IARIW 2022

IARIW 2022

11 Friday - 13 Sunday November

Measuring Multidimensional Inequality and Conflict in Africa and in a Global Comparison

Laura Radatz
University of Tübingen
laura.radatz@uni-tuebingen.de

Jorg Baten
University of Tübingen, CESifo, and
CEPR and Academia Europaea

Paper prepared for the IARIW-TNBS Conference on “Measuring Income, Wealth and Well-being in Africa”, Arusha, Tanzania November 11-13, 2022

Session 2: Inequality

Time: Friday, November 11, 2022 [10:45 AM -12:15 PM]

Measuring multidimensional inequality and conflict in Africa and in a global comparison

Laura Radatz* and Jörg Baten†

Working Paper, October, 2022

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 civil war increases consistently 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.

Multidimensional Inequality, Civil War, Well Being.
JEL: I30, N40, O15

*Corresponding author, Univ. Tuebingen, laura.radatz@uni-tuebingen.de ff
†Corresponding author, Univ. Tuebingen, CESifo, CEPR and Academia Europaea.

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 neighborhood, is a horrible and almost unimaginable experience by those who did not suffer from it. Economists can identify risk factors that increase the probability of occurrence of civil wars 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 have found no 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 beginning with the 1980s on a larger scale particularly for African countries; hence, the available data sets are small (see review of Cramer et al. (2005)). Recently, a new generation of studies have argued that only inequality between groups, rather than nation-wide inequality, is correlated with civil war onset ('horizontal 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.[1] The hypothesis that we study partially contradicts and complements these views, as we assess whether nation-wide 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.

Hence, our study tackles a long-term perspective on the relationship between nation-wide inequality and conflict onset in 193 countries for two centuries with a specific 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 not only include post-tax income inequality, but also health and land inequalities. Evidently, health is an important dimension of welfare, as human 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, Baten, Mira d'Ercole, et al., 2014; see chapter 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.

[1]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 & Mumme, 2013) are also exceptions. They have instrumented nation-wide inequality in a similar way as we do but with a much smaller data set. They also restrict inequality to health inequality.

As a preview of our findings, we observe a positive and robust link between inequality and civil war onset: high inequality increases the probability of civil war outbreaks. 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).[2] Easterly (2007) has argued that land inequality is an important component of 'structural' inequality. 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. Given that within-country inequality has risen dramatically in the United States (US) and the United Kingdom (UK) over the past four decades, one obvious question is whether the likelihood of civil war in the US, the UK and Russia has also increased. According 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 several 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) has discussed the problems around the comparability of inequality data across countries. 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 house-

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

holds. 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) has 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. Banerjee and Duflo (2011) have described how life without proper access to health care 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 (Stiglitz, 2012). 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 data set of inequality that goes beyond income inequality in terms of its time coverage and dimension while still being correlated to it.

In section 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 analyze the relationship between inequality and the probability of a civil war outbreak (using a whole battery of different models, control variables, and specifications). We then assess the robustness of our results in section 3, circumventing potential endogeneity via an IV estimation. We conclude with a discussion of our results and provide policy recommendations.

1 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.

1.1 Height Inequality

The average height of populations is currently a well-established indicator for the quality of nutrition and health care of past populations (Baten & Komlos, 1998; Fogel et al., 1982; Steckel, 1995). Insufficient or poor-quality nutrition, medical care, or shelter during the childhood period determine 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.[3] 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 therefore conclude that the literature interprets height variation to reflect the general inequality within a country to a certain degree. 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, Baten, Mira d'Ercole, et al., 2014).[4]

In our study, we combine height distribution data to obtain a comprehensive view of several facets of inequality by combining height inequality, post-tax income inequality, and land inequality as the three components of a joint index by using the appropriate weights discussed below. As average height is an output-oriented metric that reflects early life conditions in terms of nutrition and health (Blum & McLaughlin, 2019), the advantages of health inequality data may exceed that of income inequality for several reasons as follows: As 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 data set is partly based on the data collection of heights by Baten and Blum (2011), which is available via the website of Clio infra.[5] Moreover, we substantially extend this data set

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

[4]See also literature review in Blum (2014)

[5]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.

in its coverage of countries and years. One important height inequality source is derived from the Demographic and Health Surveys (DHS) program. The DHS are household surveys on a national level, mainly conducted in developing countries. The goal is to monitor and analyze representative data in the fields of population, health, and nutrition.[6] DHS data on height is available for women. Besides the DHS data on 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 height inequality as a coefficient of variation (CV) of the final heights of adults, which is measured in centimeters. We exclude individuals aged below 22 years or older than 50 years from our sample. This factor is because 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 $HeightGini = 33 + 25 \cdot CV$. They based this formula on a multi-country sample of developing countries. This formula has been confirmed by other close estimates (Van Zanden, Baten, Foldvari, et al., 2014).

For our analysis, we concentrate on ten-year periods to eliminate year-specific random fluctuations. We calculate the height CV for each birth decade of the respective country; for example, 1910 includes the years 1910–1919. After applying the described restrictions and dropping observations with missing or obviously false information, we construct a data set containing 928 height Gini values for the period of 1810–2000 for 127 countries worldwide. We include a detailed overview of countries and the periods in our sample, and its sources are given in the Appendix. 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 (Table 3).

When comparing our anthropometric inequality measures to income inequality data, we anticipate a positive correlation between height and income inequality. 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 following analysis, we compare various inequality databases and indicators. For this comparison, we use data from the OECD, World Income Inequality Database (WIID), World

[6] Demographic and Health Surveys (DHS) Program: https://dhsprogram.com/Methodology/Survey-Types/DHS.cfm, derived 26/06/2021.

Inequality Database (WID), and World Bank (Milanovic, 2013; Van Zanden, Baten, Mira d'Ercole, et al., 2014). In the case of 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 (OECD, 2020). Second, we include data for income Gini coefficients from the World Bank.[7] As the data is based on household surveys, it only covers the periods starting from 1974 (Bank, 2021). The 'All the Ginis Dataset' compiled by Branko Milanovic combines eight available databases into one broad data set covering the period of 1950–2012. This data set is available from the World Bank and includes 187 economies on a yearly basis (Milanovic, 2013). Van Zanden, Baten, Mira d'Ercole, et al. (2014) have provided a long-run data set on income inequality to study within-country inequality with observations from 1820 until 2000. They use various sources to construct a broad data set, including 'Williamson' estimates that are based on the ratio between GDP and real wages of unskilled laborers (Van Zanden, Baten, Foldvari, et al., 2014). The WID, created by the Piketty group, collects historical income data in the last decade (Alvaredo et al., 2020). By using tax statistics and different surveys, the WID team constructed a database that provides long-run data for income and wealth distribution with a huge country and time coverage.[8] Top income shares are available for up to 11 countries from 1870 to 1910 and 74% of all countries after 1990.[9] 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 inequality data fills many remaining gaps and allows us to check several dimensions of inequality beyond income. With the different databases of income inequality, we now assess how the measurement of health inequality is related to different data on income inequality. As shown in Figure 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 exhibits an equitable distribution of health in the 1960s and 1970s. In the upper right corner of Figure 1, Mexico emerges with very high differences in height distribution and high levels of income inequality in the 1970s and 1980s. Even with a high economic growth in the previous decades, high health inequality is still observed (López-Alonso & Condey, 2003). The correlation coefficient of height and income inequality is significant (ρ=0.34, P<.000). As displayed in Table 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

[7]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.

[8]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)

[9]Other databases include Milanovic, Lindert, and Williamson, 2008; and World Income Inequality Database, with earliest data from 1867.

the WID data set, we add the data based on their top 1% shares that are transformed into income Gini coefficients for comparison (Table 2). We also compare the Gini coefficients of income only from Van Zanden, Baten, Foldvari, et al. (2014).

The strength of height inequality data is to provide evidence for developing countries, especially 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 inequality is less informative for these rich countries). Hence, it is particularly important to obtain a good coverage of our component 'income inequality' 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 most pronounced in agricultural countries (such as Russia before the mid-20th century). A combination of all three measures therefore provides a unique and broad coverage.

1.2 Land Inequality

High land inequality values indicate the degree to which large landowners have control over land as production factor.[10] In the case of Latin America, Frankema (2005) has shown that the inequality of land distribution caused by colonial rule is accompanied by high income inequality (see also Baten and Juif (2014) and Qasim et al. (2020)).

Our sample for land inequality is based on those of Frankema (2005, 2010)[11] 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 FAO (2019). Furthermore, we include data from WCARRD (1988) to fill missing decades for several countries and add some missing countries, covering a total of 143 countries.

Land inequality is typically not changing significantly over time, unless a successful land reform or substantial industrialization development has occurred, during which laborers 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 collection of land reforms, we interpolate the data on Gini coefficients of land backwards in time

[10] Landowners are defined as those who produce on their own or on rented land.

[11] Frankema (2010) is the updated and corrected version.

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 coverages, we also expand the data set on land reforms. Therefore, 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 A.2.2 and Table A.2. 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 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.

1.3 Calculation of the Multidimensional Joint Index

This study aims to construct a broad data set of inequality, which goes beyond post-tax income inequality. After identifying height 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 data sets. 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, where the latest version was released in May 2021. This data set provides data for up to 170 countries from 1990 to 2019 (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 2. Using this technique, we estimate the missing Gini coefficients for 1870 to 2010 and for additional 6 countries.

Given the high number of estimated income Ginis combined with our newly collected height inequality 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 urbanization rate at country i at time t. We weight the different dimensions of inequality as follows: $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$.[12] Land inequality matters less for highly urbanized and less agricultural societies. By including a weighting for the degree of urbanization, we consider

[12]An explanation of the weighting procedure is given in the Appendix.

the degree of urbanization. For example, in Sierra Leone, more than a half of the total population of over six million people live in rural environments. Based on the data from 2013, over 60% of the working population work in the agricultural sector, where women are responsible for harvesting and processing the cassava crop, whereas men typically take care of rice cultivation and tree crops.[13] By contrast, in Chile in 2013, only 10.8% of the population live in rural areas, and agriculture accounts for 10.3% of employment.[14] By combining the data of income inequality with health and land inequality data, we construct a data set for 193 countries and cover a period of 1810–2010. This data reflects an average of 77% of the world's population over the last 200 years, as shown in Figure 2. The coverage of our joint index as percentage of the world population is displayed in Table 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 1890. When only height inequality values are available, we include these to fill gaps (controlling these cases with an appropriate dummy variable strategy). By doing so, we address some concerns mentioned above by providing a broad data set for long-term analysis and various countries.

1.4 Cross Validation of the new measure: How our anthropological inequality measure correlates with income measurements.

In Table 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 1, which implies that the new index has some remaining value added.

In Figure 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 19th century is quite stable. Inequality has decreased in the early and mid-20th century and started to increase again after the 1980s (consistent with Lakner and Milanovic (2016)).

In Figure 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 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

[13] http://www.fao.org/gender-landrights-database/country-profiles/countries-list/general-introduction/en/?country_iso3=SLE, 09/06/2021.

[14] http://www.fao.org/gender-landrights-database/country-profiles/countries-list/general-introduction/en/?country_iso3=CHL, 09/06/2021.

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.

1.5 Conflict Data

In this section, we analyze the distribution of civil war onsets worldwide. Data on conflict is available from the Correlates of War Project (COW). This database tracks different kinds of violent conflicts worldwide and provides data on non-state wars, intrastate wars, and interstate and extra-state wars. As we analyze the impact of inequality on conflicts on a country level, we focus on conflicts that occur within national borders, namely, intra-state conflicts. Conflicts arise due to complex socioeconomic interactions and motivation (Raleigh & Kniveton, 2012). In our analysis, we therefore distinguish between different types of civil wars as provided by the COW, which are civil war over central control and over local issues.[15] 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 data set on intra-state wars (v5.1) covers the period from 1816 to 2014. We exclude the decade of the 2010s, since some conflicts occurred after 2015.[16]

This data set provides the broadest coverage for conflict over the long run. For our regression analysis, we include civil war onset from all types as a dummy variable that takes the value 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 6, we report the number of civil war outbreaks over time from our sample and the unequal distribution of civil war outbreaks in different world regions on the right side of Figure 6.

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 Sub-Saharan

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

[16]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).

Africa. 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 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 8, we control for different types of civil wars. We observe that most civil wars are about central control, followed by civil war over local issues. In order to test the impact of inequality on different aspects of civil wars, we also measure civil wars by the severity of the conflicts in the respective decade.

1.6 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, the quality of institutions displayed by the polity2 index[17], colonial background, the history of wars, ethnic fractionalization, and diamond deposits. First, population size is a necessary control variable, as a large country, such as China, has almost automatically a higher likelihood of a civil conflict in one corner of the country compared with a small country, such as Portugal. Second, the quality of institutions and democratic decision processes may matter, as Collier and Hoeffler (2004) or Fearon and Laitin (2003) has found that the institutional and political contexts have an impact. A consolidated democracy faces a low risk of civil wars. We control for the colonial history of a country by including a categorical variable. A highly fractionalized country in terms of ethnicity, language, or religion may face an increased risk of a civil war outbreak; therefore, we include the measures mentioned in our analysis. We include GDP per capita as control for the economic development of a country (and for similar purposes, we also include height growth). A detailed description of the control variables and their sources are included in the Appendix. In Table 5, we show the summary of statistics if the main variables civil war and inequality are available.

[17] 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).

2 Does inequality fuel violent conflict?

A high level of inequality tends to undermine social cohesion and fuel protest and violent conflicts (e.g.,Vergolini, 2011). Does a high level of inequality increase the probability of a civil war? Three main views are discussed on the economic causes of civil wars. Collier and Hoeffler (2004) favor 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, people have lower opportunity costs in joining a violent movement in contrast with those in richer populations. This factor provides armed opposition groups with a larger number of people with low opportunity costs in poorer economies. By contrast, the argument for grievance is the civil war motivation based on inequalities. The motivation for people to change the status quo must be sufficiently high to join a violent conflict to solve those issues.

Fearon and Laitin (2003) have supported the view that civil wars mainly occur in countries with weak institutions. Therefore, in their view, state capacity matters more than the motivation of people. By contrast, civil wars in Latin America are often explained by the grievance argument that land and income inequalities are high and crucial for understanding the conflict in this region. Jensen and Sørensen (2012) have checked the association between land inequality and civil conflict using a panel of 18 Latin American countries spanning the 20$^{\text{th}}$ century. Their empirical study reports a significant relationship between land inequality and civil war onset. Their finding confirms the impact of inequality on conflict (also consistent with the model of Acemoglu and Robinson (2006)).

2.1 Regression Results

The regression results of the likelihood of civil war onset are displayed in Table 6. In nine regression analyses, we identified the potential correlates of civil war onset. We have chosen the standard model selection strategy to first assess a bivariate regression of inequality and civil war onset (column 1) and then added time-fixed effects (column 2), world region, and time-fixed effects (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 in every model 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 – 0.45 to 0.81 – expect 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 consistent with Collier and Sambanis (2002). The higher the population of a country, the higher the probability of a civil war outbreak. As expected, a high level of democratization reduces the risk of a civil war in a country. However, this effect 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 squared terms as well. Colonial background shows no significant effect here, neither positive nor negative (different results: Baten and Mumme (2013)). The same result is true for the 'greed' dummy variable diamond, which is included in models 4, 6 and 8. The 'greed' proxy for (low) income has a significant effect (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, language, or religious groups increases the risk of civil war, we observe the positive effect of ethnic fractionalization. In addition, the country's history of previous wars (*history of wars*) has a significant influence on the outbreak of future civil wars. However, we mainly observe this effect in our rare event model (column 9). Conflicts over local issues or political power can always reemerge if remain unresolved, as seen in Israel/Palestine, where civil wars arise repeatedly. In model 7, we include height growth as a proxy for economic growth. We have found no evidence that this variable reduces or increases the risk of civil war.

In order to assess potential selectivity, we study how well our data set – used in several regression specifications in Table 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 percent, and for the – always least documented – 'low-income' countries, we still have a respectable 13-15% of all possible country-decade combinations (Figure 9).

2.2 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 7. We also compare the results using country fixed effects as opposed to region fixed effects. The coefficient of inequality is virtually identical[18] (Table 8).

[18]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

Does inequality affect different types of civil wars differently? In Table 9 we consider different types of intra-state 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. We find that inequality significantly increases the risk of civil war, both over central control and over local issues, whereas the coefficient for the latter is slightly larger. If we multiply with 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 relative higher effect of inequality in fueling civil wars over local issues rather than over central control.

In model 3 we test for the effect of inequality on the severity of a civil war, measured as average civil war deaths in the particular decade and country. We find that higher with-in country inequality leads to substantially more severe conflicts.

We also control for differences between low- and high-income countries. In Table 10, we omit poor countries in model 1 and very rich countries in model 2. In both models, the estimated influence of inequality on the risk of civil war is highly significant and positive. Finally, we assessed whether the inequality effect depends only on our composite measure of inequality. We used height inequality alone in Appendix Table A4 (because height inequality covers the largest amount of observations) and we find that the inequality effect is robust using only this inequality concept.

2.3 IV Regressions

To circumvent potential endogeneity issues, we apply an IV approach. For example, reverse causality can be an issue: civil war may affect inequality in a country, as Bircan et al. (2010) have noted that during and after wars, economic activities significantly decrease. 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 labor 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, named *wheat-sugar ratio*, refers to the suitability of the soil for sugar divided by the suitability of the soil for growing wheat. This approach is implemented in several studies, sometimes with further modifications (Baten & Juif, 2014; Baten & Mumme, 2013).[19] 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 primary labor force in earlier times; for example, in the early 19$^{\text{th}}$ century Brazil,

sample size in other specifications.

[19] 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.

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 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 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.[20] 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 the total exports. As the export coefficient is insignificant and does not affect the inequality effect, this would not suggest a violation of the exclusion restriction (see column 5 of Table 10).

[20] 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.

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 joint 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 data set 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 inequality significantly increases the risk of a civil war outbreak in a country. Our results remained robust to the application of different models, including various sets of control variables and time- and region-fixed effects. We also considered the robustness when using only height 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 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%.[21] 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.

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

References

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., Martınez-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.

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

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. (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.

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.

Bircan, C., Brück, T., & Vothknecht, M. (2010). Violent conflict and inequality.

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.

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

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

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.

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., Hrazdıra, 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.

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

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.

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). *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*.

OECD. (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*(1), 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.

UNU-WIDER. (2021). 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. (2014). 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., et al. (2014). *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.

WCARRD. (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.

Tables and Figures

Figure 1: Relationsship between Height Gini and Income Gini.

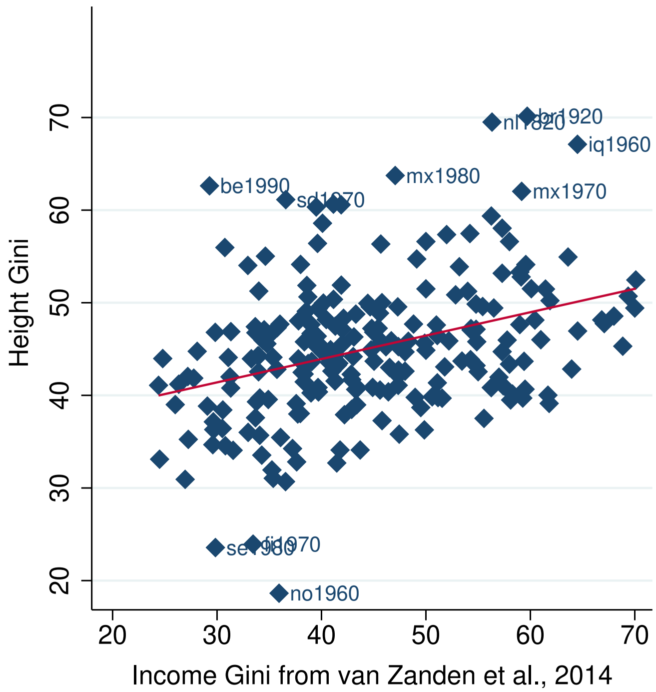


Table 1: Comparison between income Gini and height Gini: regressions based on selected different data sets.

	(1) Van Zanden et al. 2014a	(2) WID	(3) Top10% 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*

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

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

	Income Gini WID
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 top 1% income share and the Gini coefficients for post-tax income are from the World Inequality Database (WID).

Figure 2: Distribution of our Joint Inequality Index by the percentage of the world population covered.

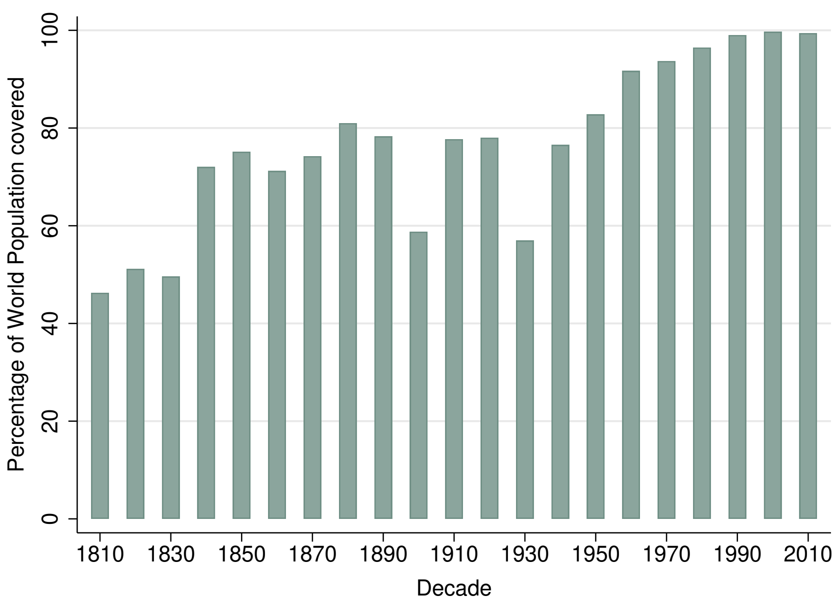


Table 3: Overview of joint inequality index – time and region coverage in percentage of the world region's population.

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 4: Correlation between multidimensional inequality index and income Ginis from different sources.

	Joint Inequality Index
Income Gini OECD	0.97***
Income Gini WID	0.63***
Income Gini Milanovic 2013	0.66***
Income Gini Van Zanden et al. 2014	0.57***

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

Figure 3: Comparison of Different Inequality Measures.

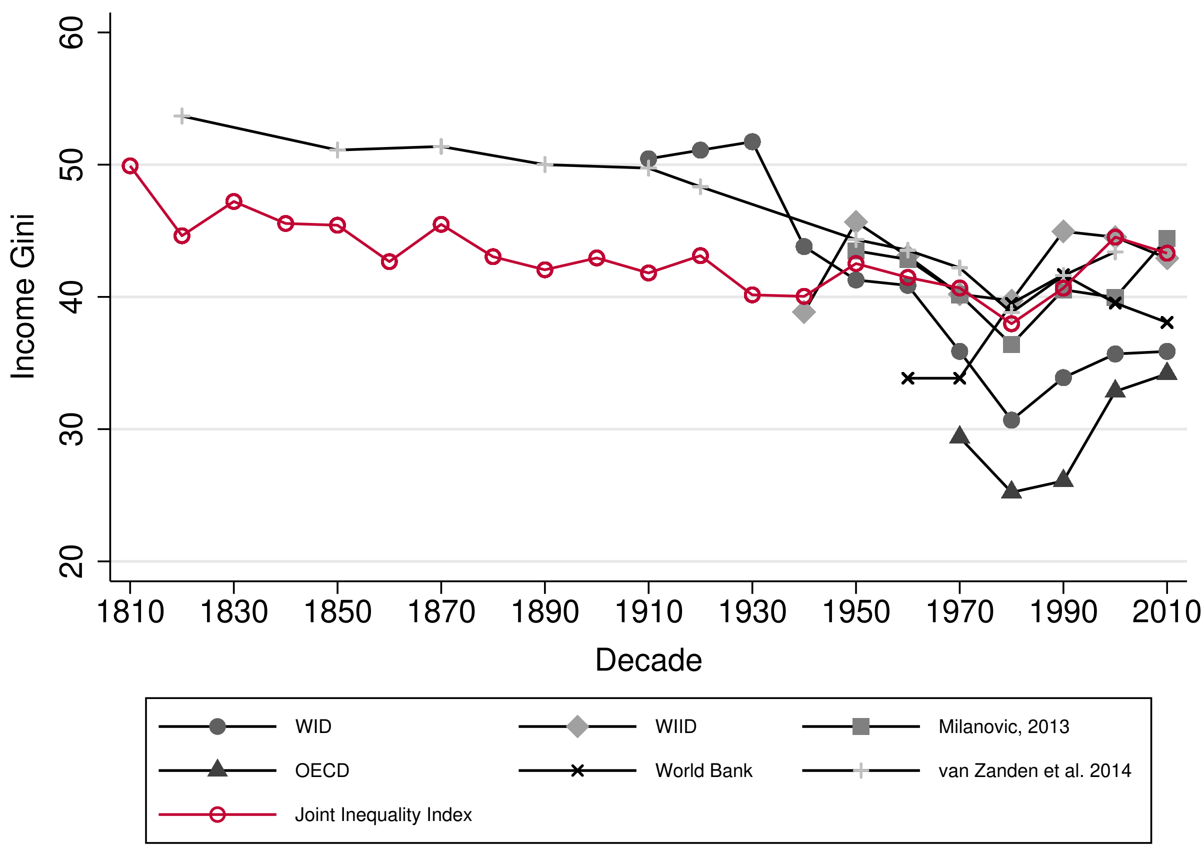


Figure 4: Inequality index, most recent data available.

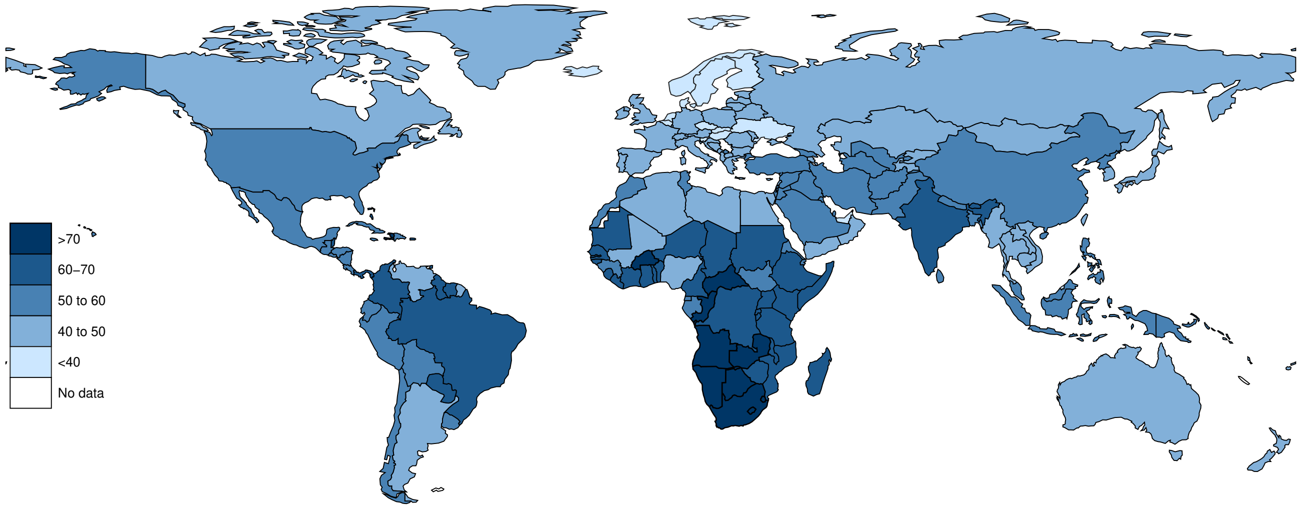


Figure 5: Inequality by World Regions.

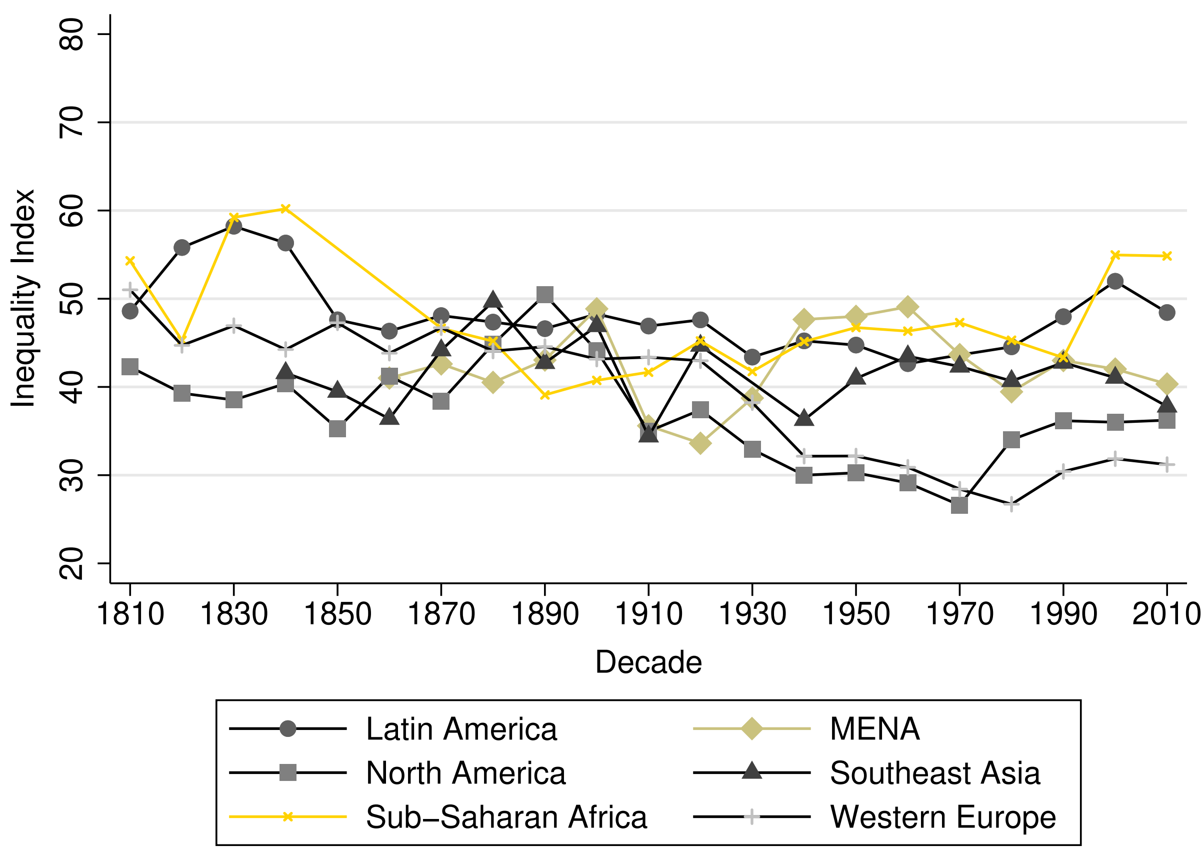
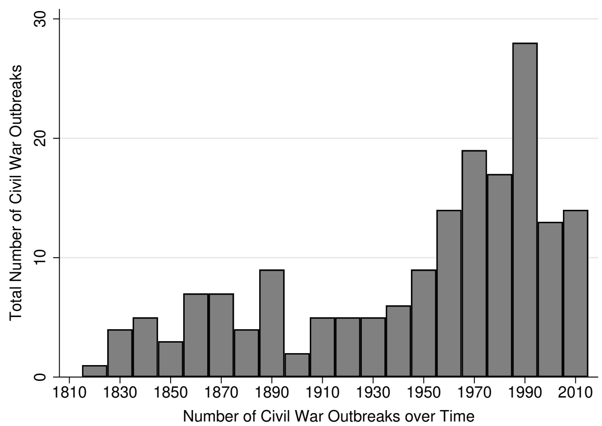


Figure 6: Number of Civil War Outbreaks over Time and by World Region.



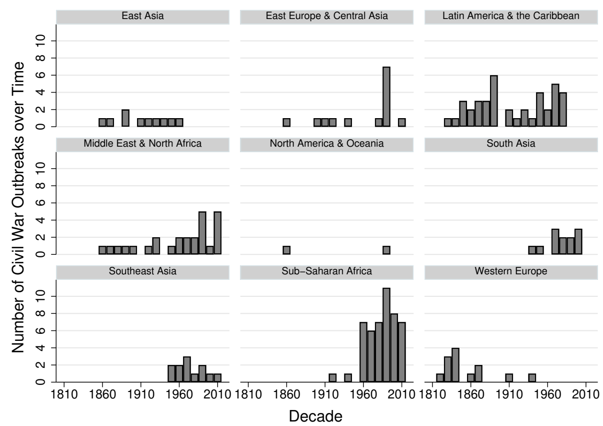


Figure 7: Civil War Outbreaks worldwide - 2000.

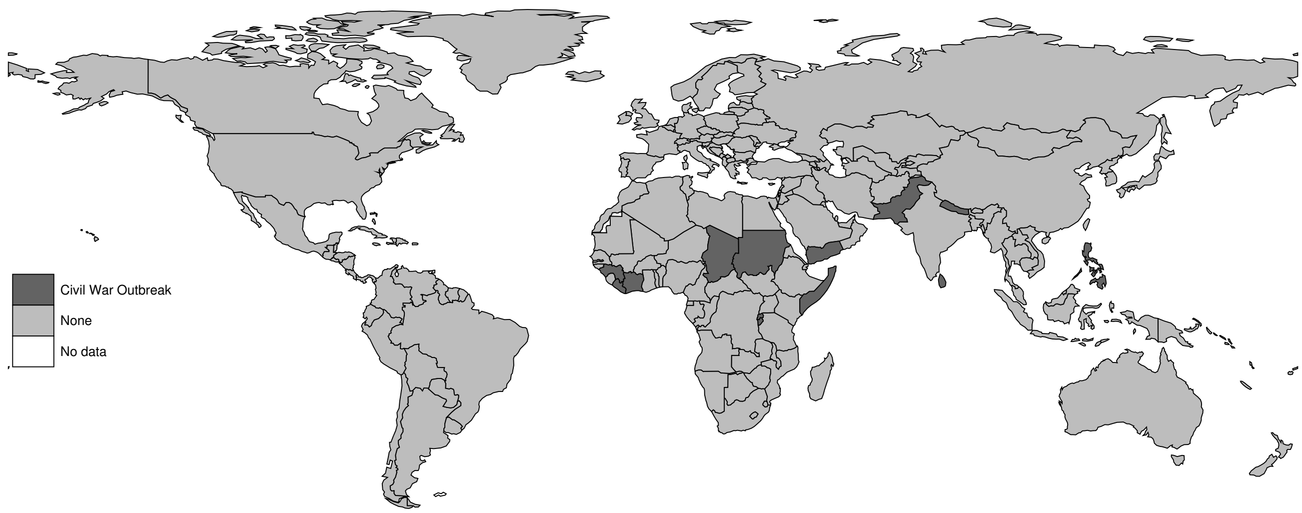


Figure 8: Civil War Outbreaks by Type.

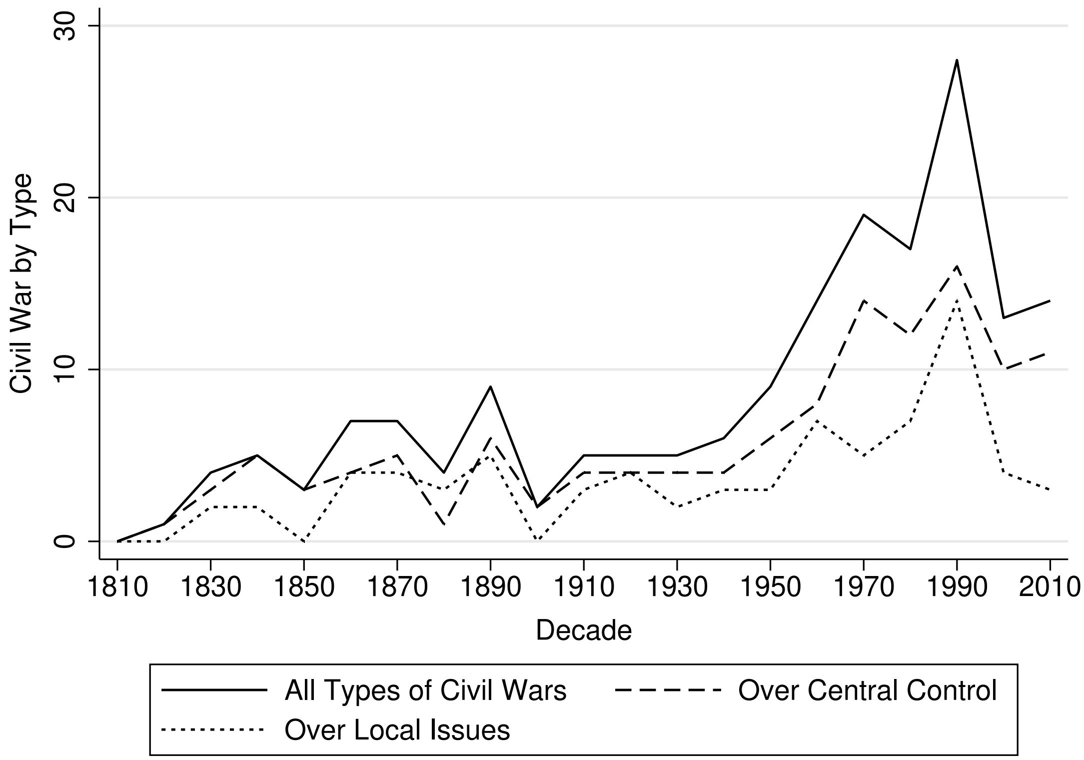


Table 5: Summary statistics.

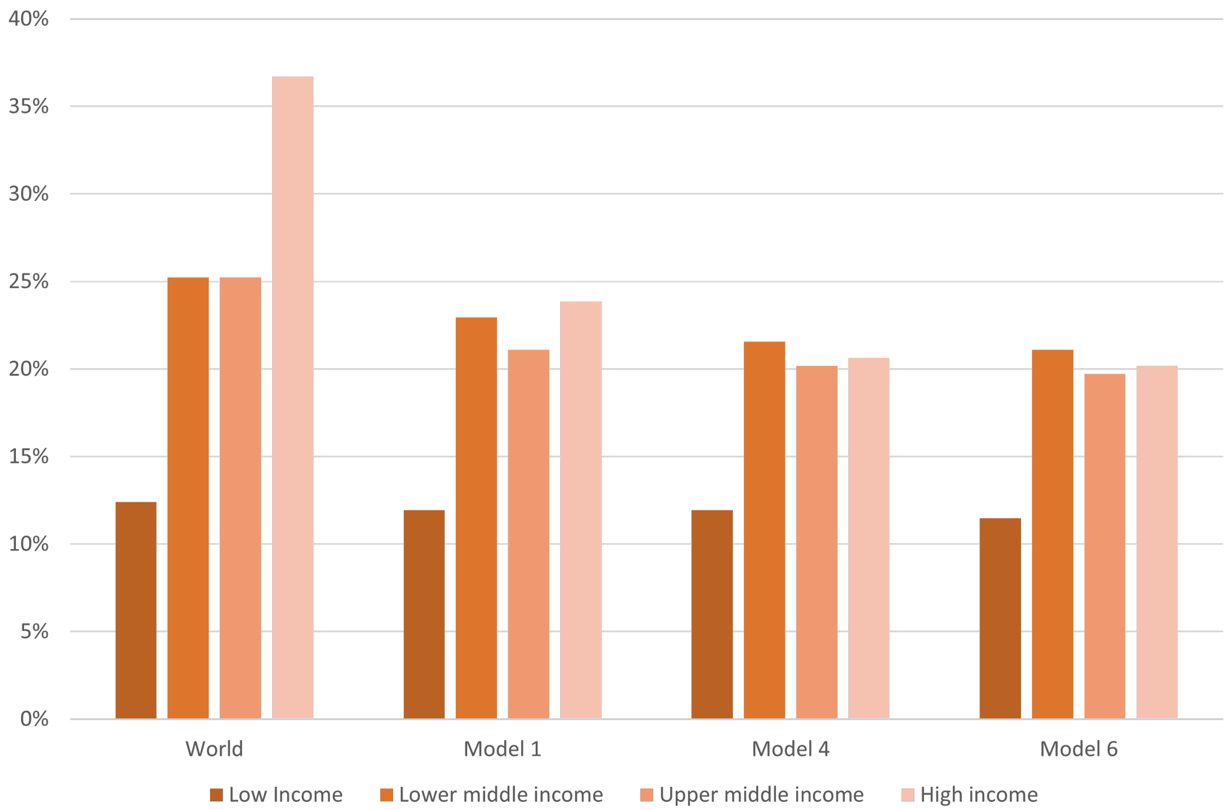
	N	Mean	SD	Min	Max
Civil War Onset	1190	0.11	0.32	0	1
Severity of Civil War	1190	2459.16	20529.86	0	515612.5
Inequality (lag)	1190	0.42	0.09	.1723938	.8133366
Population (log)	1182	15.83	1.56	11.30382	20.95647
Democracy	873	0.01	0.07	-.1	.1
Democracy^2	873	0.48	0.36	0	1
Diamond Deposits	1189	0.17	0.38	0	1
History of Wars	1190	0.22	0.74	0	6
Colony	1189	0.13	0.34	0	1
Ethn. Frac.	1140	0.43	0.27	0	.9302
Height Growth	657	0.00	0.02	-.084287	.1259581
GDP per capita	933	7624.22	9562.11	485.735	77798.96

Table 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. fract.					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	1182	1182	850	850	827	827	447	753	873
Time Fixed Effects	*Y*	*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). Diamond and colonial history are dummy variables. Fractionalization measures are time invariant. For expository purpose, democracy and democracy squared are divided by 100 before running the regressions. GDP per capita is divided by 10,000 before running the regression. Inequality is composed by income Gini, height Gini and land Ginis which are all lagged by 1 decade.

Figure 9: Data selection.



Notes: The graph shows the data selection in the models of Table 6 by income group as defined by the World Bank. Income group classification is derived from the World Bank, World Development Indicators: 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 7: OLS Regression of Civil War Onset.

	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) Pooled OLS	(5) Random effect model
Inequality	0.70***	0.56***	0.45**	0.45**	0.85***
	(0.128)	(0.140)	(0.180)	(0.176)	(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		−0.14	0.11	0.03	0.43
		(0.206)	(0.216)	(0.228)	(0.379)
Democracy^2		−0.16***	−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. fract.			0.10	0.09	
			(0.091)	(0.078)	
GDP p.c.				0.02	
				(0.013)	
Height growth					−1.03
					(2.004)
Constant	−0.85***	−0.58***	−0.82***	−0.66***	−0.99**
	(0.147)	(0.177)	(0.202)	(0.240)	(0.451)
Observations	1182	873	850	789	460
R-squared	0.0911	0.1726	0.1584	0.1999	0.1806
Time Fixed Effects	*Y*	*Y*	*Y*	*Y*	*Y*
Region Fixed Effects	*N*	*N*	*Y*	*Y*	*Y*

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

Table 8: OLS Regression with regional and country fixed effects.

	(1) Pooled OLS	(2) Pooled OLS
Inequality	0.53^{***}	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	1182	1182
R-squared	0.1116	0.3189
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 by income Ginis, height Ginis and land Ginis where data is available and lagged by one decade.

Table 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**	14.14**
	(0.116)	(0.148)	(5.996)
Population (log)	0.02**	0.03***	1.69**
	(0.007)	(0.008)	(0.672)
Democracy	−0.03	−0.03	−2.95
	(0.186)	(0.142)	(10.586)
Democracy^2	−0.13***	0.00	−2.54
	(0.034)	(0.025)	(2.199)
History of wars (by type)	0.10***	0.06**	2.55***
	(0.024)	(0.026)	(0.942)
Constant	−0.33**	−0.56***	−35.69**
	(0.135)	(0.166)	(13.991)
Observations	873	873	873
R-squared	0.139	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. For interpretation, democracy and democracy squared are divided by 100 before running the regressions. Inequality is composed by income Ginis, height Ginis and land Ginis, which are all lagged by one decade.

Table 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^{***}	0.70^{***}
	(0.152)	(0.196)
Population (log)	0.04^{***}	0.04^{***}
	(0.011)	(0.014)
Democracy^2	-0.18^{***}	-0.24^{***}
	(0.051)	(0.059)
Democracy	-0.46^{**}	-0.38
	(0.216)	(0.309)
Diamond	-0.04	-0.02
	(0.043)	(0.059)
Colony	-0.03	-0.01
	(0.095)	(0.059)
Observations	687	666
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. Diamond and colonial history are dummy variables. For expository purpose, democracy and democracy squared are divided by 100 before running the regressions. Inequality is composed by income Ginis, height Ginis and land Ginis, which are all lagged by one decade.

Table 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^{***}	5.36^{***}	5.31^{***}	2.68^{***}
	(0.247)	(1.661)	(1.674)	(1.630)	(0.334)
Population (log)	0.07^{***}	0.06^{***}	0.06^{***}	0.06^{***}	0.08^{***}
	(0.014)	(0.021)	(0.021)	(0.021)	(0.018)
Democracy	0.62^{**}	0.95^{*}	0.95^{*}	0.94^{**}	
	(0.248)	(0.499)	(0.498)	(0.479)	
Democracy^2	-0.06	-0.14^{*}	-0.14^{*}	-0.14^{*}	
	(0.051)	(0.084)	(0.084)	(0.081)	
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	-2.44^{***}	-2.66^{***}	-2.67^{***}	-2.66^{***}	-2.21^{***}
	(0.247)	(0.745)	(0.756)	(0.741)	(0.276)
Observations	843	843	843	843	328
Adj. R-squared	0.218	0.346	0.345	0.347	0.09
Time Fixed Effects	*Y*	*Y*	*Y*	*Y*	*Y*
Region Fixed Effects	*N*	*Y*	*Y*	*Y*	*N*
F-statistic	153.62	20.65	20.45	15.75	82.33
Kleinbergen-Paap rk LM statistic	Exactly identified				
Hansen J statistic	Exactly identified				

P-values in parentheses * p 0.05, ** p 0.01, *** p 0.001. Heteroscedasticity- and cluster-robust standard errors. Dependent variable in first stage is inequality. Dependent variable in second stage is civil war onset. Diamond and colonial history are dummy variables. For expository purpose, democracy and democracy squared are divided by 100 before running the regressions. Inequality is composed by income Ginis, height Ginis and land Ginis, which are all lagged by one decade.

A Appendix

A.1 Variable definitions

Civil War. We use the outbreak of civil war as a dependent variable. Civil war is coded as a dummy variable, which takes the value 1 if a civil war outbreak has occurred in the regarded country and decade. The Correlates of War Project (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 disputes 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 for the severity of a conflict. Source: Sarkees, Meredith Reid and Frank Wayman (2010). Resort to War: 1816–2007. Washington DC: CQ Press.

Colony. We control for colonial history, where 1 indicates that the country was a colony and 0 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 is a highly consolidated democracy. We use democracy and democracy squared in our regressions. Source: Polity5 Project, https://www.systemicpeace.org/polityproject.html.

Diamond. Coded as a dummy variable that takes the value 1 if a diamond deposit is/was present in a country and 0 if not.

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: Maddison Project Database, version 2020. Bolt, Jutta, and Jan Luiten van Zanden (2020), 'Maddison style estimates of the evolution of the world economy. A new 2020 update'.

Height growth. Growth rate of heights between two periods.

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

Population (log). The log of the total population of a country at the beginning of a ten-year period. Source: World Bank and Maddison (2001).

History of Wars. Indicates whether a civil war occurred in the previous period. It counts the number of decades. Source: Correlates of War Project.

A.2 Construction of multidimensional inequality

To construct our multidimensional inequality index, we combine income, health, and land inequality data.

A.2.1 Health Inequality

Our data set on height Ginis is partly based on the data collection of heights by Baten and Blum (2011), which is available via the website of Clio Infra. We extend this data set with data from different surveys and individual height studies. In Table A1, we provide an overview of the updated data country coverage, period, and sources used in our data set. For the data used from the data set of Baten and Blum (2011), an overview of the data sources can be found via Clio Infra.[22]

A.2.2 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 data set with the data from the latest 2019 FAO report. We calculate the Gini coefficients for 91 new observations based on the formula from Frankema (2005):

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

where $n = amount\ of\ decile\ shares = 10$.

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 of land inequality after the land reform by 5.57 Gini points. However, if land reform is unsuccessful, then land inequality remains unchanged. Following the methods of Baten and Juif (2014), we replicate the effect of land reform and obtain an estimated average effect of a land reform 4.47 Gini points (Table A3). To do so, we also 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) (5.57 Gini points) may be explained by a high country and time coverage, including recent years from 2000–2010, where land reforms may not have that high effect on land inequality it had during the 1900s. Based on the estimated average impact of a land reform on land inequality, we adjust our data set by subtracting 4.47 Gini points to the following period if a land reform is successful. In this manner, we can gain several observations for our analysis.

[22]https://clio-infra.eu/Indicators/HeightGini.html.

A.2.3 Joint Inequality Index

We construct our inequality variable by combing our three indicators, income, health and land inequality to one joint inequality index (Table A2 report our sources for income Ginis). We used the common method of assigning equally weight to each dimension. We hereby follow a normative approach on the construction of our joint index, which is also done to construct the Human Development Index (HDI), which takes the geometric mean of three dimensions to construct on single indicator, the Gender-related Development Index (UNDP, 2013) and used by Alkire and Foster (2011) on the construction of a multidimensional poverty index. In Table A4 we include sensitivity analysis for the joint index due to different construction decisions and weighting. Our results remain robust. In model 1 we do not weight height Gini by the degree of urbanisation. In model 2-3 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 2 and $\alpha_1 = 0.2, \alpha_2 = 0.4$ and $\alpha_3 = 0.4$ in case of model 3. Model 4 exclude the observations for income Ginis, which we calculated based on the top 1% income share from the WID. Finally, model 5 displays the regression results if we use health inequality as only indicator for inequality.

Table A1: 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
Gambia	gm	1960 - 1990	Demographic and Health Survey

Table A1: (continued)

Country	ccode	Birth Decade	Source
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 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 A1: (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, Frederick P. 1959
		1980	Hossain, Lestrel and Ohtsuki 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 A2: 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

Table A2: (continued)

Country	ccode	Decade	Source
		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
Bosnia and Herzegovina	ba	1980 - 1990	WID, 2021
		2000 - 2010	WIID, 2021
Botswana	bw	1980 - 2010	WIID, 2021
Brazil	br	1970 - 2010	WIID, 2021

Table A2: (continued)

Country	ccode	Decade	Source
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 A2: (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

Table A2: (continued)

COUNTRY	CCODE	DECADE	SOURCE
Gambia	gm	1990 - 2010	WIID, 2021
Georgia	ge	1980	WID, 2021, calc. top 1 % income share
		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

Table A2: (continued)

Country	ccode	Decade	Source
Israel	il	1980 - 2010	WIID, 2021
Italy	it	1940	WIID, 2021
		1960 - 2010	WIID, 2021
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

Table A2: (continued)

Country	ccode	Decade	Source
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
		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

Table A2: (continued)

Country	ccode	Decade	Source
Nicaragua	ni	1990 - 2010	WIID, 2021
Niger	ne	1960	WIID, 2021
		1990 - 2010	WIID, 2021
Nigeria	ng	1950	WIID, 2021
		1980 - 2010	WIID, 2021
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

Table A2: (continued)

Country	ccode	Decade	Source
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
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

Table A2: (continued)

Country	ccode	Decade	Source
		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
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

Table A2: (continued)

Country	ccode	Decade	Source
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
		1970	WIID, 2021
		1990 - 2010	WIID, 2021
Zimbabwe	zw	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 World Inequality Database (WID).

Table A3: 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

Robust standard errors in parentheses *** p0.01, ** p0.05, * p0.1

Table A4: Sensitivity test for the construction of the joint inequality index.

	(1) Joint Index I	(2) Joint Index II	(3) Joint Index III	(4) Joint Index IV	(5) Height Gini
Inequality	0.37**	0.49***	0.46***	0.51***	0.71***
	(0.164)	(0.162)	(0.149)	(0.157)	(0.207)
Population (log)	0.03***	0.03***	0.04***	0.03***	0.04**
	(0.010)	(0.010)	(0.010)	(0.011)	(0.015)
Democracy	−0.05	−0.01	0.03	−0.10	0.06
	(0.213)	(0.208)	(0.208)	(0.228)	(0.304)
Democracy^2	−0.12***	−0.13***	−0.13***	−0.13***	−0.13***
	(0.038)	(0.038)	(0.038)	(0.041)	(0.048)
Diamond	0.00	0.01	0.01	0.01	0.03
	(0.045)	(0.045)	(0.045)	(0.048)	(0.064)
History of Wars	0.08***	0.08***	0.08***	0.08***	0.05***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Colony	−0.02	−0.02	−0.02	−0.01	−0.02
	(0.065)	(0.064)	(0.064)	(0.068)	(0.069)
Constant	−0.63***	−0.71***	−0.70***	−0.67***	−0.85***
	(0.193)	(0.191)	(0.189)	(0.211)	(0.283)
Observations	873	873	873	807	630
R-squared	0.18	0.185	0.186	0.184	0.175
Time Fixed Effects	*Y*	*Y*	*Y*	*Y*	*Y*
Region Fixed Effects	*Y*	*Y*	*Y*	*Y*	*Y*

Notes: Heteroskedasticity-robust clustered standard errors in parentheses. ***, **, * significant on the 1, 5, and 10%-level, respectively. Note: Diamond and colonial history are dummy variables. Fractionalization measures are time invariant. For interpretation, democracy and democracy squared are divided by 100 before running the regressions. Inequality is composed by income gini, height gini and land ginis which are all lagged by 1 decade. Alternative calculation for joint inequality index is used in model (1)-(4).