2019/asia/ssptw18asia.pdf

Social Security Administration. (2020). Social security programs throughout the world: The americas, 2019. https://www.ssa.gov/policy/docs/progdesc/ssptw/2018-2019/americas/ssptw19americas.pdf

WHO, L., Murauskiene, Janoniene, R., Veniute, M., van Ginneken, E., Karanikolos, M., et al. (2013). Lithuania: Health system review.

D.10.5 Tables and figures

Table D.7: Sources for health insurance legislation

Country	ccode	Birth Decade	Source
Australia	au	1970	Cutler and Johnson (2004)
Austria	at	1890	Cutler and Johnson (2004)
Belgium	be	1940	Cutler and Johnson (2004)
Canada	ca	1970	Cutler and Johnson (2004)
Chile	cl	1980	SSA (2020)
Colombia	co	1960	SSA (2020)
Costa Rica	cr	1940	SSA (2020)
Czech Republic	cz	1890	SSA (2018)
Denmark	dk	1930	Cutler and Johnson (2004)
Estonia	ee	1920	SSA (2018)
Finland	fi	1960	Cutler and Johnson (2004)
France	fr	1930	Cutler and Johnson (2004)
Germany	de	1880	Cutler and Johnson (2004)
Greece	gr	1920	SSA (2018)
Guyana	gy	1970	SSA (2020)
Hungary	hu	1890	SSA (2018)
Iceland*	is	1940	SSA (2018)
Ireland	ie	1910	SSA (2018)
Israel	il	1950	Rosen et al. (2015)
Italy	it	1940	Cutler and Johnson (2004)
Japan	jp	1930	Cutler and Johnson (2004)
Kazakhstan	kz	1910	Goudima and Rybalko (1996)
Latvia	lv	1920	SSA (2018)
Lithuania	lt	1990	WHO et al. (2013)
Luxembourg*	lu	1900	SSA (2018)
Netherlands	nl	1940	Cutler and Johnson (2004)
New Zealand*	nz	1940	Cutler and Johnson (2004)
Norway	no	1910	Cutler and Johnson (2004)
Poland	pl	1920	SSA (2018)
Portugal	pt	1930	Cutler and Johnson (2004)
Romania*	ro	1930	SSA (2018)
Russia	ru	1910	Goudima and Rybalko (1996)
Singapore*	sg	1950	SSA (2019)
Slovakia	sk	1990	SSA (2018)

Table D.7: (continued)

Country	ccode	Birth Decade	Source
Slovenia	si	1920	SSA (2018)
South Korea	kr	1980	SSA (2019)
Spain	es	1940	Cutler and Johnson (2004)
Sweden	se	1930	SSA (2018)
Switzerland	ch	1990	Cutler and Johnson (2004)
Taiwan	tw	1950	SSA (2018)
Turkey	tr	1950	SSA (2018)
United Kingdom	uk	1910	Cutler and Johnson (2004)
Uruguay*	uy	1970	SSA (2020)

Notes. Countries marked with a star $*$ are not included in the regression analysis due to missing data for height Gini.

E Summary and Outlook

This dissertation provides important insights into the long-run development of inequality. Specifically, it provides new evidence through the construction of a large and consistent dataset on inequality by the usage of anthropological measures, spanning from 1810 to 2000. The main objective was to identify the determinants of inequality and its implications by empirically examining the relationship between inequality and civil conflict, pandemics, and health insurance expansion in more detail.

The first study, Chapter B, therefore provides an alternative measure of inequality by constructing a joint inequality index that accounts for the multidimensionality of inequality. It then assesses the impact of high inequality on civil war conflict in sub-Saharan Africa and globally for over 200 years, from 1810 to 2010. To construct this joint index, several dimensions of inequality are taken into account, namely income, health and land inequality. The measurement of health inequality is based on the use of an anthropological measure: the distribution of heights, which reflects the general level of inequality within a country. In this study, the dataset on height inequality based on Baten and Blum (2014) has been substantially expanded and updated, both in terms of country and time coverage. It then combines all three components—income, health and land inequality—into a joint inequality index and empirically examines its impact on civil wars. Particularly in the area of the inequality-conflict nexus, data scarcity and consistency have been problematic and are seen as one of the main reasons for inconsistent and mixed findings (Cramer et al., 2005). This study contributes to the literature by providing critical long-term evidence on inequality and conflict for 193 countries worldwide, in particular new evidence for developing countries. The empirical evidence suggests a positive relationship between inequality and civil war: an increase in inequality led to a higher risk of a civil war outbreak in a country. Findings support the grievance argument discussed in the literature (see Collier and Hoeffler, 2004). However, the negative impact of inequality on civil war is not limited to the poorest countries in sub-Saharan Africa. The United States (US) and the United Kingdom (UK) have experienced significant increases in inequality over the past four decades, and the cross-country evidence in this study suggests that these countries are at increased risk of civil conflict.

The second chapter contributes to the literature and recent developments by examining the relationship between pandemics and inequality. Motivated by the COVID-19 outbreak in early 2020 and its multiple observable impacts on societies in terms of health, social and economic implications, this study looks at a major past pandemic: the 1918 influenza pandemic. It studies whether and how this health shock can be linked to the observed increase in income inequality and in health inequality between 1910 and 1920. Other studies have focused on the link between pandemics and economic development (e.g., Barro et al., 2020; Eichenbaum et al., 2021; Jordà et al., 2022), rather than inequality, and have mainly focused on one specific (industrialised) country (e.g., Almond, 2006; Galletta and Giommoni, 2022; Guimbeau et al., 2022). By examining the relationship between the severity of the pandemic in terms of mortality and inequality for countries around the world, including evidence for low-income countries, this study provides a much broader picture. Possible pathways that could explain an increase in inequality include the asymmetric health risk between poorer and richer people, labour market dynamics and the shock to aggregate demand. Although the regression results show a positive correlation between pandemic mortality and inequality, the relationship appears to be statistically insignificant. Therefore, there is no evidence that the 1918 influenza pandemic is associated with long-term differences in social and economic outcomes. In the case of the recent COVID-19 pandemic, however, appropriate policy interventions should nevertheless be considered to mitigate possible short- and long-term effects.

Finally, the third chapter of this dissertation investigates a social determinant of inequality, namely Universal Health Coverage (UHC), and its impact on reducing disparities in height. Covering the period from 1810 to 2000, this long-term study examines globally whether the implementation of health insurance legislation has reduced height inequalities in subsequent decades, controlling for the achievement of full coverage of the population (as of 2010). As the introduction of health insurance was legislated in many countries at the end of the 19^{th} and beginning of the 20^{th} centuries, data to measure its impact are scarce (Bauernschuster et al., 2020). This chapter provides evidence for a large number of countries worldwide, dating back to the first introduction of a health insurance scheme in Germany in 1883. This can be achieved by again using the anthropological measures of height to account for the level of inequality within a country, providing a rich dataset on health insurance implementation and height inequality. The estimates show that within-country disparities in height decrease with health insurance expansion, given that it effectively covers the whole population. Improving access to the healthcare system for all, and thus improving the health of individuals, is considered to be particularly beneficial to poorer people in society, as they would otherwise have experienced difficulties in accessing healthcare due to, for example, financial barriers (Einav & Finkelstein, 2018). The study, therefore, concludes by calling for the adaptation of appropriate public policies, as the expansion of UHC can have a significant impact on reducing health inequalities.

As outlined in this dissertation, high inequality can have several negative implications, and there are several reasons why we should reduce inequality as it disadvantages many people in our society. Furthermore, inequality is not given and can be addressed through appropriate policy interventions (Alvaredo et al., 2018). The main call of this dissertation is therefore obviously to reduce inequality. How can this reduction in inequality be achieved?

First, different dimensions of inequality, such as unequal access to education or healthcare, could be tackled by ensuring that everyone has the same opportunities. Educational policies should aim to increase school attainment and the quality of the school system, which could lead to higher levels of education also for people from low-income households, thereby reducing inequalities in income and wages. Health policies should target equal access to the health system. Our results in Chapter D suggest that the introduction and achievement of UHC in a country leads to a significant reduction in height inequalities. Promoting health coverage for the whole population, combined with investment in the health system, could therefore reduce health disparities, as has been successfully demonstrated in Brazil and South Africa (UN-DESA, 2018). Another effective tool, but much less favoured by politicians, is to raise taxes on high-income earners and wealthier people in society.

Combined with government redistribution, this could be used to explicitly target the poorest in society and disadvantaged social groups. In times of crisis, such as the recent COVID-19 pandemic, these instruments could include (unconditional) cash transfers to low-income households to cushion short-term income shocks. They could also include tax cuts and subsidies to firms to avoid lay-offs, which often hit unskilled workers first, as discussed in Chapter C.

This dissertation gives important insights into understanding different dimensions of inequality and how it relates to civil conflict, pandemics, and social policy. Although it provides important cross-country evidence on inequalities within countries, policies to address these issues are often implemented at the national and sub-national levels and need to take into account regional differences. Therefore, in order to provide more detailed policy recommendations, further research should aim to complement these findings at a more disaggregated level. This could be done by expanding the dataset for anthropometric indicators at the regional level to provide even more evidence for developing countries and earlier periods. Further identification of the root causes of inequality may explain current social, political or economic outcomes and help to reduce inequality worldwide.

E.1 References

Almond, D. (2006). Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. *Journal of political Economy*, *114*(4), 672–712.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2018). *World inequality report 2018*. Belknap Press.

Barro, R. J., Ursúa, J. F., & Weng, J. (2020). *The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity* (tech. rep.). National Bureau of Economic Research.

Baten, J., & Blum, M. (2014). Why are you tall while others are short? agricultural production and other proximate determinants of global heights. *European Review of Economic History*, *18*(2), 144–165.

Bauernschuster, S., Driva, A., & Hornung, E. (2020). Bismarck's health insurance and the mortality decline. *Journal of the European Economic Association*, *18*(5), 2561–2607.

Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563–595.

Cramer, C., et al. (2005). *Inequality and conflict: A review of an age-old concern*. United Nations Research Institute for Social Development Geneva.

Eichenbaum, M. S., Rebelo, S., & Trabandt, M. (2021). The macroeconomics of epidemics. *The Review of Financial Studies*, *34*(11), 5149–5187.

Einav, L., & Finkelstein, A. (2018). Moral hazard in health insurance: What we know and how we know it. *Journal of the European Economic Association*, *16*(4), 957–982.

Galletta, S., & Giommoni, T. (2022). The effect of the 1918 influenza pandemic on income inequality: Evidence from italy. *Review of Economics and Statistics*, *104*(1), 187–203.

Guimbeau, A., Menon, N., & Musacchio, A. (2022). Short-and medium-run health and literacy impacts of the 1918 spanish flu pandemic in brazil. *The Economic History Review*.

Jordà, Ò., Singh, S. R., & Taylor, A. M. (2022). Longer-run economic consequences of pandemics. *Review of Economics and Statistics*, *104*(1), 166–175.

United Nations Department of Economic and Social Affairs. (2018). Promoting inclusion through social protection: Report on the world social situation 2018. *New York: United Nations*.