DOES INEQUALITY LEAD TO CIVIL WARS? A GLOBAL LONG-TERM STUDY USING ANTHROPOMETRIC INDICATORS (1816-1999)

This is the last working paper version before this study was submitted and accepted.

Please cite as Baten, Joerg and Mumme, Christina (2011). "Does Inequality Lead to Civil Wars? A global long-term study using anthropometric indicators (1816-1999)", *European Review of Political Economy* 32 (2014), 56-79. Online:

https://doi.org/10.1016/j.ejpoleco.2013.06.007

JOERG BATEN

Univ. Tuebingen and CESifo
email: joerg.baten@uni-tuebingen.de
postal address: Mohlstrasse 36
72074 Tuebingen, Germany
Phone: +49 - (0)7071 - 29 - 72 98 5
Fax: +49 - (0)7071 - 29 - 51 19

CHRISTINA MUMME

Univ. Tuebingen
email: christina.mumme@uni-tuebingen.de
postal address: Mohlstrasse 36
72074 Tuebingen, Germany
Phone: +49 - (0)7071 - 29 - 78 19 2
Fax: +49 - (0)7071 - 29 - 51 19

Abstract

We test the influence of absolute and relative deprivation – proxied by anthropometric methods – on civil war risk. A comprehensive height data set allows extending the study period as far back in time as 1816, for a global sample. We measure absolute deprivation by using lagged human stature and we use height inequality within birth cohorts to measure relative deprivation. We take care that selectivity caused by missing values does not bias the results. We find that relative economic deprivation within populations had in fact a strong and consistent impact on the propensity to start civil wars. Absolute deprivation was significant in most but not all specifications. Included are controls for a number of variables such as population size, regime type, ethnic fractionalization, colonial status and other factors. We also take care of potential endogeneity by introducing an instrumental variables approach.

Keywords: civil wars, anthropometric welfare indices, economic deprivation

JEL codes: H56, N30, N40, 015

I. Introduction and Literature Overview

Since the end of World War II, 244 armed conflicts have been documented, 134 of which continued after the end of the Cold War. The majority of these conflicts took place in the intrastate realm, with only seven interstate wars listed for the period from 1989 through 2003 (Harbom et al. 2006; Eriksson and Wallensteen 2004; Gleditsch et al. 2002).[2] For the year 2011 37 ongoing armed conflicts were registered, 27 of them being pure intra-state conflicts, 9 internationalized conflicts and one interstate conflict.[3] The consequences of these wars are devastating; they do not only cause tremendous human suffering – from 1945 to 1999 internal conflicts resulted in 16.2 million dead victims and an even higher number of refugees (Fearon and Laitin 2003: 75) – but they also ruin the economies of the states they are situated in, condemning their population to dwell in economic underdevelopment for years after the conflicts end.[4] Children in conflict-affected countries are twice as likely to be undernourished and three times as likely to be out of school compared to children living in another developing country (World Bank 2011: 51). Moreover, internal conflicts almost always affect and involve neighboring states, undermining regional stability, and they often engage the interests of distant powers and international organizations (Brown 1996: 3). Furthermore, the response to intrastate conflicts during and after their occurrence costs the international community sizeable sums every year.[5] These funds could otherwise be directed towards much needed developmental aid.

A new wave of intrastate conflicts has emerged in 2010 with the “Arabian Spring” in Northern Africa, drawing worldwide attention to the topic of civil wars and its determinants.

---

[2] Brown (2003) only lists four interstate conflicts for this period: the 1991 Gulf War, the 1998-2000 border war between Ethiopia and Eritrea, the 1999 Kargil conflict between India and Pakistan, and the 2003 war in Iraq.
[3] For a detailed overview of minor, intermediate and major armed conflicts 1946-2003, refer to the Uppsala Conflict Data Program: http://www.pcr.uu.se/research/UCDP/ (last accessed: 06-19-2011).
[4] According to Elbadawi et al. (2001) one out of every five people in Sub-Saharan Africa is directly affected by civil wars (cited in Kimenyi and Njuguna 2005: 123).
[5] The approved budget for the UN Department of Peacekeeping Operations for the period from 1st of July 2005 to 30st of June 2006 alone is approximately US$5 billion (cf. United Nations (2006): United Nations Department of Peacekeeping Operations Fact Sheet. Retrieved 27.06.2006 from:).

Because of the weight of these implications, the World Bank devoted the 2011 World Development Report entirely to "Conflict, Security, and Development" (World Bank 2011), underlining the fact that conflicts and civil wars are at the core of current economic and political debates.

Factors that contribute to the outbreak of civil wars have been studied extensively, and competing explanations thereof continue to exist.[6] The debate about the determinants of civil wars has been influenced, among others, by Paul Collier and Anke Hoeffler's various works (2004, 2009).[7] The authors ask whether civil wars can be explained best by 'grievance', such as inequality or a lack of political rights, or by 'greed', which is to be understood as the revenue motivation for rebellion, for example lootable natural resources, minus the costs. Arnson (2005: 12) states that although resources are central to the duration and intensity of war, the roots and objectives of war can be explained best by grievances. The most important factors that have been commonly found to facilitate civil war outbreaks are poverty, ethnic or religious discrimination, a lack of democracy, and opportunities for rebellion (Collier and Hoeffler 2004).

In our paper we put the focus on inequality as the main determinant of civil war onset. There have been certain attempts to explain how inequality affects the probability of starting a civil war, but both the theoretical approaches and the attempts to measure the impact empirically have led to quite contradictory results. The latter have also been hampered by severe problems due to data constraints so far. Earlier theoretical studies have often put the focus on finding the right functional form of the relationship between inequality and civil war onset. This is also important in the empirical context, as a functional form misspecification might lead to biased estimation results. However, the resulting theories were far from leading

[6] Relevant studies include Sambanis 2001, de Soysa 2002, Collier and Hoeffler 2004, Ross 2004, Collier, Paul and Nicholas Sambanis (Eds.) 2005, Ron 2005, Lujala, Gleditsch and Gilmore 2005, Fearon 2005, Montalvo and Reynal-Querol 2005, Blattmann and Miguel 2010 etc.

[7] Especially for material on the literature review and the description of the variables, we thank Sonja Bohnet, born Rabus; see also Baten and Rabus (2007).

into the same direction.[8] Muller (1985), for example, assumes a curvilinear, positively accelerated relationship between inequality and political violence, such that discontent as a result of inequality is leading to political violence. The likelihood of political violence increases exponentially with discontent due to an ever higher mobilization. The author mentions, however, that very high inequality might also lead to a lower probability of civil war if we take into consideration that inequality goes hand in hand with mighty elites that use their power to suppress conflict. Other authors propose a U-shaped relationship between inequality and civil war, assuming that violence occurs either if inequality is very high or very low but is hardly ever observed at intermediate levels (Davis 1954, Havrilesky 1980, Parvin 1973, see the review of Lichbach 1989). The idea behind the U-shaped curve is that the government typically decides to redistribute from rich to poor if inequality is low. Therefore, the rich minority of the society will have a high level of dissatisfaction and might use its financial power to organize political violence. At high levels of inequality, the poorer strata will be dissatisfied and is likely to rebel. Nagel (1974), on the other hand, argues that conflict would rather occur at intermediate levels of inequality. According to his theory, discordance due to inequality depends on the grievance that results from an individual's comparison with others and his individual tendency to compare.[9] Given Nagel's theses that the first factor increases and the latter decreases with inequality[10], he suggests a U-shaped relationship of inequality and civil war onset.[11] Although this theory seems quite plausible and intuitive, Nagel (1974) did not find a non-monotonic relationship between inequality and

---

[8] Lichbach (1989) provides a detailed literature survey on that topic.

[9] According to Nagel, (Discontent due to Inequality) = (Tendency to Compare) * (Grievance Resulting from Comparison); or, symbolically: D=TG. G= a_2g with g=inequality. T= 1- a_1g. The maximum point of discontent would be at $D_{max} = a_2 /4a_1$.

[10] The rationale behind this is that because people rather tend to compare themselves with people from the same strata.

[11] The U-shaped form could also be explained in a more intuitive way. If we assume that grievance due to inequality leads to civil war at a certain point of discontent, the probability of civil war onset increases with inequality. But as unequal societies are often ruled by small but powerful elites with military power, the probability that the government will suppress revolts either through concessions or repression is relatively high. In this case, the probability of civil war onset decreases at a certain point with increasing levels of inequality.

conflict in his data. In sum, there are several different theories about non-monotonic relationships and their contents are quite contradictory. As those theories take plausible but different views, the empirical study might be even more important.

In the few studies that tried to assess the relationship between inequality and civil war empirically (see for example Alesina and Perotti 1996, Cramer 2003, Muller and Seligson 1987, among others), positive and negative empirical findings balance each other out. Some authors admit that the explanatory power of the results is rather limited as there is a severe data problem, especially for the poor and conflict-prone regions in the world. That is caused by the fact that poverty and inequality have only been measured in monetary terms. Unfortunately, only incomplete or imprecise estimates of GDP per capita or Gini indices are available for many regions, in particular those that are notably affected by civil war. Apart from the missing data problem, GDP estimation might be imprecise in countries where many people do not receive wages as they are self-employed, earn their money in the shadow economy or are unemployed. Miguel et al. (2004, p. 728) conclude that as a result of lacking reliable data on income inequality, especially with respect to the African case, it is not possible to assess the importance of these potential determinants. Or as Fearon and Laitin (2003: 85) put it: “The poor quality of the inequality data, available for only 108 countries, does not allow us to go beyond the claim that there appears to be no powerful cross-national relationship between inequality and onset (…).”

Inspired by the lively debates about the causes of conflict, this study intends to contribute to the discussion by analyzing the impact of relative and absolute deprivation on civil war onset. By using our comprehensive dataset we are able to overcome the data problem that has so far hindered many authors to empirically assess this topic. We hypothesize that populations in which people feel deprived compared to others are more likely to start a civil war and also expect a positive relationship between absolute deprivation and civil war onset.

One of the major novelties in this study is the fact that we do not define poverty or inequality in purely monetary terms but take an approach that is close to Amartya Sen's capability-emphasizing approach to human well-being.[13] Martha Nussbaum, who refined the approach by developing a list of the most fundamental capabilities a society should have, understands bodily health - defined as "being able to have good health, including reproductive health; to be adequately nourished; to have adequate shelter" – as one of the central human capabilities (Nussbaum 2003, p. 41). Our proxies for well-being and inequality – average height[14] and the inequality thereof – are determined to a large extent by this capability.[15] Unlike most recently published studies that limit their scope to the time period since World War II or even the post-Cold War period, our method allow us to take into account a much longer period. We assess civil wars from as far back as 1816, amounting to 306 conflicts. We also take into account potential endogeneity issues, a problem that has been ignored in many earlier empirical studies. The rest of the paper is organized as follows. Section II presents our data and introduces our proxy for absolute and relative deprivation as well as the subsequent hypotheses. Section III presents the estimation framework and results. Section IV concludes.

II. Data, methods and hypotheses

A. Data

Because of the extended time period covered, we use an enlarged version of the Correlates of War (COW) project's data[16] on intrastate wars, beginning in 1816.[17] The onset of civil war is coded as a dichotomous variable adopting the value 1 if at least one new civil war started in a

---

[13] Capabilities in Amartya Sen's approach are opportunities that depend on peoples' personal and social circumstances.

[14] Following Baten/Blum (2010), we use estimates that were transformed into male equivalents, since the vast majority of historical height estimates are for males. For further discussion of the topic and advantages of this method, see Baten/Blum (2010), p. 8 pp.

[15] Of course, bodily health is also positively related to monetary wealth as richer people have better access to food and health services.

[16] Cf. Sarkees (2000).

[17] For further information, see Singer, and Small (1972), or: http://www.correlatesofwar.org.

given country during a ten-year period. The COW database only contains information on those conflicts that caused at least 1,000 battle-related deaths (of all participants) per conflict year.[18] Furthermore, a civil war is only included in the data set if one of the conflict parties is the government of a state[19] and both sides are organized armed forces, capable of "effective resistance".[20] As the COW database excludes colonies, we apply the same definition to Clodfelter's compilation of civil wars (Clodfelter 2002) and hence obtain global coverage, as far as historical sources commented on such large-scale civil war events.

Information on adult male height was obtained from Baten and Blum (2012) who collected a very large number of height estimates, comprising all world regions.[21] Height is always organized by birth cohort, because the influence of the environment on an individual is strongest during the first three years after birth.[22] The existing height evidence is aggregated by country and birth decade. By dividing the data into country-specific ten-year periods, we transform it into panel data. The countries and birth decades covered are shown in table A1. One could imagine the potential problem that heights were measured in a selection of countries, for example, countries with a long tradition of measurement and academic study. This could cause sample selection bias due to missing values for the other countries. Fortunately, the data coverage of poor world regions is very comprehensive. Still,

---

[18] The war participant state has to contribute at least 1,000 troops or suffer 100 deaths while the war participant non-state group contributes at least 100 troops or suffers 25 deaths.

[19] This criterion differentiates classical intrastate conflicts from so-called 'New Wars' in which the government does not have to be involved (see Münkler 2003, Chojnaki 2005). Please note that according to the new definitions of the COW project, war types 6 and 7, Regional internal and intercommunal wars are also considered to be intra-state wars, but no civil wars. Civil wars are either Civil wars for central control (type 4) or Civil wars over local issues (type 5).

[20]Effective resistance means that either "(a) both sides had to be initially organized for violent conflict and prepared to resist the attacks of their antagonists, or (b) the weaker side, although initially unprepared, is able to inflict upon the stronger opponents at least five percent of the number of fatalities it sustains." The definition of wars is explained in the Intra-State war Codebook, available at http://www.correlatesofwar.org/COW2%20Data/WarData_NEW/WarList_NEW.html (last access, January 16th, 2013).

[21] Baten and Blum obtained the data on the basis of a very large range of available height estimates, plus the Measure DHS (Demographic and Health Surveys) project data (http://www.measuredhs.com/), where available, plus estimates using the World Health Organization's (WHO) Global Database on Child Growth and Malnutrition (http://www.who.int/gdgm/p-child_pdf/).

[22] The height data points and their limitations are documented in detail there and in the Baten and Blum study that is attached as an appendix.

we take care that selectivity caused by missing values does not bias the results. Sources and detailed explanations of other explanatory variables than the main variables described above can be found in the appendix.

B. Methods and Hypotheses

Economic deprivation has often been mentioned as one of the major determinants of civil conflict, but so far it has always been measured in purchasing power in the literature. Moreover, it has never been possible to assess this factor over two centuries. Instead of using GDP per capita or inequality of income estimates, we introduce adult male height as an indicator of the biological components of well-being, with low heights functioning as a proxy for absolute deprivation.[25] As a second main contribution, we assess relative deprivation measured by the coefficient of height variation. This variable measures how unequal the access to good nutrition and healthcare is for different groups in the society. Our approach to measure absolute and relative deprivation is quite different from common ways to measure these variables. We therefore will describe the idea, its origin and the measurement process in detail below.

Mean adult height is by now a well-established indicator of biological welfare (Fogel 1982; 1986, Steckel 2009, Komlos 1985). Environmental conditions prevailing during a cohort's first three years of life have a strong influence on adult height, because the growth rate in human stature is highest during this age (Eveleth and Tanner 1990, Baten 2000a).[26] Hence, stature measurements of adults that were recorded decades later can be used to shed light on the period of the first years after birth (cf. Moradi and Baten 2005). There is comprehensive anthropometric theory documented in the literature that cannot be reported

[25] For collections of recent examples, see Komlos and Baten (1998), Steckel and Floud (1997), and Komlos and Cuff (1998). On height inequality research, see Soltow (1992) and Quiroga and Coll (2000).

[26] Admittedly, the environmental influence on growth during later ages, especially during the adolescent growth spurt, remains – theoretically speaking – a potential void. However, Baten (2000b) finds in a multi-country empirical study that this effect is negligible compared with the impact of the first three years, as long as individuals have reached their final height.

here in detail (Komlos 1985, Steckel 1995, Baten 2000b; and on height inequality especially Baten 2000a). These studies provide proof that genetic factors matter strongly at the individual level, while population averages are mostly determined by nutrition and health conditions. If a person's parents were tall, she is also tall for genetic reasons, but, at a population level, the Dutch were very short during a period of severe protein malnutrition during the mid-19th century. Many patterns that earlier anthropologists considered to be attributable to genetics (such as tall Masai and Tutsi) turned out to be the results of special nutritional and health environment features.[27] The nutrition and health impact on height is confirmed by studies such as the one by J.P. Habicht et al. (1974), who found that upper class groups of African countries were of similar height as the average U.S. citizen, whereas the middle and lower class in their countries were severely stunted.[28] Graitcer and Gentry (1981) confirm this finding for Egypt, Togo, and Haiti. Correspondingly, educated young Chinese in Beijing today are only marginally shorter than the U.S. standard would predict. There may be exceptions with very isolated populations, such as the Pygmies, for whom genetic height potential may play a limiting role at the population level. In other words, for those populations even optimal diet and health conditions would not produce height levels similar to Western averages. But those exceptional populations have never accounted for a substantial share of the world population during the past two centuries.

Occasionally, purchasing power and the biological components of welfare (life expectancy, health, quality of nutritional intake) diverge from each other. In such situations, heights and real incomes also diverge. For example, it has been observed that heights fell, even though real incomes grew, when the U.S. economy switched from being characterized by farmers living in remote frontier areas to a highly urbanized industrial society. These

[27] In late 20th century, there may be a modest effect of intergenerational height transmission at a population level in wealthy populations due to nutritional customs (for example, in Japan; Baten 2006).
[28] The following discussion is based on Moradi and Baten (2005).

deviations are, of course, interesting by themselves: they allowed important anthropometric contributions to crucially influence debates in economic development, as, for example, the standard of living discussion or the resolution of fundamental research puzzles (such as the "early industrial growth puzzle").[29] We therefore use a data set of adult male height from 165 countries since 1810 that has been compiled only recently (cf. Baten and Blum 2012).[30] The dataset is more closely documented in this study.

Malnutrition has been previously assessed in the literature as a potential cause of conflict, uprisings, and civil wars. Grain price shocks, for example, were often associated with periods of malnutrition. Increasing grain prices as a determinant of conflicts turned out to be particularly important for research on the French Revolution and the conflicts of 1830 and 1848 that occurred after seasons of unusually bad harvests (Labrousse 1948). Marie Antoinette's recommendation to eat cake when bread was lacking was hardly of advantage to the starving population. Berger and Spoerer (2001) found that in the European case "it was economic misery, rather than 'ideas' that caused the outbreak of revolutions in early 1848." More specifically, conflicts only occurred if grain prices strongly exceeded people's price expectations, placing them in a catastrophic nutritional shortage. More recently, Moradi (2005) studied the effect of malnutrition on the probability of civil war in African regions. He found that an increase in calories per day from 1800 to 2500 reduced the risk of civil war outbreak by 11 percent. These examples suggest that internal armed conflicts typically break out one or more years after a famine. This finding suggests that lags in estimating the effect of anthropometric deprivation on civil war occurrence should be taken into account. Following findings in hitherto existing literature, we assume that a civil war is less likely to occur if people are economically well-off and therefore hypothesize:

[29] The 'antebellum puzzle' or the early 'industrial growth puzzle' (see Komlos 1985, 1998, Margo and Steckel 1989) count among them.

[30] The dataset will be available via the Clio Infra-internet site soon.

Hypothesis 1: *The higher the extent of absolute deprivation (measured with the average height of adult people that were born during this decade) in a country during the decade prior to the time of observation the more likely a civil war is to break out in that country.*

As aforementioned, we are using adult male height as a proxy for well-being during a certain decade. To be more precise, we have data on average heights of adults and we know their exact time of birth. On the basis of this information, we estimate the decade in which the individuals were born. If the average height of people (that are adults at the time of measurement) that were born during a certain decade is relatively low (high) compared to individuals born during other decades, we know that there is a high probability that malnutrition and a bad (good) health environment prevailed during the birth decade.[31] Bad living conditions in turn lead to dissatisfaction of the society, which might at a certain point result in the onset of a conflict. It is important to keep in mind that it would not be the individuals from our data who would lead the revolt as these are still children at that point of time. The individuals in our data do merely shed light on the living conditions during the years after their time of birth.

Of course, discontent does not always lead to violent conflict. But *if* conflict results from discontent, it generally takes some time until the dissatisfaction is big enough to result in the onset of civil war. We therefore look at the living conditions in the decade *prior* to the onset of a civil war. Firstly, because we have to take into consideration the time span of the first three years of life in which living conditions are exceptionally important for the later final height of our individuals in the data. Secondly, because we assume that it takes about five years until dissatisfaction amounts to a maximum level of discontent and people are able

[31] For example, we have data on a large number of people who are now adults (and therefore full-grown) and were born in the 1960s and 1970s. If the average height of people that were born during the 1970s is relatively low compared to those born in the 1960s, we can assume that living conditions were worse in the 1970s and that these living conditions hampered the growth of people born during this decade.

to organize in order to start a violent conflict. We are aware of the fact that this might seem to be a slightly imprecise calculation, but we would like to emphasize that every civil war and the time until dissatisfaction results in violent conflict (or not) differs; consequently, we have to make assumptions, which is common in economics. Moreover, the aggregation to certain time periods – unavoidable due to the availability of evidence here – is common in the civil war literature[32] as it might also remove some of the random noise.

Despite our expectation that there is a negative relationship between well-being and civil war onset, it might also be the case that it is not absolute deprivation of an entire society that makes people engage in conflict, but rather relative deprivation. Nafziger and Auvinen (1997) argue that conflict is rather likely to occur if people feel disadvantaged relative to others. Moreover, some parts of the population might feel especially deprived if they observe different consumption patterns of the richer strata even if severe deprivation in absolute terms does not exist (Nafziger and Auvinen 1997: 8). Barron et al. (2004) show that there is a significant and positive relationship between (income) inequality and conflict in Indonesia, especially on the village level. Furthermore, recruiting activities might be most successful in regions with high inequality, as Macours (2011) shows for Nepal between 1996 and 2006. Therefore our research project will make a valuable contribution to civil war literature by examining the relationship between inequality and civil war in depth.

We use the coefficient of height variation (CV from here) to estimate height inequality – and thereby the inequality of living standards – of a population within a certain birth decade. Baten (2000a) showed that the coefficient of height variation is a good proxy for overall inequality within societies, as the two measures are correlated (see also Moradi and Baten 2005, Van Zanden et al. 2013). However, human height is influenced by factors that are not traded, but rather provided by public goods, such as the healthcare system within the

[32] Most of the other studies use 5-year-periods, see for example Collier/Hoeffler (2004), Collier/Hoeffler/Rohner (2009).

society. Therefore, the correlation between income inequality and height inequality is not a perfect one. But we argue that anthropometric inequality measures might be even advantageous compared to income-based inequality indicators in civil war studies as height captures important aspects of living standards, in particular nutrition and health (Komlos 1985, Steckel 1995). More specifically, although income inequality is an important factor for dissatisfaction, people might be even more inclined to rebel if they feel deprived of very basic factors of well-being while other groups of the society have access to these. We therefore hypothesize:

Hypothesis 2: *Higher inequality within a country, measured by the coefficient of height variation, leads to a higher probability of civil war onset.*

We calculate the coefficient of height variation on the basis of our height data. For a country i and birth decade t, the CV is defined as:

$$CV_{it} = \frac{\sigma_{it}}{\mu_{it}} \cdot 100$$

A high coefficient of height variation therefore implies that a society is highly unequal (see also Baten and Blum 2011 for more details).

III. Estimation Framework and Results

We analyze the probability of a civil war onset within a ten-year period in a certain country from the 1810s to 1990s.[33] Our basic model with which we test our hypotheses will be a logit regression of the form

$$\text{Civil War}_{it} = \beta_0 + \beta_1\, height_{i,t-1} + \beta_2\, inequality_{i,t-1} + BX + \eta_i + \gamma_t + \varepsilon_{it}$$

where t refers to a decade, and i refers to the country in question. X is the vector of control variables, while γ_t denotes time fixed effects that have been added to control for

[33] Other studies prefer using five-year-periods (Collier and Hoeffler 2004, Collier et al. 2009, among others) which is not possible with our data. However, as we argued before, discontent mounts over time so that certain years are needed until the discontent results in civil war.

unobserved factors. Geographic fixed effects η_i are controlled for with world region dummies or country fixed effects estimation. We use clustered standard errors to avoid potential consequences of serial correlation.[35]

Civil War is an indicator that equals one if a new civil war started in country i during decade t and equals zero otherwise, so we only look at the determinants of war onsets, not ongoing wars. We are lagging height and inequality levels in all but the instrumental variables specifications to avoid contemporaneous correlation, because heights as well as inequality might be influenced by civil wars. Civil wars are likely to increase absolute deprivation by destroying infrastructure and harvests and therefore complicating access to food and healthcare. Inequality might rise for the same reasons mentioned above, especially because already poor people will have less access to scarce resources. Moreover, one could imagine that the poorer strata of the society that have neither political nor economic power will be disproportionally burdened with the costs of reconstruction and other costs that incurred during and after the war. On the other hand, it is also possible that inequality is reduced if the assets of richer people are heavily affected by civil war and the society therefore becomes relatively more equal.

Apart from our key variables that are proxies for absolute and relative deprivation, there are several other variables that might be potential determinants of civil war and that were included in our models. In the following, we briefly present the theoretical consideration behind the explanatory variables of our basic model and discuss their expected signs. The sources and exact definitions of these variables can be found in the Appendix.

Democracy illustrates the extent of political freedom of a country. Intuitively, we would expect that a democratically organized country would have less risk of facing a civil war as democracies typically have at their disposal strong institutions that tend to resolve

[35] Our estimates are in fact heteroskedasticity-robust and cluster-robust, see Cameron/Trivedi (2009), p. 85.

struggles for power and conflicts through political arrangements (Skaperdas 2008). Moreover, dissatisfaction within the population is most likely in repressive regimes. On the other hand, one could imagine that repressive regimes try to impede the outbreak of civil wars through repression or concession to remain in power.[36] This would support Hegre et al.'s (2001) finding of a nonlinear relationship between political regime and civil war: According to the authors, pure democracies and pure dictatorships are less conflict-prone than intermediate regimes. We test the non-monotonic hypothesis by introducing the squared term of this variable in one of our models. We also test a specification with a percentage change in democracy from the previous to the current period to control for policy changes that have been driven by the threat of political unrest.

A large strand of literature is engaged in disentangling the effects of the *ethnic composition* of a country's population on civil war outbreak. Sambanis (2001) investigated whether ethnically and nonethnically-based civil wars have the same causes. In his quantitative analysis, he finds that this question must be answered with 'no': He discovers the main determinants of ethnic conflicts to be rooted in ethnic heterogeneity of a state's population, the lack of democracy in the country under observation as well as in a neighboring state, and the occurrence of civil wars in neighboring countries. On the other hand, Fearon and Laitin (2003) find that it is not the countries' ethnic or religious characteristics that best explain the risk of a civil war outbreak, but rather the conditions that favor insurgency, such as poverty, political instability, or large populations. We use Alesina et al.'s (2003) ethnic fractionalization variable that takes into account linguistic as well as ethnic characteristics and reflects the probability that two randomly selected individuals belong to different groups. We also test specifications with the (only) linguistic- and religious fractionalization indices.

[36] Aidt and Jensen (2011) as well as Acemoglu and Robinson (2000) showed that European governments increased voting rights from the 19^{th} to 20^{th} century when they were faced with the threat of a conflict.

In addition, we include *population size* as a control variable. The rationale for doing so is that a larger population makes it more likely that the critical threshold of 1,000 battle-related deaths will be reached. For example, China experienced many civil wars simply because its population was so large. We also consider the country's *colonial background*. Colonial empires can influence internal armed conflicts either because the colonial armies prevent conflict by force or because they provoke conflict via unfair treatment of the colonized. We can guess that colonial empires prevented major civil wars of such a destructive magnitude within their colonies, as they were extremely harmful to the colony's economy and prevented risk-averse investors from engaging actively in the colony. The colonial powers were superior in military technology. Therefore, costs remained within bounds. On the other hand, colonies were often created ignoring ethnic or tribal lines, which might explain ethnic conflicts. Therefore, it is not immediately obvious in which ways a colonial background might influence the propensity for starting a civil war.

Finally, we test whether the idea holds true that rebels finance their activities based on natural resources, and therefore introduce a "greed" component as proposed by Collier and Hoeffler (2004). *Diamonds* are an example of a natural resource that is relatively easy to loot.[39] Ross (2006) shows that the number of civil wars increased sharply in diamond-producing countries from the 1970s to the 1990s and he finds a positive correlation between diamonds and civil wars. Smillie (2002) supports this finding in his study on Sierra Leone. However, Ross cautions the reader about generalizations, since civil wars in diamond-producing states are still quite rare, even though they increase the number of civil wars. Considering the argument that rebels finance themselves through natural resource revenues, we conclude that a civil war is more likely to occur if a country has diamond deposits.

---

[39] This is the case, at least, for secondary diamonds. Including a dummy for diamonds therefore seems more reasonable than including oil as an example of a natural resource. A potential shortcoming of this paper is that we do not differentiate between primary (difficult to loot) and secondary (easy to loot) diamonds.

Descriptive Statistics

How did the general trend of civil war occurrence develop over the last two centuries? In general, the number of conflicts was relatively constant until WWII, with approximately 10-15 major civil wars per decade exceeding the threshold of 1,000 battle-related deaths. After WWII, there was a clear increase in the number of civil wars, followed by a boom in the most recent period (over the past decade, the number has slightly declined again; see Figure 2).

[Figure 2 about here]

If we compare this development with the evolution of height over the past two centuries, we actually observe a modest correlation – one that is opposite to the expected direction. Heights were relatively constant in the first period (until approximately 1860-1890) and started to increase in the 20th century, most notably in Western and Eastern Europe and European settlements. In Africa and East Asia, a certain increase in height between the 1890s and 1950s can also be observed (Figure 3). Hence, if we observe some influence of height on the probability of civil war outbreak, it is most likely not due to a trend correlation – the trend developments of the two series would suggest a correlation in the opposite direction. Notwithstanding these observations, Africa experienced a period of decreasing heights afterwards (Moradi 2005).

[Figure 3 about here]

Looking at inequality trends by world region, no overall trend is visible (Figure 3.2). The levels of inequality were generally quite high in Latin America, at least since the late 19th century, which is also confirmed in other studies (Van Zanden et al. 2011). In Europe (both East and West) and Asia inequality levels were generally lower. In Western Europe's inequality levels decreased between the 19th and 20th centuries. Sub-Saharan African inequality in contrast increased substantially between the first and the second half of the 20th century, consistent with what we know about civil war trends.

An examination of the descriptive statistics already offers insight into possible outcomes of our regression analyses. We observe, for instance, that the mean population (in logs) in countries that have experienced civil wars (9.31) is substantially larger than in such countries that did not (7.74), thus supporting the hypothesis that more populous countries are at higher risk of a large civil war (1,000 battle deaths/conflict year).

[Table 1 about here]

Furthermore, civil war-ridden countries are less democratic (-2.57) than their peaceful counterparts (-0.21). Lagged height was in fact lower in the observations with incidence of civil war, and ethnic fractionalization was roughly equal in the civil-war ridden countries. Most important for the present study, we do observe higher inequality within countries that experienced a civil war. The coefficient of height variation equals 3.93 in countries that experience a civil war onset compared to 3.75 in countries without a new civil war. In the upper panel, we document all of the cases, whereas the bottom panel shows the descriptives for the cases of table 2 in which none of the explanatory variables are missing. The correspondence between both panels is high. We will assess missing value selectivity in greater detail below. We also took a closer look at the functional forms of inequality, height and democracy. Theories of non-monotonic relationships between inequality and civil war cannot be confirmed for our data. Figure 1 shows that civil war probability is rising almost in a linear way, at the most a slightly exponential function can be assumed.

[Figure 1 about here]

When we tested the relationship by controlling for inequality and its squared term in a regression analysis, we did not find a significant non-monotonic relationship either.[46] We therefore abandoned the idea that inequality might affect civil war probability in a non-monotonic way and do not use other functional forms of inequality in our regressions. The

[46] Regressions not reported here, but available from the authors.

probability of civil war onset subject to democracy is decreasing linearly with higher levels of democracy (Figure A1). However, regression analysis suggests that there might also be a non-monotonic relationship between the political regime and civil war probability. To avoid functional form misspecifications, we always use democracy and its squared term in the regression analyses.

As units are probably statistically related over time, we have to account for potential serial correlation problems. Beck et al. (1998) propose to use the discrete logistic model of the form

$P(y_{i,t} = 1|x_{i,t}) = h(t|x_{i,t}) = \frac{1}{1 + e^{-(x_{i,t}ß + \kappa_{t-t0})}}$ where κ_{t-t0} denotes a dummy variable measuring the time of peace until time t. This model has been applied, among others, by Aidt and Jensen (2009). However, temporal dummies should *not* be applied if the null hypothesis of temporal independence is not rejected as unnecessary multicollinearity would be the consequence (Beck et al. (1998), p. 1269). This is the case in our data (p=0.4312), possibly not because there is no path dependence at all, but rather because we use decades instead of yearly observations. Although the above mentioned model does not seem to fit for our data, we are taking potential serial correlation seriously. We therefore cluster standard errors at the country level in every specification and employ variables that control for war persistency.[47] War persistency is either measured by the absolute number of peace decades, which is adopted from the idea of introducing spell counters above, just using the absolute values instead of dummy variables.[48] As an alternative measure we introduce the variable “warlast” that measures how many years of conflict took place in country i during the previous decade. While the former variable takes into account the entire period of time, the latter shifts the

[47] This is the standard solution to this problem in civil war literature, see for example Miguel et al. (2004).
[48] The underlying assumption would be that the probability of civil war onset decreases linearly with peace duration, which might be a rather strong assumption. However results do not change if we only take into account the previous decade.

focus to the short run, relying on the assumption that it might be rather violent action in the recent past that leads to the outbreak of new conflicts. We ran the regressions with both specifications and received the same outcome for our main variables, so we generally keep to the long-term perspective of war persistency and use the short-term variable only in some of the specifications to prove the robustness of the results.

We use different sets of variables and models to test for the robustness of our results. We expect that inequality increases the likelihood of civil war, whereas height is negatively related with civil war onset.

[Table 2 about here]

In models 1 to 4 of table 2, we use the same set of variables with pooled logit, panel logit and rare events logit models. The rare events model takes into account that civil war onsets are rather rare compared to periods without civil war onsets and generates approximately unbiased and lower-variance estimates of logit coefficients and their variance-covariance matrix.[53] We find that inequality in fact increased the probability of civil war. Similarly, height – as a measure of non-deprivation – reduced this probability. Hence, both hypotheses were confirmed and statistically significant throughout the models. The relative importance of our inequality variable is not trivial. A one standard deviation increase in inequality leads to a 0.23 increase in the log odds of having a civil war (Table 2, Model 2). With full standardization, the results show that an increase in inequality by one standard deviation results in 0.1 (if we use Model 1 of Table 2) to more than half a standard deviation increase (0.67 in Model 2 of the IV Table 5) in civil war onset, the latter being certainly a substantial economic effect. The higher impact in the IV models is probably a result of the reduced measurement error compared to previous models without instrumental variables.

[53] It was not possible to report marginal effects for the rare events logit model with our statistical software. But it is possible to calculate the probability of civil war onset manually: If we set all explanatory variables at their mean and increase inequality from the 20th to 80th percentile, civil war risk increases by 0.0305, or by 0.164 if we increase inequality from the mean to the maximum level.

Nunn (2008, pp. 159-163) emphasized that the second function of instrumental variable techniques - to estimate with less measurement error - results sometimes in larger coefficients of the second stage, compared to OLS estimates. If the instrumented variable is measured with a certain amount of error, while the instruments are not, it is not astonishing if the coefficients are larger. This is certainly the case for our variable "inequality", because all inequality evidence is known to suffer from some degree of measurement error. Hence, the seriousness of this caveat is reduced with the instrumental variable techniques as well.

Apart from our two main variables, there are certain other variables that seem to have an important impact on the onset of a civil war. Countries under colonial rule were at higher risk of facing civil strife. The result is significant in every specification in which the variable is included. As expected, a larger population is also associated with higher civil war risk, although the coefficients become insignificant in some of our specifications. In some cases, the relationship between democracy and civil war onset seems to be bell-shaped. However, the turning point would be beyond all data points we observe in our panel, therefore our dataset only covers the observations left to the turning point.[54] We tentatively conclude that civil war onset is less likely in countries with very high levels demoracy, as squared democracy has a negative sign. But this squared term of the political system variable is only significant in some of the specifications.

Our long-term measure of peace duration indicates that countries that have not faced civil war for longer time periods are less prone to face conflicts in the future. In contrast, the *warlast* variable suggests that the number of months in the previous decade does not play a role: Countries that were susceptible to war during the previous decade do not face higher risk of having a new war in the near future.

[54] Please note that democracy and democracy squared have been divided by 100 for expository purpose. The turning point would be at approximately 0.26 which is above the maximum of 0.10 (transformed, original value 10).

In model 5, we test whether a policy change since the previous period affected the likelihood of a civil war onset. We obtain a negative coefficient, so a shift towards more democracy might potentially reduce the probability of a civil war onset, but the coefficient is not significant.

Model 6 introduces the growth rate of height to assess whether changes in well-being rather than absolute or relative well-being are responsible for dissatisfaction. For example, people might feel more deprived if the overall well-being in previous periods was better compared with the current situation. Changing conditions in well-being proxied by height growth does not seem to generate sufficient dissatisfaction for conflict (Model 6).

Ethnic fractionalization and the existence of diamonds are added in models 7 and 8. None of these variables turns significant. In most of the models, we include time and world region effects, but omit them in some of the models to see whether our results stay robust. In sum, inequality and absolute deprivation remain significant in every model.

Table 2a includes solely our deprivation variables and different fractionalization measures. As fractionalization has been mentioned as one of the main drivers of civil wars in the previous literature on civil wars, we introduce ethnic fractionalization, ethno-linguistic fractionalization and religious fractionalization separately to assess whether perhaps fractionalization rather than deprivation is the actual driver of a conflict, or whether the deprivation coefficients change dramatically if we include fractionalization measures. As a result, fractionalization does not turn significant in any of the specifications in our long-run study. As most studies that focus on the second half of the 20th century find an impact, our results contribute important findings for colonial effects. Namely, ethnic factors were apparently no major driving forces in pre-colonial times and also played no major role during the colonial period. Dissatisfaction was suppressed or sometimes channeled to activity against the colonial powers. Only during the post-colonial period, with artificially tailored borders, ethnic fractionalization increased civil war probability, which has been found by many other studies on this topic which include only the post-colonial period.[55]

As a robustness test, we also ran linear probability regressions (Table 3). Our main variable inequality remains positive and significant throughout the different models, while absolute deprivation only becomes significant in the random and fixed effects models.[56]

In table 4, we test whether the results remain stable across different income levels by omitting either very rich or very poor countries.[57] Inequality turns insignificant if we exclude least developed countries. If we exclude rich countries, inequality is significant and positively related with the onset of civil wars and the coefficients are much bigger than in our baseline specifications. We therefore conclude that the impact of inequality on a civil war onset is much higher in poor countries. Absolute deprivation plays a role across all income levels.

[Table 4 about here]

When looking at the different models, differences in the number of observations between the different models can be observed, which is mainly due to missing values for some of the explanatory variables. We want to determine whether these missing values might have resulted in a missing value sample selection problem. In other words, do we have a very special subsample due to missing values in some of the specifications? We tested this possibility by assessing whether countries of a certain income range have been over- or underrepresented in one of the models. We also checked whether the distribution of observations within certain income groups changed between the models (Figure 4). We found that our models with few explanatory variables draw an almost perfect mirror image of the

---

[55] Alesina et al.'s (2003) ethnic fractionalization contains not only language- but also other ethnic characteristics, we preferred to use this measure of fractionalization in the following regressions.

[56] Note however, that we do not regard the fixed effects specification as the best model for our data. Statistical literature argues that using fixed effects might lead to inefficient results if the between-variation is better measured than the within variation, because it puts much less focus on the between variation (see Cameron and Trivedi 2009, p.244, among others). Moreover, if we use fixed effects we have to assume that unobserved heterogeneity is consistent over time, which is a very strong assumption for such a long time period. Therefore, we only include some fixed effects specifications as robustness tests.

[57] To exclude poor countries, we drop countries with a GDP p.c. less than 1,000 US $. We define rich countries as those that have a GDP with more than two standard deviations above the mean GDP of our sample.

real income distributions in the world (Table 4, Model 10 compared to calculations of the World Bank).

[Figure 4 about here]

In the other models, high income countries are modestly overrepresented, while lower middle income countries are modestly underrepresented. However, income levels do not differ vastly from the real distribution in any of the models. Low income countries are always sufficiently covered. It is also important to note that particularly war struck regions such as Africa and Latin America are quite well covered in our data.[58] Therefore, we conclude that missing value sample selection that might lead to omitting the poorest countries is not a problem here.

Instrumental Variable Models

In what follows, we will dwell on endogeneity issues. The results of the ordinary least squares regressions could be affected by reverse causality. Civil war may lead to a decrease in children's height (see, for example, Bundervoet et al. 2009) and perhaps also a change in height dispersion. Instrumental variable (IV) estimation allows us to circumvent these issues of endogeneity. Which instrumental variables could be suited for this purpose? Easterly (2007) and Galor et al. (2009) have recently advocated the use of climatic, geological and similar variables, which proxy types of agriculture that correlate either with higher or lower efficient sizes of scale. Our instrumental variables for inequality will be: (1) the ratio of soil and climatic suitability for sugar cane relative to wheat and (2) low population density in 1500 interacted with southern location. Sugar plantation and cattle-raising for hide and beef exports are typical examples for economies of scale in agricultural production: The owner of a sugar plantation in Cuba or Brazil will aim at producing on a large scale by employing a high number of slaves or – later – free laborers on large plantations, because higher scale

[58] Table available from the authors.

increases revenues much more than costs. Likewise, an Argentinean or Australian cattle baron will increase his herd and the amount of pasture to enormous values. As cattle-related variables might be endogenous, we instrument this with the second IV. Relative to sugar cane, wheat production is already highly productive on much smaller farm units as has been amply demonstrated in the agricultural economics literature. Sugar and wheat production requires relatively clear-cut climatic and soil characteristics. Hence, we can use those to proxy the suitability for the “inequality crop” sugar and the “equality crop” wheat. In the spirit of Easterly’s (2007) famous instrumental variable, we take the ratio between the two.

The cattle-raising activity associated with inequality is unfortunately less concentrated on specific soils. But there is a clear correlation with Southern location and initially low population density. Consequently, we use the interaction term of these two variables as additional instrumental variable. Argentina became the prototype of this type of land use, because the indigenous Indian population on its great plains was always very thin, and in contrast to the United States, European immigration was relatively limited in numbers until the late 19th century. Apart from Argentina and Australia, also South Africa, New Zealand and Uruguay had very unequal distributions following the colonial land grab (Eastwood et al. 2010; Juif and Baten 2013). Thus, we instrument this kind of inequality agriculture with Southern latitude interacted with population density in 1500.

The advantage of the ratio between the climatic and geological suitability ratio of sugar and wheat is its intrinsically exogenous nature, whereas the actual crop use could be influenced by civil wars. Similarly, population density around 1500 is a very popular instrumental variable, because it captures human development in a very early time period (it was prominently used by Acemoglu, Johnson, and Robinson 2002, for example). The autocorrelation of early population densities and those of the 20th century is quite limited due to the unequal population increase in some world regions and the massive migration movements. Similarly, it is likely that Southern latitude is exogenous. The second IV in only included in some of the specifications, as our first stage results show that the IV's correlation with inequality might not be strong enough. We find that the results do not depend on it.

For height levels, we use lactose tolerance as an intrumental variable. Baten and Blum (2012) have summarized the literature on diminished protein consumption as a result of lactose intolerance and find clear empirical evidence for this relationship in their econometric analysis. As we use lactose tolerance instead of lactose intolerance as an instrument, we expect positive signs for this IV.

In addition to an IV probit model, we test Two-Stage Least Squares (from now on: 2SLS) models. Angrist and Pischke (2009, p. 197 f.) argue that using OLS models might also be preferable in the case of limited dependent variables, as they have conceptual robustness advanteges. Wooldridge (2002, p. 594 and 596) confirms that the 2SLS estimates of a linear model provide good approximates for the average partial effects of bivariate probit model. We use Limited Information Maximum Likelihood (LIML) models in cases where our F-statistics were relatively low.

The results of the two-stage-least-squares regressions confirm that we have strong instruments for inequality and height.[60] The respective instruments for height and inequality are strongly correlated with our endogenous variables, as shown in the first stage of table 5. The Angrist-Pischke multivariate F-test is above or near the critical threshold of 10 (Stock and Yogo 2002) in case of the 2SLS specifications (models 2 and 3). The F-value for the inequality instruments is below the critical threshold in the LIML specifications, but the Anderson-Rubin test, which is robust to the presence of weak instruments, rejects the null hypothesis that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero in every model. We conclude that we do not have a weak instruments

[60] As we have two endogenous variables, we use the Angrist-Pischke F-statistics, as the use of the overall F-statistic might dissemble a well-identified first stage, although only one instrument is the reason for a high F.

problem here. The Kleibergen-Paap statistic clearly shows that our reduced form model is not underidentified.

We would argue that the instruments influence the dependent variable only through the potentially endogenous variable, inequality (see the discussion on the exclusion restriction below). The p-value of the Durbin-Wu-Hausman test, which could be calculated for models 2 (p-value: 0.037) and 3 (p-value: 0.178), indicates that the IV approach is not strictly necessary in every case, as the results are not significantly different from OLS estimates, at least if control for democracy and its square. Any endogeneity among the explanatory variables would not have deleterious effects on OLS estimates. Nevertheless, given the general suspicion that all macroeconomic variables could be endogenous, it is reassuring that the significant impact of inequality remains a consistent determinant of a civil war onset. In contrast, height levels are not significant once instrumental variables are used. This might also be caused by the lower number of cases available for this estimation – the sizes of the coefficients are not small. Most of the other controls are insignificant.

The exclusion restriction

One of the biggest challenges in any instrumental variable approach is the requirement of the exclusion restriction, which implies that the instrumental variables do not have a direct influence on the ultimate dependent variable except via the potentially endogenous variable. In his seminal paper, Easterly (2007) studied the applicability of the exclusion restriction of relative soil and climatic suitability by using both theoretical reasoning and econometric tests. One possibility for such a direct causal channel is the possibility that wheat and sugar have different effects on the wealth of the local population. This wealth difference could be a potential direct causal influence on dissatisfaction and conflict. On the other hand, Easterly argues convincingly that the difference in the wealth effects of those agricultural goods are quite limited compared to all of the other goods that countries are producing.

Another potential violation of the exclusion restriction could stem from the widely discussed concept of the “natural resource curse”. Exceptionally high incomes from raw material exports might generate rents that in turn could lead to political economy problems (Sachs und Warner 1995, Auty 1993, see also the recent review by Frankel 2010). Sugar cane is a primary product that might produce such high windfall profits, for example. Isham et al. (2005) have developed a theory of “point-source” agricultural exports. Typical cases are exports such as sugar cane. The idea is that the “point-source” export revenues can more easily be captured by ruling elites than “diffuse” exports such as wheat. Easterly (2007) argues that if these “resource curse” effects operate via inequality, the exclusion restriction is of course not violated. Most of the studies discussing these issues emphasize that the behavior of rich elites and their interactions with the institutional environment is the main issue, which is consistent with the inequality story (Easterly 2007, Isham et al. 2005).

Nevertheless, one can still imagine that the resource curse works through other channels. One strategy to address these issues is to directly include additional controls for a resource-oriented export structure and determine whether inequality, measured with the sugar/wheat suitability variable, turns insignificant. We again follow the literature in applying this strategy (Easterly 2007; Juif and Baten 2013). We construct a variable of the share of raw material and mining exports relative to the country’s total exports. The “resource curse” variable is insignificant and does not affect the significance of inequality measured by the sugar-wheat-suitability variable (Table 6).

What about the theoretical properties of the instrumental variable “low population density in 1500 interacted with southern location”? Low population density could, for example, have a negative impact on civil wars, because the probability of a civil war onset decreases in sparsely populated countries. However, the population density around 1500 does not have much in common with the population density in the 19^{th} and 20^{th} centuries. To be on the save side we only included this IV in model 5 of table 5, and the ineqaulity coefficient does not differ much from the others.

In the case of the last instrument, lactose tolerance, the discussion of the exclusion restriction is relatively straightforward, as lactose tolerance is difficult to imagine having a separate impact on a civil war onset, except via the anthropometric causal channel. We conclude that the issue of the exclusion restriction does not pose a major problem.

IV. Conclusions

Several recent publications have documented the negative effect of economic deprivation on civil war risk for the last few decades. In most cases, GDP per capita or its growth rate has been used to indicate an economy's level of development. When assessing the 306 civil wars that took place between 1816 and 1999, we find that anthropometric measures are powerful indicators for explaining the onset of civil war. The supposed negative correlation of absolute welfare levels and the probability of civil war outbreak was confirmed in most of our models, but did not pass the instrumental variables tests. Relative deprivation measured by the height inequality, on the other hand, turned out positive and significant in all of the models and passed robustness tests. Inequality therefore is a consistent determinant. Other variables, such as colonial rule or population size, had some influence on the probability of a civil war. We therefore conclude that the discontent of the colonized, caused by the creation of borders across tribal lines and bad treatment by the colonizers, was stronger than the foreign military forces that could suppress conflicts.

We also explicitly addressed the issue of endogeneity of absolute and relative deprivation by introducing instrumental variables. We used lactose tolerance as well as climatic and geological factors that are correlated with inequalities. Those turned out to be strong instruments. The IV regressions confirmed the robustness of our results regarding the impact of inequality on civil war onset. The inequality coefficients remained negative and significant throughout every specification, passing robustness tests as well as instrumental variables estimation.

This study took a long-term view, analyzing civil wars over the last two centuries. This insight was only possible by the use of innovative methods of measuring absolute and relative deprivation. Relative deprivation has hardly ever been introduced to empirical civil war literature as there is a lack of data on inequality in several countries. Our study succeeded to overcome the problem of data scarcity thanks to an extensive data base for inequality in well-being. The impact of inequality on civil war probability can now be systematically tested and a significant impact is found in our study.

References

Acemoglu, Daron, Simon Johnson and James A. Robinson (2002): Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. The Quarterly Journal of Economics 117 (4): 1231-1294.

Acemoglu, Daron, and James A. Robinson (2000): Why did the West extend the Franchise? Democracy, Inequality and Growth in Historical Perspective. The Quarterly Journal of Economics 125(4):1167-1199.

Aidt, Toke S., and Peter S. Jensen (2011): Workers of the World, Unite! Franchise Extensions and the Threath of Revolution in Europe, 1820-1938. CESifo Working Paper Series 3217.

Aidt, Toke S. and Peter Jensen (2009): The Taxman Tools Up: An Event History Study of the Introduction of the Personal Income Tax, 1815-1938. Journal of Public Economics 93: 160-175.

Alesina, Alberto, and Roberto Perotti (1996): Income distribution, political instability, and investment. European Economic Review 40: 1203-1228.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg (2003): Fractionalization. Journal of Economic Growth 8: 155-194.

Angrist, Joshua D., and Jörn-Steffen Pischke (2009): Mostly harmless econometrics. An empiricist’s companion. Princeton and Oxford: Princeton University Press.

Arnson, Cynthia J. (2005): The Political Economy of War: Situating the Debate. In: Cynthia J. Arnson and I.William Zwartman (eds.): Rethinking the Economics of War. The Intersection of Need, Creed, and Greed. Washington, D.C. and Baltimore: Woodrow Wilson Center Press.

Auty, Richard M.: *Sustaining Development in Mineral Economies: The Resource Curse Thesis*. London: Routledge, 1993.

Baten, Joerg (2000a): Economic Development and the Distribution of Nutritional Resources in Bavaria, 1797-1839. Journal of Income Distribution 9: 89-106.

Baten, Joerg (2000b): Height and Real Wages: An International Comparison. Jahrbuch fuer Wirtschaftsgeschichte. 2000-1: 17-32.

Baten, Jörg and Sonja Rabus (2007): The Political Economy of Civil War Risk: Was there an Impact of Economic Deprivation, 1816 - 2004?" Working Paper.

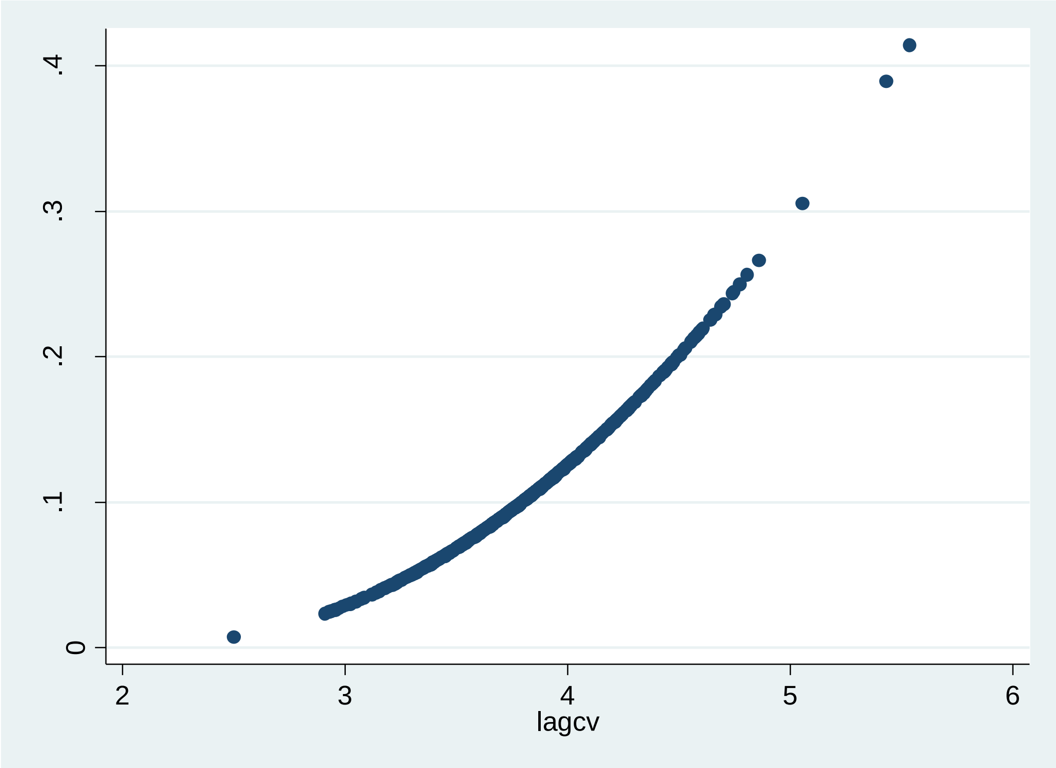
Baten, Joerg (2006): Global Height Trends in Industrial and Developing Countries, 1810-1984: An Overview. WP Tuebingen.
Baten, Joerg and Matthias Blum (2011): Anthropometric within-country Inequality and the Estimation of Skill Premia with Anthropometric Indicators. *Review of Economics -- Jahrbuch fuer Wirtschaftswissenschaften* 62 (2): 107-138.
Baten, Joerg and Matthias Blum (2012): New Findings and New Background Evidence on Anthropometric Welfare in 156 Countries, 1810-1989. Economic History of Developing Regions 27(1)
Barron, Patrick, Kai Kaiser, and Menno Pradhan (2004): Local Conflict in Indonesia. Measuring Incidence and Identifying Patterns. World Bank Policy Research Working Paper 3384
Beck, Nathaniel, Jonthan N. Katz, and Richard Tucker (1998): Taking Time Seriously: Time Series-Cross-Section Analysis with a Binary Dependent Variable. American Journal of Political Science 42 (4): 1260-1288.
Berger, Helge and Mark Spoerer (2001): Economic Crises and the European Revolutions of 1848. Journal of Economic History 61: 293-326.
Blattman, Christopher and Edward Miguel (2010): Civil War. Journal of Economic Literature 48 (1): 3-57.
Brown, Michael E. (1996): Introduction. In: Michael E. Brown (ed.): The International Dimensions of Internal Conflict. Cambridge, MA and London: The MIT Press: 1-31.
Brown, Michael E. (2003): Introduction. Security Challenges in the Twenty-first Century. In: Michael E. Brown (ed.): Grave New World. Security Challenges in the 21st Century. Washington, D.C.: Georgetown University Press: 1-13.
Bundervoet, Tom, Philip Verwimp and Richard Akresh (2009): Health and Civil War in Rural Burundi. The Journal of Human Resources 44(2): 536-563.
Cameron, A. Colin and Pravin K. Trivedi (2009): Microeconometrics Using Stata. Texas: Stata Press.
Chojnacki, Sven (2005): Gewaltakteure und Gewaltmärkte: Wandel der Kriegsformen?. In: Frech, Siegfried, Trummer, Peter I. (Hrsg.), Neue Kriege. Akteure, Gewaltmärkte, Ökonomie. Schwalbach/Ts., Wochenschau Verlag: 73-99.
Clodfelter, Michael (2002): Welfare and Armed Conflicts. A Statistical Reference to Causality and Other Figures, 1500-2000. Jefferson, North Carolina, and London: McFarland&Company.
Collier, Paul, Anke Hoeffler and Dominic Rohner (2009): Beyond Greed and Grievance: Feasibility and Civil War. Oxford Economic Papers 61: 1-27.
Collier, Paul and Anke Hoeffler (2004): Greed and grievance in civil war. Oxford Economic Papers 56(4): 563-595.
Collier, Paul and Nicholas Sambanis (Eds. 2005): Understanding Civil War. Volume 1: Africa. Washington DC: The World Bank.
Cramer, Christopher (2003): Does inequality cause conflict? Journal of International Development 15: 397-412.
Davis. (1954). *Political Statistics*. Evanston,IL: Principia Press.
De Soysa, Indra (2002): Paradise Is a Bazaar? Greed, Creed, and Governance in Civil War, 1989-99. Journal of Peace Research 39(4): 395-416.
Easterly, William (2007): Inequality does cause underdevelopment: Insights from a new instrument. Journal of Development Economics 84: 755-776.
Eastwood, R. Lipton, M., and Newell, A. (2010). Farm Size, in Handbook of Agricultural Economics, edited by R. Evenson and P. Pingali, 3323-3397. Amsterdam: Elsevier.
Elbadawi, Ibrahim A.,N. S. Ndung'u, and S. Njuguna (2001): The Economics of Civil Wars and Post-Conflict Recovery. Working Paper, Development Research Group, Africa

Region,World Bank,Washington D.C.
Eriksson, Mikael und Peter Wallensteen (2004): Armed Conflict, 1989-2003. Journal of Peace Research 41(5): 625-636.
Eveleth, Phyllis B. and and James M. Tanner (1990): Worldwide Variation in Human Growth. Cambridge: Cambridge University Press.
Fearon, James D. (2005): Primary Commodity Exports and Civil War. Journal of Conflict Resolution 49(4): 483-507.
Fearon, James D. and David D. Laitin (2003): Ethnicity, Insurgency, and Civil War. American Political Science Review 97(1): 75-90.
Fogel, Robert W., Stanley L. Engerman and James Trussell (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: 401-421.
Frankel, J. (2010). The Natural Resource Curse: A Survey. Harvard Kennedy School Working Paper 10-005.
Galor, Oded, Omer Moav and Dietrich Vollrath (2009): Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. Review of Economic Studies, 76(1): 143-179.
Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand (2002): Armed Conflict 1946–2001: A New Dataset. Journal of Peace Research 39(5): 615–637.
Graitcer, Philip L. and Eileen M. Gentry (1981): Measuring children: One reference for all. The Lancet 2: 297–299.
Habicht, J. P., R. Martorell, C. Yarbrough, R. M. Malina and R. E. Klein (1974): Height and weight standards for preschool children: How relevant are ethnic differences? The Lancet 6: 611–615.
Harbom, Lotta; Stina Högbladh and Peter Wallensteen (2006): Armed Conflict and Peace Agreements. Journal of Peace Research 43(5): 617-631.
Havrilesky, Thomas (1980): The Discordance-Inequality Tradeoff. Public Choice, 35(3): 371-377.
Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Peter Gleditsch (2001): Toward a democratic civil peace? Democracy, political change, and civil war, 1816-1992. American Political Science Review 95 (1): 33-48.
Isham, J., Woolcock, M., Pritchett, L., and Busby, G. (2005). The varieties of resource experience: natural resource export structures and the political economy of economic growth. *The World Bank Economic Review,* 19:141–174.
Juif, Dácil-Tanja, and Joerg Baten (2013): A Story of Large Land Owners and Maths-Skills: Inequality and Human Capital Development in Long-Run Development, 1820-2000. Working Paper.
Kimenyi, Mwangi S. and Njuguna S. Ndung'u (2005): Sporadic Ethnic Violence: Why has Kenya not Experienced a Full-Blown Civil War? In: Paul Collier and Nicholas Sambanis (eds.): Understanding Civil War. Volume 1: Africa. Washington DC: The World Bank.
Komlos, John (1985): Stature and Nutrition in the Habsburg Monarchy: The Standard of Living and Economic Development. American Historical Review 90: 1149-1161.
Komlos, John and Joerg Baten (Eds. 1998): The Biological Standard of Living in Comparative Perspective. Stuttgart: Steiner Verlag.
Komlos, John and Timothy Cuff (Eds. 1998): Classics of Anthropometric History: A Selected Anthology. St. Katharinen: Scripta Mercaturae.
Komlos, John (2003): An Anthropometric History of Early-modern France. European Review of Economic History Vol. 7.2003, 2, p. 159-189.
Labrousse, Ernest (1948): 1848, 1830, 1789. Comment naissent révolutions. In : Actes du congrès historique du centenaire de la révolution. Paris.

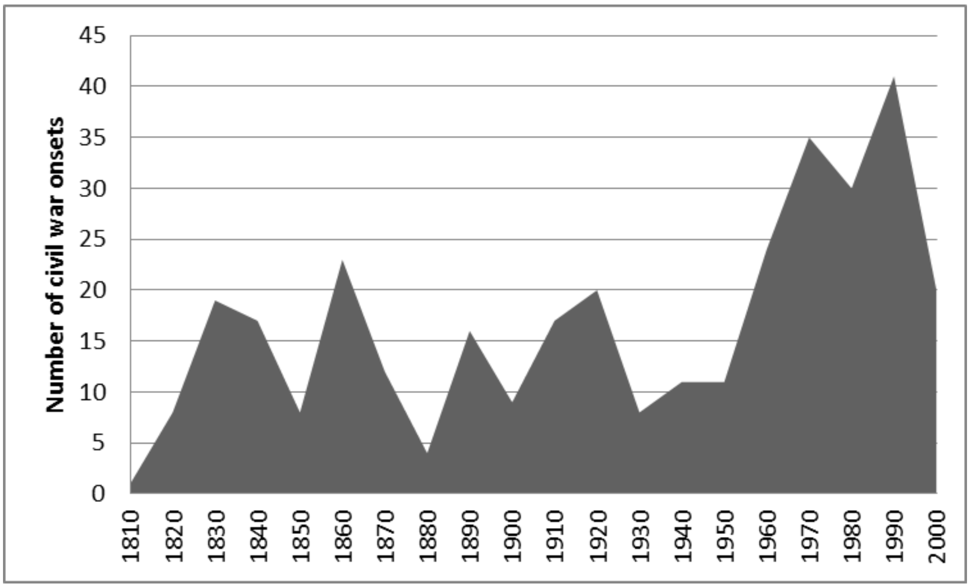
Lichbach, M. I. (1989). An Evaluation of “Does Economic Inequality Breed Political Conflict?” Studies. *World Politics, 41*(4), pp. 431-470.
Lujala, Päivi, Nils Petter Gleditsch and Elisabeth Gilmore (2005): A Diamond Curse? Civil War and a Lootable Resource. Journal of Conflict Resolution 49(4): 538-562.
Macours, Karen (2011): Increasing Inequality and civil Conflict in Nepal. Oxford Economic Papers 63: 1–26.
Maddison, Angus (2001): The World Economy: A Millennial Perspective. OECD Development Centre.
Margo, Robert and Richard H. Steckel (1983): Heights of Native Born Northern Whites during the Antebellum Period. Journal of Economic History 43, 167-74.
Miguel, Edward, Shanker Satyanath, and Ernest Sergenti (2004) "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," Journal of Political Economy, 112 (4):725-753.
Montalvo, José G., and Marta Reynal-Querol (2005): Ethnic Polarization, Potential Conflict, and Civil Wars. The American Economic Review 95(3): 796-816.
Moradi, Alexander (2005): Ernährung, wirtschaftliche Entwicklung und Bürgerkriege in Afrika südlich der Sahara (1950-2000). Inaugural Dissertation, Tübingen.
Moradi, Alexander und Joerg Baten (2005): Inequality in Sub-Saharan Africa: New Data and New Insights from Anthropometric estimates. World Development 33(8): 1233-1265.
Muller, Edward N., and Mitchell A. Seligson (1987): Inequality and insurgency. American Political Science Review 81 (2): 425-452.
Muller, Edward N (1985): Income equality, regime repressiveness,and political violence. American Sociological Review 50: 47-61.
Münkler, Herfried (2003): Die neuen Kriege. Reinbek: Rowohlt Verlag.
Nafziger, E. Wayne, and Juha Auvinen (1997): War, Hunger, and Displacement: An Econometric investication into the sources of humanitarian emergences. UNU/WIDER Working Papers Vol. 142.
Nagel, Jack (1974): Inequality and Discontent: A Nonlinear Hypothesis. World Politics 26 (4): 453-472.
Nunn, Nathan (2008): The long-term effects of Africa’s slave trades. The Quarterly Journal of Economics 123 (1): 139-176.
Nussbaum, Martha (2003): Capabilities as fundamental entitlements. Sen and social justice. Feminist Economics 9(2-3): 33-59.
Parvin. (1973). Economic Determinants of Political Unrest: An Econometric Approach. *Journal of Conflict Resolution*, 271-96.
Quiroga, Gloria and Sebastian Coll (2000): Income Distribution in the Mirror of Height Differences: The Case of Spain: 1895–1950. Journal of Income Distribution 9(1): 107–131.
Ron, James (2005): Paradigm in Distress? Primary Commodities and Civil War. Journal of Conflict Resolution 49(4): 443-450.
Ross, Michael L. (2006): A Closer Look at Oil, Diamonds, and Civil War. Annual Review of Political Science 9: 265-300.
Ross, Michael L. (2004): What Do We Know About Natural Resources and Civil War? Journal of Peace Research 41(3): 337-356.
Sachs, J. D., and Warner, A.M. (1995). Natural resource abundance and economic growth. NBER Working Paper No. 5398.
Sambanis, Nicholas (2001): Do Ethnic and Nonethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1). Journal of Conflict Reolution 45(3): 259-282.
Sarkees, Meredith Reid (2000): The Correlates of War Data on War: An Update to 1997. Conflict Management and Peace Science 18(1): 123-144.

Schubert, Hermann (2008): Anthropometrische Geschichte der Französischen Revolution. München: Verlag Europäische Wirtschaft.
Singer, J. David and Melvin Small (1972): The Wages of War, 1816-1965: A Statistical Handbook. New York.
Skaperdas, Stergios (2008): An Economic Approach to Analyzing Civil Wars. Economics of Governance 9(1): 25-44.
Smillie, Ian (2002): Dirty Diamonds. Armed Conflict and the Trade in Rough Diamonds. Fafo-Report 377. Oslo, Norway: Fafo. http://www.fafo.no/pub/rapp/377/377.pdf
Soltow, Lee (1992): Inequalities in the Standard of Living in the U.S., 1798–1875. In: R.E. Gallman and J.J. Wallis (Eds.): American Economic Growth and Standards of Living Before the Civil War. Chicago: Univ. of Chicago Press: 121-171.
Steckel, Richard H. (2009): Heights and human welfare: Recent developments and new directions. Explorations in Economic History 46: 1-23.
Steckel, Richard H. (1995): Stature and the Standard of Living. Journal of Economic Literature 33(4): 1903–1940.
Steckel, Richard H. and Roderik Floud (Eds. 1997): Health and Welfare During Industrialization. Chicago: Univ. of Chicago Press.
Stock, James H., and Motoshiro Yogo (2002): Testing for Weak Instruments in Linear IV Regression. NBER Technical Working Paper 284.
Van Zanden, Jan Luiten, Joerg Baten, Peter Foldvari, and Baas van Leeuwen (2011): World Income Inequality 1820-2000. Working Paper Univ. Utrecht/Tuebingen.
Wooldridge, J. (2002): Econometric Analysis of Cross-Section and Panel Data. MIT Press, 2nd edition.
World Bank (2011): World Development Report 2011. Conflict, Security, and Development. The World Bank: Washington, D.C.

Figure 1: The probability of civil war onset subject to inequality

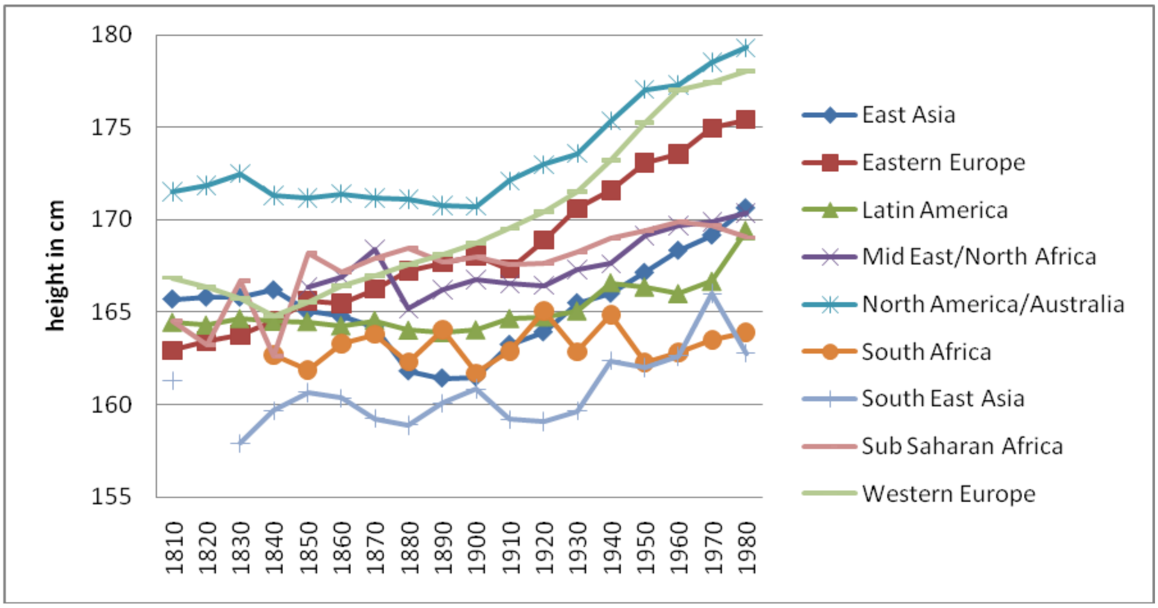


<u>Figure 2</u>: <u>Civil Wars 1820-2000</u>



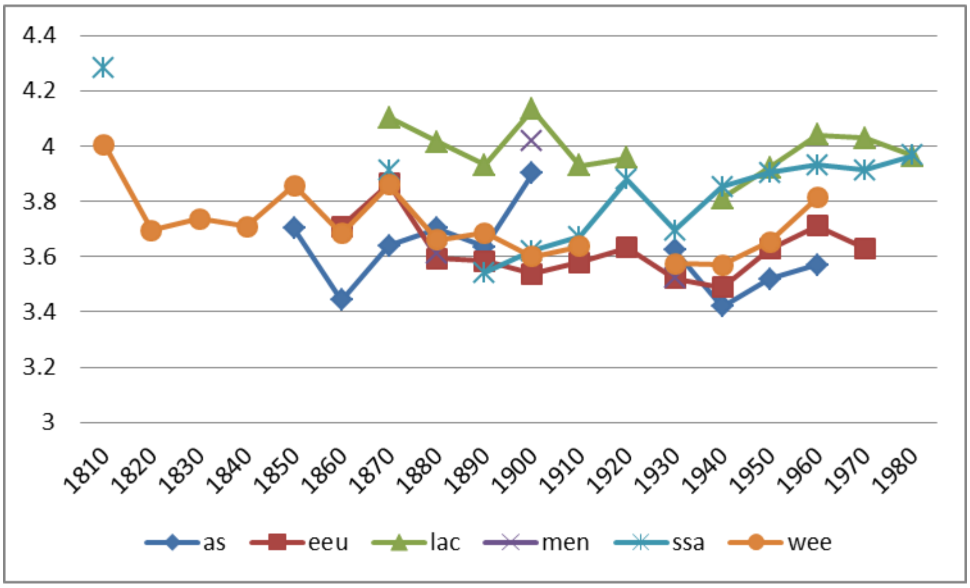
Sources: Data taken from the Correlates of War Project.

Figure 3.1: Height trends estimated by world regions



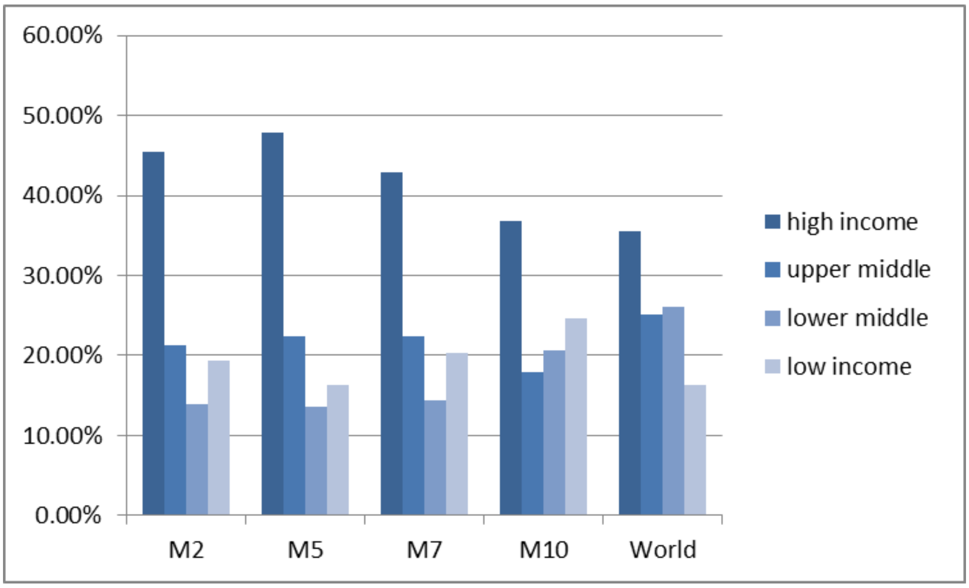
Source: Baten and Blum (2012)

Figure 3.2: Inequality trends estimated by world regions



Source: Baten and Blum (2012). Note: We omitted data points with less than four countries per given decade in a world region.

Figure 4: Data selection within the models of Table 2 by income group compared representative income distributions as reported by the World Bank



Source of data for models: Data taken from Baten and Blum (2012)
Notes: Income levels as defined by World Bank: low income: 1,005$ or less, lower middle income: 1,006$-3,975$, upper middle income: 3,976-12,275$, high income: 12,276$ or higher. Source of Income Classifications and Distribution of Income: World Bank, World Development Indicators.

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Mean if cwon=1	Mean if cwon=0
Civil War Onset	3380	0.07	0.25	1	0
Inequality	708	3.77	0.37	3.93	3.75
Height	1330	1.68	0.05	1.66	1.68
Av. Conflict years t-1	3211	0.19	0.99	1.07	0.13
Colony	3380	0.22	0.41	0.09	0.23
Democracy	1525	-0.54	6.69	-2.57	-0.21
Democracy^2	1525	0.45	0.35	0.31	0.47
Diamond	3380	0.13	0.34	0.24	0.12
Ethn. Fract.	3140	0.46	0.26	0.47	0.46
Language Fract.	3100	0.40	0.28	0.38	0.40
Religious Fract.	3200	0.43	0.23	0.37	0.43
Population (log)	3199	7.85	1.84	9.31	7.74
Peace Duration	3380	7.39	5.70	5.20	7.55
Only nonmissing values for main specification of Table 2 (Model 1 and 2)					
Civil War Onset	403	0.16	0.36	1	0
Inequality	403	3.80	0.36	3.94	3.77
Height	403	1.69	0.05	1.66	1.69
Av. Conflict years t-1	403	0.41	1.45	0.96	0.30
Colony	403	0.06	0.24	0.08	0.06
Democracy	403	-0.16	6.78	-2.29	0.23
Democracy^2	403	0.46	0.35	0.29	0.49
Diamond	403	0.16	0.37	0.24	0.14
Ethn. Fract.	383	0.41	0.28	0.44	0.4
Language Fract.	375	0.33	0.3	0.31	0.33
Religious Fract.	384	0.44	0.24	0.39	0.45
Population (log)	403	9.43	1.39	9.61	9.40
Peace Duration	403	7.53	5.96	4.10	8.16

Note: The sample includes 224 country-decade observations in which a civil war took place. The absolute number of civil wars within the observed time period is higher (306 civil wars). Height and inequlity are lagged by one decade Source: see Appendix 1.

Table 2: Pooled Logit with clustered se, Panel Logit with clustered se Regressions of Civil War Onset

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Estimation Method	PLOG	XTLOG	XTLOG	RELOG	PLOG	XTLOG	XTLOG	XTLOG	XTLOG	XTLOG
Marginal Effects?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Inequality	0.051*	0.055*	0.059*	0.685**	0.067**	0.070*	0.088**	0.059*	0.069*	0.079***
	(0.066)	(0.075)	(0.068)	(0.031)	(0.013)	(0.057)	(0.019)	(0.078)	(0.094)	(0.005)
Height	-0.885*	-0.989**	-1.064**	-9.701**	-0.923*	-1.148**	-1.069**	-1.079**	-1.267***	-0.819***
	(0.051)	(0.032)	(0.019)	(0.045)	(0.063)	(0.026)	(0.035)	(0.019)	(0.006)	(0.005)
Population (log)	0.028*	0.033**	0.038**	0.208	0.018	0.035**	0.020	0.016	0.043**	
	(0.056)	(0.036)	(0.014)	(0.106)	(0.101)	(0.037)	(0.150)	(0.252)	(0.012)	
Democracy	0.001	0.038	0.021	-3.649	-0.352	-0.036	-0.245	-0.350	0.193	
	(0.997)	(0.907)	(0.951)	(0.188)	(0.119)	(0.918)	(0.499)	(0.237)	(0.581)	
Democracy^2	-0.057	-0.068	-0.078*	-0.920*		-0.046	-0.109*	-0.109*		
	(0.211)	(0.139)	(0.093)	(0.082)		(0.426)	(0.056)	(0.051)		
Colony	0.144**	0.151**	0.151**	1.224**		0.179***	0.146**	0.106*		
	(0.030)	(0.021)	(0.015)	(0.043)		(0.010)	(0.049)	(0.065)		
Peace Duration	-0.006*	-0.004		-0.076**	-0.010***	-0.003	-0.005	-0.008**		
	(0.064)	(0.225)		(0.040)	(0.001)	(0.320)	(0.135)	(0.032)		
Warlast			0.010							
			(0.200)							
Policy Change					-0.001					
					(0.737)					
Height Growth						-2.002				
						(0.445)				
Diamond							0.024	0.024		
							(0.699)	(0.666)		

Ethn. Fract.							0.031	0.030		
							(0.727)	(0.645)		
Time dummies?	Y	Y	Y	N	Y	Y	Y	N	Y	Y
Wreg dummies?	Y	Y	Y	N	N	Y	N	N	Y	N
N	403	403	403	375	356	364	383	383	403	663

Notes: (Country) Clustered standard errors applied in every model. P-values in parentheses, ***, **, * significant on the 1, 5, and 10%-level respectively. Note: Diamond mining and colonial status are defined as dummy variables. Fractionalization measures are time invariant. For expository purposes, we divided the variables democracy and democracy squared by 100 before running the regressions. Marginal Effects reported except for rare events logit model (Model 5). Inequality is proxied by height cv, absolute deprivation is proxied by height, both lagged by 1 decade. Peace duration: sequence of decades where no war has started up to current period. Abbreviations: PLOG: pooled logit, XTLOG: panel logit, RELOG: rare events logit. Wreg dummies: Dummies for world regions.

Table 2a: Comparison of Fractionalization Measures

	(1)	(2)	(3)
Marg. Effects?	Y	Y	Y
Inequality	0.054*	0.049*	0.052*
	(0.053)	(0.062)	(0.066)
Height	-0.607**	-0.429	-0.574*
	(0.047)	(0.116)	(0.076)
Peace duration	-0.005	-0.006*	-0.004
	(0.162)	(0.098)	(0.215)
Ethn. Fractionalization	-0.075		
	(0.216)		
Language Fract.		-0.034	
		(0.523)	
Religious Fract.			-0.000
			(0.996)
Time dummies?	Y	Y	Y
Wreg dummies?	Y	Y	Y
N	637	632	643
Log lik.			
Chi-squared	153.0	154.8	142.9

(Country) Clustered standard errors applied in every model. P-values in parentheses, ***, **, * significant on the 1, 5, and 10%-level respectively. Inequality is proxied by height cv, absolute deprivation is proxied by height, both lagged by 1 decade. Peace duration: sequence of decades where no war has started up to current period. Other notes: see Table 2.

Table 3: Linear Probability Model of Civil War Outbreak

	(1)	(2)	(3)	(4)
Estimation Method	XTREG	RE	XTREG	FE
Inequality	0.121**	0.135**	0.124**	0.084*
	(0.014)	(0.011)	(0.018)	(0.056)
Height	-0.47	-0.965*	-0.626	-1.351**
	(0.469)	(0.055)	(0.370)	(0.037)
Population (log)	0.026	0.018	0.028*	
	(0.201)	(0.202)	(0.092)	
Democracy	0.298	-0.176	0.382	
	(0.521)	(0.579)	(0.428)	
Democracy^2	-0.056	-0.0762	-0.065	
	(0.341)	(0.202)	(0.260)	
Colony	0.161*	0.143*	0.157	
	(0.093)	(0.091)	(0.105)	
Spell Counter	-0.006	-0.009**	-0.005	
	(0.199)	(0.021)	(0.291)	
Ethn. Fract.			-0.051	
			(0.686)	
Diamond			0.019	
			(0.810)	
Time dummies?	Y	Y	Y	Y
Wreg dummies?	Y	N	Y	N
Constant	0.196	1.124	0.414	2.041*
	(0.853)	(0.216)	(0.713)	(0.062)
N	403	403	383	663
Log lik.				-17.06
Chi-squared	274.6	60.41	345.1	
R^2	0.169	0.169	0.052	0.052

Heteroskedasticity-robust clustered standard errors. P-values in parentheses, ***, **, * significant on the 1, 5, and 10%-level respectively. Note: Diamond mining and colonial status are defined as dummy variables, Fractionalization measures are time invariant. For expository purposes, we divided the variables democracy and democracy squared by 100 before running the regressions. Inequality is proxied by height cv, absolute deprivation is proxied by height, both lagged by 1 decade. Peace duration: sequence of decades where no war has started up to current period. In model 4, the within-R^2 is reported. Other notes: see Table 2.

Table 4: Robustness test: Panel Logit Omitting the Very Rich and Very Poor Countries

	(1)	(2)	(3)	(4)
Omitted	GDP p.c.<1000 US$		GDP p.c.>13200 US$	
Inequality	0.023	0.063	0.106**	0.098*
	(0.497)	(0.126)	(0.041)	(0.087)
Height	-1.113***	-0.986**	-0.790	-1.275***
	(0.004)	(0.046)	(0.159)	(0.006)
Population (log)	0.027*	0.003	0.014	0.020
	(0.072)	(0.863)	(0.474)	(0.245)
Democracy	0.114	-0.191	0.132	-0.094
	(0.651)	(0.555)	(0.774)	(0.817)
Democracy^2	-0.044	-0.082	-0.033	-0.071
	(0.321)	(0.142)	(0.554)	(0.264)
Colony	0.062	0.020	0.155**	0.160**
	(0.491)	(0.872)	(0.038)	(0.031)
Peace duration	-0.001	-0.006	-0.005	-0.006
	(0.660)	(0.139)	(0.189)	(0.120)
Diamond		0.044		0.025
		(0.607)		(0.631)
Ethn. Fract.		0.043		0.032
		(0.621)		(0.649)
Time dummies?	N	Y	Y	Y
Wreg dummies?	Y	N	Y	N
N	303	284	287	284
Log lik.	-95.32	-100.1	-97.62	-98.46
Chi-squared	342.5	245.0	147.1	157.8

Notes: Estimation: Pooled Logit Clustered standard errors. P-values in parentheses, ***, **, * significant on the 1, 5, and 10%-level respectively. Marginal Effects reported. Note: Diamond mining and colonial status are defined as dummy variables. Fractionalization measures are time invariant. For expository purposes, we divided the variables democracy and democracy squared by 100 before running the regressions.

Table 5: IV-Regressions of Civil war:

	(1)	(2)	(3)	(4)	(5)
Estimation Method	IVPROBIT	TSLS	TSLS	LIML	LIML
First stage		**Dependent Variable: Inequality**			
(1) Lactose Tolerance	0.235**	0.240	0.235	0.197	0.243
	(0.028)	(0.104)	(0.208)	(0.328)	(0.221)
(2) Wheat/Sugar	-0.610***	-0.623***	-0.610***	-0.626**	-0.613***
	(0.000)	(0.001)	(0.003)	(0.010)	(0.004)
(3)Lowpop*Southern					0.000
					(0.891)
First stage		**Dependent Variable: Height**			
(1) Lactose Tolerance	0.691***	0.067***	0.069***	0.074***	0.068***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
(2) Wheat/Sugar	0.024*	0.028	0.024	0.059*	0.020
	(0.093)	(0.315)	(0.483)	(0.085)	(0.575)
(3) Lowpop*Southern					0.000
					(0.904)
Second stage		**Dependent Variable: Civil War Onset**			
Marg. Effects?	N	Y	Y	Y	Y
Inequality	2.614**	0.464***	0.515**	0.502**	0.527**
	(0.050)	(0.004)	(0.036)	(0.045)	(0.044)
Height	-8.218	-1.758	-1.666	-1.664	-1.762
	(0.184)	(0.113)	(0.220)	(0.216)	(0.208)
Population (logs)	0.231**	0.025	0.037	0.043	0.037
	(0.021)	(0.255)	(0.277)	(0.201)	(0.294)
Democracy	-2.755		-0.253	-0.365	-0.243
	(0.211)		(0.602)	(0.491)	(0.622)
Democracy^2	-0.126		0.002	0.046	0.015
	(0.792)		(0.988)	(0.686)	(0.901)
Colony	0.316	-0.0913*	0.0347	0.136	-0.026
	(0.616)	(0.059)	(0.777)	(0.273)	(0.822)
Peace Duration	-0.000	0.004	0.002	-0.000	0.001
	(0.996)	(0.526)	(0.716)	(0.967)	(0.833)
Constant	0.430	1.085	0.634	0.295	0.754
	(0.973)	(0.598)	(0.816)	(0.918)	(0.780)
Time dummies?	N	N	N	Y	N
N	250	368	250	250	242
AP-F height		33.82	24.48	31.13	13.2
AP-F inequality		11.65	9.53	6.49	4.6
Kleinbergen-Paap Lm stat. p-val.		0.002	0.005	0.014	0.023
Anderson-Rubin p-val.		0.000	0.013	0.003	0.012
Hansen p-val.		exactly identified			0.578

Notes: Heteroskedasticity- and cluster-robust standard errors. Instruments: (1) Lactose Tolerance; (2) Sugar/Wheat suitability ratio; (3) Low Population density * southern latitude. It was not possible to obtain marginal effects for ivprobit twostep estimates (Model 1), therefore size of coefficients cannot be directly compared. P-values in parentheses, ***, **, * significant on the 1, 5, and 10%-level respectively. AP-F are Angrist-Pischke F-tests. Colonial status is defined as dummy variable. For expository purposes, we divided the variables democracy and democracy^2 by 100 before running the regressions.

Table 6: IV-Regressions of Civil War Onset, Controlling for Primary Goods Exports:

	(1)	**(2)**	**(3)**
	IVPROBIT	TSLS	LIML
First stage	**Dep. Variable: Inequality**		
(1) Lactose Tolerance	0.156	0.072	0.082
	(0.158)	(0.692)	(0.667)
(2) Wheat/Sugar	-0.708***	-0.855***	-0.858***
	(0.000)	(0.000)	(0.000)
(3) Pop*Southern			0.000
			(0.790)
First stage	**Dep. Variable: height**		
(1) Lactose Tolerance	0.075***	0.082***	0.080***
	(0.000)	(0.000)	(0.001)
(2) Wheat/Sugar	0.025*	0.069*	0.069*
	(0.116)	(0.062)	(0.066)
(3) Pop*Southern			-0.000
			(0.829)
Second stage	**Dep. Variable: Civil War Onset**		
Marg. Effects?	N	Y	Y
Inequality	2.156*	0.431**	0.435**
	(0.067)	(0.019)	(0.019)
Height	-8.760	-1.847	-1.721
	(0.212)	(0.121)	(0.145)
Natural Resources	-0.002	-0.002	-0.002
	(0.685)	(0.272)	(0.299)
Population (logs)	0.022	-0.015	-0.018
	(0.854)	(0.401)	(0.323)
Democracy	-1.048	0.147	0.132
	(0.653)	(0.790)	(0.813)
Democracy^2	-0.027	0.042	0.042
	(0.957)	(0.723)	(0.719)
Colony	0.254	0.147	0.099
	(0.686)	(0.159)	(0.348)
Peace Duration	0.008	0.003	0.001
	(0.811)	(0.535)	(0.832)
Constant	4.996	1.521	1.320
	(0.722)	(0.535)	(0.587)
Time dummies?	N	Y	Y
N	228	228	225
AP-F height		17.98	5.23
AP-F inequality		10.6	9.17
Kleinbergen-Paap Lm stat. P-val.		0.020	0.065
Anderson-Rubin p-value		0.001	0.000
Hansen p-val.	exactly identified		0.316

Notes: Heteroskedasticity- and cluster-robust standard errors. Marginal effects reported. P-values in parentheses, ***, **, * significant on the 1, 5, and 10%-level respectively. Instruments: (1) Lactose Tolerance; (2) Sugar/Wheat suitability ratio; (3) Low Population density * southern latitude. Angrist-Pischke F-values are reported. Colonial status is defined as dummy variable. For expository purposes, we divided the variables democracy and $democracy^2$ by 100 before running the regressions**.**

Appendix 1: Definition of the Variables

Civil war is coded as a dichotomous variable adopting the value 1 if civil war broke out in given country and five-year period. It is defined as sustained combat with at least 1,000 battle-related deaths per year that takes place between the armed forces of a government and forces of another entity for central control or for local issues. Military and civilian deaths are counted. Sources: Correlates of War Project and Uppsala Conflict Data Project: http://www.correlatesofwar.org/

Colonydummy is coded as a dichotomous variable adopting the value 1 if a country was a colony. Source: Correlates of War 2 Project. *Colonial/Dependency Contiguity Data, 1816-2002*. Version 3.0

Democracy indicates the openness of democratic institutions in a country and is measured on a scale of -10 (low) to 10 (high). Source: Polity IV Project.

Diamond is coded as a dichotomous variable adopting the value 1 if a country had diamond deposits that could be extracted.

Ethnic fractionalization: based on a combination of racial and linguistic characteristics and defined as 1 minus the Herfindahl index of group shares of these characteristics. Source: Alesina et al. (2003). Data available at http://www.nsd.uib.no/macrodataguide/set.html?id=16&sub=1.

Height describes the average adult male height in the respective country ten years prior to the time of observation (except for the instrumental variables regressions, where we take current height); measured in centimetres. Sources: Baten (2006), Measure DHS (Demographic and Health Surveys) project and the World Health Organization's (WHO) Global Database on Child Growth and Malnutrition.

Height growth is the growth rate of heights between two decades: $(height_t - height_{t-1}) / height_{t-1}$

Inequality is the coefficient of height variation at time t in the case of instrumental variables regressions and t-1 for all other specifications.

Language Fractionalization is defined as 1- Herfindahl Index of linguistic group shares, which reflects the probability that two randomly selected persons belong do different groups. Source: Alesina et al. (2003). Data available at http://www.nsd.uib.no/macrodataguide/set.html?id=16&sub=1.

Natural resource exports: percentage value of raw materials and mining products, relative to total exports, around 1980. Source: World Bank Data 1999 (CD-Rom)

Religious Fractionalization is defined as the probability that 2 randomly selected individuals belong to different religious groups. Source: Alesina et al. (2003). Data available at http://www.nsd.uib.no/macrodataguide/set.html?id=16&sub=1.

Peace duration is the number of decades that country i has not been affected by conflict. Source: Intra-State war dataset (version 4.1, posted in March 2011) from the Correlates of War project

Population (log) is the log of a country's population at the beginning of a ten-year period. Sources: The World Bank and Maddison (2001).

Warlast counts the number of months a country has experienced internal conflicts during the previous decade and divides them by 12 to get the average number of years. We include all intrastate conflicts (apart from civil wars also regional internal and intercommunal conflicts) as we want to capture the overall dissatisfaction. Source: Intra-State war dataset (version 4.1, posted in March 2011) from the Correlates of War project.

Table A1: World regions, individual countries, and birth decades: coverage of the data set (yellow indicates that real data was available and was accepted for height and civil war)

		Birth Decade																			
		18..									19..										20..
wreg	co	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90	0
as	bd																				
as	cn																				
as	hk																				
as	id																				
as	in																				
as	jp																				
as	kh																				

as	kp																				
as	kr																				
as	la																				
as	lk																				
as	mm																				
as	mn																				
as	my																				
as	np																				
as	pg																				
as	ph																				
as	pk																				
as	sg																				
as	th																				
as	tp																				
as	tw																				
as	vn																				
eeu	al																				
eeu	am																				
eeu	az																				
eeu	ba																				
eeu	bg																				
eeu	by																				
eeu	cs																				
eeu	cz																				
eeu	ee																				
eeu	ge																				
eeu	hr																				
eeu	hu																				
eeu	kg																				
eeu	kz																				
eeu	lt																				
eeu	lv																				
eeu	md																				
eeu	mk																				
eeu	pl																				
eeu	ro																				
eeu	ru																				
eeu	si																				
eeu	sk																				
eeu	tj																				
eeu	tm																				
eeu	ua																				
eeu	uz																				
lac	ar																				
lac	bo																				

lac	br																				
lac	cl																				
lac	co																				
lac	cr																				
lac	cu																				
lac	do																				
lac	ec																				
lac	gt																				
lac	gy																				
lac	hn																				
lac	ht																				
lac	jm																				
lac	mx																				
lac	ni																				
lac	pa																				
lac	pe																				
lac	pr																				
lac	py																				
lac	sv																				
lac	tt																				
lac	uy																				
lac	ve																				
men	af																				
men	dz																				
men	eg																				
men	er																				
men	il																				
men	iq																				
men	ir																				
men	jo																				
men	kw																				
men	lb																				
men	ly																				
men	ma																				
men	om																				
men	sa																				
men	sy																				
men	tn																				
men	tr																				
men	ye																				
naa	au																				
naa	ca																				
naa	nz																				
naa	us																				
ssa	ao																				

ssa	bf																				
ssa	bi																				
ssa	bj																				
ssa	bw																				
ssa	cd																				
ssa	cf																				
ssa	cg																				
ssa	ci																				
ssa	cm																				
ssa	et																				
ssa	ga																				
ssa	gh																				
ssa	gm																				
ssa	gn																				
ssa	gq																				
ssa	gw																				
ssa	ke																				
ssa	km																				
ssa	lr																				
ssa	ls																				
ssa	mg																				
ssa	ml																				
ssa	mr																				
ssa	mu																				
ssa	mw																				
ssa	mz																				
ssa	na																				
ssa	ne																				
ssa	ng																				
ssa	rw																				
ssa	sd																				
ssa	sl																				
ssa	sn																				
ssa	so																				
ssa	sz																				
ssa	td																				
ssa	tg																				
ssa	tz																				
ssa	ug																				
ssa	za																				
ssa	zm																				
ssa	zw																				
wee	at																				
wee	be																				
wee	ch																				

wee	cy																				
wee	de																				
wee	dk																				
wee	es																				
wee	fi																				
wee	fr																				
wee	gr																				
wee	ie																				
wee	it																				
wee	nl																				
wee	no																				
wee	pt																				
wee	se																				
wee	uk																				

Note: Migrant heights (unadjusted), with the number of birth decades in parentheses, was used in the following countries: Algeria (2), Armenia(1), Bangladesh (4), Croatia (Hrvatska) (1), Czech Republic (1), India (6), Israel (1), Korea (North) (6), Malawi (1), Mozambique (1), Pakistan (1), Poland (2), Romania (1). Source: Baten and Blum (2012).”men” denotes countries from the middle East and North Africa. The group “naa” includes North America, Australia and New Zealand.

Figure A1: The Probability of civil war onset subject to democracy

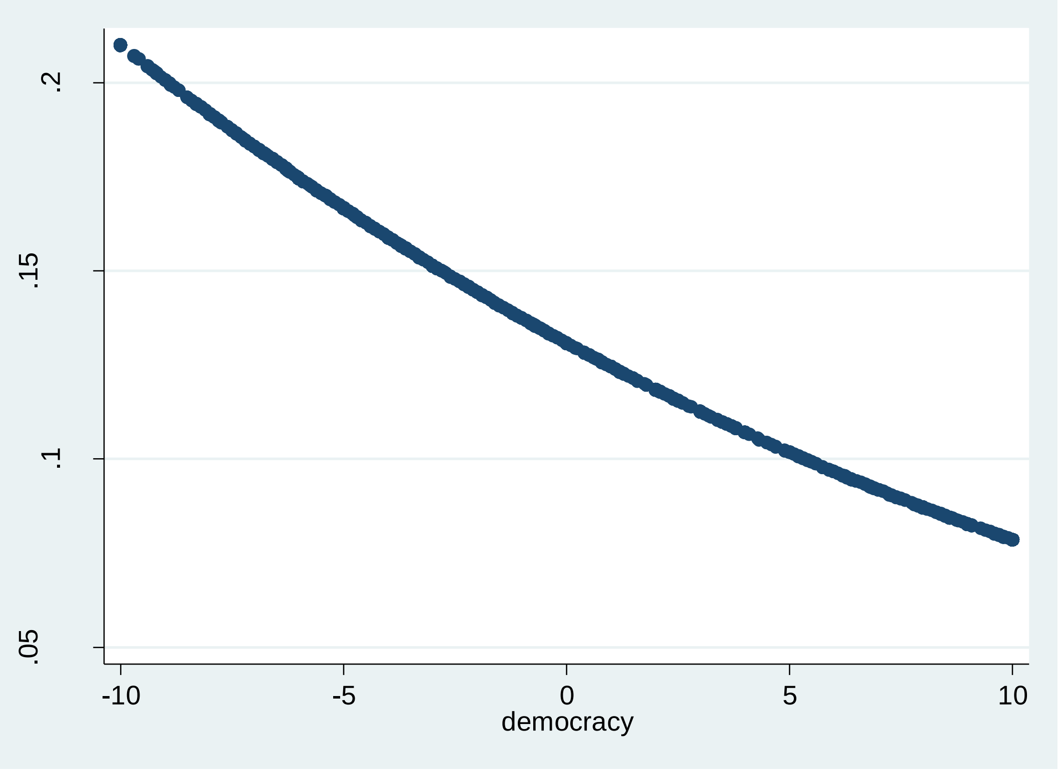


Table A2: Multicollinearity Statistics. Variance inflation Factors for the Independent Variables and Correlation between Height inequality and lagged average height

Variable	VIF	1/VIF
Inequality	1.17	0.855
Lag height	1.45	0.689
Democracy	1.40	0.717
(Democracy)^2	1.51	0.663
Population (logs)	1.31	0.765
Ethnic polarization	1.17	0.856
Colony	1.16	0.862
Diamond mining	1.18	0.845
Mean VIF	1.29	

	Lag height	Inequality
Lag height	1	
Inequality	-0.045	1

Table A3: Horserace between Inequality and lagged inequality

	(1)
	Panel Logit
Inequality	0.041
	(0.143)
Inequality t-1	0.040*
	(0.094)
Height	-0.464
	(0.173)
Population (log)	0.013
	(0.172)
Peace Duration	-0.002
	(0.302)
Colony	-0.023
	(0.431)
Time dummies?	Yes
N	454
Chi-squared	101.7

Table A4: Within- and between variation of variables

Variable		**Mean**	**Std. Dev.**
Civil War Onset	overall	0.07	0.25
	between		0.10
	within		0.23
Inequality	overall	3.77	0.37
	between		0.28
	within		0.27
Height	overall	1.68	0.05
	between		0.03
	within		0.03
Democracy	overall	-0.01	0.07
	between		0.05
	within		0.05
Democracy^2	overall	0.45	0.35
	between		0.23
	within		0.27
Colony	overall	0.22	0.41
	between		0.26
	within		0.32
Diamond	overall	0.13	0.34
	between		0.34
	within		0.00
Ethn. Fract	overall	0.46	0.26
	between		0.26
	within		0.00
Population (logs)	overall	7.85	1.84
	between		1.67
	within		0.80