Youth Bulge and Mid-Life Moderation: Large Cohort Size Effects, Economic Perspectives and Civil Conflict in Sub-Saharan Africa.

Ludwig, Markus

Department of Business and Economics, University of Basel

December 2013

Youth Bulge and Mid-Life Moderation: Large Cohort Size Effects, Economic Perspectives and Civil Conflict in Sub-Saharan Africa.*

Markus Ludwig†

Department of Business and Economics, University of Basel, Switzerland

This version: December 2013

Abstract

I argue that large youth cohorts consisting of males aged 15–19 increase the risk of civil conflict, by boosting the pool of potential recruits (PR) along the extensive (cohort size) and intensive (marginalization) margin. As these large cohorts transit out of the pool of PR in their mid-40s, the risk of civil conflict declines again. My estimates show that the size of male cohorts aged 15–19 is strongly positively related to the risk of the onset of civil conflict for a sample of African countries from 1960 to 2009. A one-percent increase in this population group increases the likelihood of the onset of civil conflict by 1.4 percentage points. The results further show that better economic perspectives—in particular, high contemporaneous rainfall, higher agricultural output and foreign aid—considerably hamper this effect. Also urbanization mitigates the impact of youth bulges. This suggests that civil conflict is more likely when several adverse factors coincide and that economic conditions affect civil conflicts via a marginalization of youth population. In turn, a one-percent increase in the male population aged 40–44 reduces the risk of conflict by 0.8 percentage points. To avoid endogeneity and omitted variable bias, I use rainfall variation in the birth year of the respective age cohorts as an instrument for the cohort sizes. To support my argument, I then show that rainfall affects the infant mortality rate, and hence birth cohort sizes. Finally, I show that youth bulges also drive low-level social violence in Africa.

Keywords: Civil conflict, Youth bulge, Rainfall shocks, Instrumental variable regression
JEL: D7, J1, Q1

*I thank Frank Krysiak, Michael Wagner, Matthias Flückiger, Hermione Miller-Moser, Daniel Müller, the participants of the Scottish Economic Society 2013 Annual Conference and the participants of the Spring Meeting of Young Economists 2013 for helpful discussions and comments.

†Corresponding author. Tel.:+41 (0)61 267 33 59 ; fax:+41 (0)61 267 32 51 ; Postal address: Peter Merian-Weg 6, Postfach, CH-4002 Basel. Email addresses: markus.ludwig@unibas.ch

1 Introduction

Civil wars and the accompanying violence not only inflict severe human suffering, but are also seen as a main impediment to economic development and a main constraint on meeting the Millennium Development Goals (World Bank, 2011). Especially, Sub-Saharan Africa (SSA) has been plagued by civil wars, it also comprises most of the world's least developed countries. War and post-war economies are heavily burdened by death and displacement and as well as labor force and human capital losses. While the results of civil wars are undoubtedly devastating, there is limited agreement on what causes the onset of civil conflicts.

Young men are generally identified as being among the main perpetrators in civil conflicts, since rebel forces rely on this population group. Accordingly, a link between the size of youth cohorts and political violence is highlighted in the prominent theoretical arguments of Goldstone (1991) and Goldstone (2010). As a result, there is an increasing interest to establish and quantify the effect of fluctuations in the cohort size of the male youth population (youth bulges)[1] on the likelihood of political violence and civil wars (Urdal, 2006; Collier, Hoeffler, and Rohner, 2009). In the case of SSA, the impact of large youth cohorts on civil wars is particularly important. This is because SSA is a region, with a very youthful and fast-growing population which will experience a high level of youth population pressure in the years to come—especially, in the light of its rigid economic systems and already strained labor markets (World Bank, 2009; Urdal, 2012).

Assessing the impact of cycles of youth bulges on the onset of civil conflicts is difficult, owing to the problem of omitted variable bias and endogeneity. To my knowledge, these problems are not rigorously addressed in the existing literature, which makes it difficult to leap from correlation to causality. For example, endogeneity might weigh heavily, as civil conflict affects population sizes through death and displacement. The bias is even amplified in the case of the standard measure[2] of relative youth bulges—youth relative to total (adult) population—which is essentially constructed from two variables. For example, civil war might hurt the elderly and children disproportionately (Blattman and Miguel, 2010). This increases the relative size of the youth bulge at the onset of civil conflict and, as a consequence, leads to a spurious correlation.[3]

To avoid these traps, I employ the absolute size of the age cohort-specific population. This is feasible, since I only use time-detrended and hence trend-stationary within-country variation. Basically, I evaluate wether we can observe on average more onsets of civil conflicts in years with country-specific positive shocks in the size of the male youth population, holding all other social and economic conditions constant.

[1]In this study, I refer to the youth bulge as the absolute size of the respective male youth population cohort. The relative youth bulge refers to the relative size of the youth population to the total population.

[2]See for example Collier and Hoeffler (2004) and Urdal (2006).

[3]There are additional problems that stem from the relative measure of youth bulges which I discuss below.

To address the endogeneity and omitted variable problem, I use exogenous historical rainfall variation in the birth years of the respective age cohorts as an instrument for the latter's size. Rainfall is a plausible instrument in the case of developing countries, since rural income in rainfed farming systems and hence child nutrition and health services depend on it (Devereux, 2009). Furthermore, it is well known that infants are especially vulnerable, owing mostly to malnutrition, to negative income shocks which result in increased infant mortality (Baird, Friedman, and Schady, 2011; Rice et al., 2000). I argue and show that, in regions where available income and consequently nutrition and medical care depend on rainfed agricultural output, rainfall variation leads to fluctuations in the infant mortality rate and hence to cohort-size fluctuations. On the other hand, I show that the signal of historical short-term rainfall shocks disipates quickly enough such that by the time a given age cohort reaches the age of 15 no effect can be attributed to current socioeconomic effects. Thus, I identify the size of the male youth population with weather shocks in the respective birth years. These weather shocks are independent of any other variables influencing civil conflict, in particular current income.

My study concentrates on SSA, as this region relies to 96 percent[4] on rainfed agriculture and experiences the world's most dynamic population. Additionally, the region suffers from high child mortality, owing to malnutrition and poor medical care. It also includes most of the least developed countries and is one of the regions most prone to civil war.

Combining the Uppsala/PRIO dataset on civil conflicts of Gleditsch et al. (2002) with the GPCC rainfall data, I show for SSA, 1960-2009, that the population size of male youth aged 15–19 is significantly positively related to the risk of civil war onset. My point estimates suggest a considerable leverage effect. An increase of one percent in the male population aged 15–19 increases the likelihood of the outbreak of civil conflict by 1.4 percentage points. Rainfall shocks affect both the size of the cohort aged 15–19 and the onset of civil war, with a time lag of 17 to 19 years. Thus, on average, there is a youth bulge effect on the onset of civil conflict after 17 to 19 years. This places the relevant cohort size for triggering civil war at an age (15–19) when labor-force entry and transition into adulthood in SSA take place (Garcia and Fares, 2008) and the critical age with respect to male physical violence and aggression is reached. In turn, the data reveals a significant negative relationship between the risk of conflict and the size of the male population aged 40–44. Increasing this cohort by one percent decreases the risk of conflict by 0.8 percentage points.

I argue by means of a simple formalized relationship that large population cohorts contribute to the *change* over time in the risk of civil conflict at two points: At the age of 15–19 when they boost the size of the pool of potential recruits (PR)—by higher marginalization (intensive) and larger numbers (extensive)—and again at the age of 40–44 when the respective cohort transits out of the

[4] See World Bank (2008).

pool of PR.[5] Large youth cohorts experience a higher marginalization rate, since transition into the pool of PR and into the labor market takes place at the same age (15–19). This transition towards the labor market is, however, associated with a bottleneck effect which spurs marginalization, due to rigid economic and political institutions. Thus, a large male cohort elevates the risk level at the age 15–19 until the age 40–44 when the cohort drops out of the pool of PR. This is consistent with the literature which indicates a heavy concentration in the mass of battle related deaths for the male population aged 15–44 (Murray et al., 2002). Importantly, my results suggest that youth bulges trigger conflict by inducing fluctuations in the size of the pool of PR. However, they do not imply that male youth aged 15–19 account for the majority of the rebel army recruits.

The data also shows that various factors that improve economic perspectives hamper the impact of the youth bulge. For example, high contemporaneous rainfall and, as a consequence, higher agricultural income reduces the risk of civil conflict attributable to youth bulges. Thus, the results suggest that youth bulges increase the number of marginalized youth, since the rigid economic and political systems in SSA are not capable of absorbing exceptionally large youth cohorts. Marginalized youth with low opportunity costs reduce recruitment costs and make insurgencies more likely. This highlights the importance of the marginalization effects in the context of youth bulges. To confirm these findings, I show that youth bulges also trigger the onset of low-level social violence (SCAD dataset).

The second main result is that, in fact, contemporaneous rainfall shocks negatively affect infant mortality via agricultural output in the SSA sample. This links weather shocks directly to the age cohort-specific population size. My results, therefore, stress the importance of fluctuations in the size of the male youth population for the likelihood of civil conflicts. Since large population cohorts are a recurrent phenomenon, my results are also able to explain the cycles of civil conflicts found in many countries (World Bank, 2011).

In Section 2, I discuss the underlying mechanisms and review the relevant literature. In Section 3 and 4, I discuss data and methodology. Section 5 and 6 present the results, and Section 7 and 8 include additional consistency checks. Section 9 concludes.

2 Discussion of Mechanisms and Literature

In this section, I discuss the mechanisms between birth-year rainfall and cohort-size fluctuations as well as youth bulges and civil conflict. Additionally, I review the relevant literature.

[5] That is, only when males aged 15–19, on average, become recruitable.

Weather Shocks and Infant Mortality

In rainfed agricultural systems, such as those in SSA, agricultural output and hence available annual household income depends on a sufficient rainy season. Insufficient rainfall can cause crop failure and a food crisis for large rural populations that lives as subsistence oriented farmers (Devereux, 2009). Under such circumstances rural income is determined by exogenous weather shocks.

Agricultural rural households employ several coping strategies when exposed to a poor rainy season and hence to a negative income shock. Three central strategies are: reducing food consumption, increasing distress maternal labor supply, and cutting health service expenses (Devereux, 2009; Bhalotra, 2010). All three are determinants of the likelihood of infant survival. Additionally, ample evidence exists that infant survival is particularly, if not almost exclusively, vulnerable to malnutrition.[6] UNICEF (2012) reports that child mortality is highest in SSA, and more than a third of child deaths can be attributed to malnutrition. The region also exhibits a high risk of neonatal deaths, emphasizing the importance of the first months of life for survival.

Baird, Friedman, and Schady (2011) exploit panel data on income and child mortality for a sample of developing countries, including SSA. The authors find a negative relationship between income and child mortality; in particular, infant mortality in the first months of life. All these studies point out that infant mortality in rural areas depends crucially on the income that is determined by sufficient rainfall. The findings imply that a large share of rural dwellers in developing countries—and especially in SSA—live under conditions, where fluctuations in rainfall impacts the survival of infants.

In regions that depend on rainfed agricultural output, variation in rainfall levels is expected to cause fluctuations in the annual size of male (and female) birth cohorts. Building on this literature, I confirm below that infant mortality is affected by levels of agricultural output and rainfall.

Youth Bulges and Civil Wars

My paper is related to the empirical literature on the determinants of civil war. Blattman and Miguel (2010) give a comprehensive review of the literature. Several of these studies include demographic measures of youth population in their analysis, because they fit into several theoretical structures on the causes of civil war onset. The literature, however, delivers ambiguous results on the interdependence between relative youth bulges and civil conflict onset. Both Collier and Hoeffler (2004) and Fearon and Laitin (2003) find no evidence to support a link between the percentage of young men in the total population and the outbreak of civil war. Within the same strand of literature, Urdal (2006) finds that the number of young men as a share of the adult population

[6] See for example Rice et al. (2000) for a review.

is positively and significantly[7] associated with an increased risk of armed civil conflict, terrorism and riots. These papers employ the share of young men in the total (adult) population, which seems problematic, owing to the issues described below. As an additional distinction, I address the endogeneity and omitted variable problem by applying an instrumental variable framework. As described in the introduction, both phenomena are likely to play a prominent role in the regression estimations, leading to reversed causality and spurious correlations.

Another strand of literature relates country-specific total population size to civil conflict (Brückner, 2010; Collier and Hoeffler, 2002, 2004). My study is set apart from this literature, since it stresses the importance of distinguishing between different population age groups. For example, the setup in Brückner (2010) does not allow him to identify and isolate effects of particular population groups, since the first-stage instrument affects the total population across all population groups. In turn, I am able to identify and separate the effect of youth population size on the onset of civil conflict.

My study relates to Miguel, Satyanath, and Sergenti (2004), who show that negative rainfall shocks, as an income instrument for a sample of SSA countries between 1981 and 1999, lead to a significant increase in the probability of civil war. The authors find that exogenous short-run variation in rainfall growth affects contemporaneous economic growth and, as a result, triggers the incidence of civil war in the following period. In this paper, I use level rainfall variation which generates, via agricultural output, fluctuations in the birth cohort population size for a much longer sample period (1960-2009). I then show how such fluctuations in cohort sizes, induced by rainfall shocks, trigger civil conflict onsets with an average lag of 17 to 19 years. For my sample, I find no significant effect of a contemporaneous or a one-year level of lagged rainfall on the onset of conflict. Also, GDP per capita is not significantly related to contemporaneous rainfall. However, in my sample a strongly positive and significant relationship can be observed between contemporaneous rainfall and agricultural output per capita. This indicates that for my sample the signal of contemporaneous rainfall—through agricultural output—is not strong enough to induce observable changes in aggregate GDP, which also includes the service and the manufacturing sector.[8] In particular, the noise generated by the latter sectors seems to override the agriculture production signal.

Nevertheless, I am able to confirm the findings of Miguel, Satyanath, and Sergenti (2004), in the sense that youth bulge effects are more pronounced in years of low agricultural income (low rainfall years). Similarly, my results directly relate to the work of Nunn and Qian (2012) who show that food aid increases the risk of conflict by increasing rebel resources (looting). Even though I

[7] Similarly, Collier, Hoeffler, and Rohner (2009) argue in favor of the existence of a youth bulge effect. However, the youth bulge measure in Collier, Hoeffler, and Rohner (2009) is mostly non-significant at the ten-percent level.

[8] Please see Miguel and Satyanath (2011) and Ciccone (2011) for further discussion on the interrelationship between rainfall and GDP over time.

also find a negative overall effect of foreign aid, my results highlight that foreign aid reduces the risk of civil conflict through youth bulges.

My work resonates with Goldstone (1991) who points out that population pressure, induced by variation in the mortality rate, caused revolutions in the pre- and early industrialized era. It also presents direct links to other studies which stress the interrelationship between climate conditions and civil conflict (Hsiang, Meng, and Cane, 2011; Burke et al., 2009).

The Mechanism and Measure of Youth Bulges and Civil Conflict

In this subsection, I, first, briefly discuss the mechanism between youth bulges and conflict, and afterwards discuss some important aspects of defining youth bulges as a measure.

The Mechanism of Youth Bulges and Civil Conflict

Africa's rebel armies mostly operate, hide and recruit in the (rural) hinterland where government power is weak (Herbst, 2000, 2004; Fearon and Laitin, 2003; Sommers, 2003). For an insurgency to be feasible, a sufficiently large pool of potential recruits (PR) must exist, since Africa's insurgencies depend almost exclusively on foot soldiers (Okumu and Ikelegbe, 2010).[9] This stock of PR usually consists of members of the marginalized male population aged 15–44, e.g., Murray et al. (2002) show that the majority of battle-related deaths are concentrated within this group. Thus, age cohorts enter this pool when they transit from adolescence into adulthood at ages ranging from about 15 to 20 and leave this pool in their mid-forties.

In turn, economic and social marginalization is a precondition that motivates a fraction of this particular age group (15–44) to actually join rebel ranks. This relates especially to large youth cohorts which transit from school into the labor market: here, both transition events, towards adulthood and towards labor market participation, usually coincide between the age of 15 and 20 (Garcia and Fares, 2008; Smith, 2011). In particular, the labor markets and political systems in rural and subsistence agriculture dominated SSA—with its rigid socio-economic structure, still accounting for *two-thirds* of the African labor force (CTA, 2012)—are unable to absorb exceptionally large youth cohorts and provide an adequate quantity and quality of labor opportunities (Garcia and Fares, 2008; Chandrasekhar, Ghosh, and Roychowdhury, 2006). Thus, when a youth bulge hits the labor market, an excess labor supply for available jobs develops, the bottleneck effect occurs, consequently increasing the fraction of marginalized young males.

This allows me to formalize a simple relationship. Let $X(T)$ be the population size of a male cohort born in year T that carries its specific marginalization rate, $\alpha(X(T))$, and where $\alpha(X(T)) \in [0, 1]$. Thus, the marginalization rate depends on the cohort size; e.g., through a bottleneck effect on

[9] Collier, Hoeffler, and Rohner (2009) discuss a similar argument.

the labor market. The age of this particular cohort in year t is $age(t) = t - T$. Then the pool of PR is defined as $PR(t) = \int_{T=t-out}^{t-in} \alpha(X(T))\, X(T)\, dT$. It is determined by all cohorts aged between the transition-age 'into', $age(t) \geq in$, and the transition-age 'out', $age(t) \leq out$, weighted by the fraction of the marginalized population within each cohort. Now, let the probability of conflict, onset(t), depend on the pool of PR

$$\text{onset}(t) = \text{onset}(PR(t)) = \text{onset}\left(\int\limits_{T=t-out}^{t-in} \alpha(X(T))\, X(T)\, dT \right). \tag{1}$$

with

$$\frac{\partial \text{onset}(t)}{\partial t} = X\left(T_{in}(t)\right) \alpha\left(X(T_{in})\right) \text{onset}'\left(PR\right) - X\left(T_{out}(t)\right) \alpha\left(X(T_{out})\right) \text{onset}'\left(PR\right), \tag{2}$$

with $T_{in}(t)$ being the birth year of the cohort that transits into the pool of PR in year t; and $T_{out}(t)$ being the upper-bound birth year which governs the transition out of PR; and onset$'\left(PR\right) > 0$ being the derivative of conflict risk to PR.

Equation (2) implies that changes over time in the risk of civil conflict are determined by the size of the cohort fluctuations that shift into and out of the pool of PR. That is, as an exceptionally large youth cohort enters the critical age in, it raises the risk of civil conflict to a new level. This, however, takes place along two dimensions. Firstly, the population size between in and out owing to a youth shock. Secondly, the particular large youth cohort which transits into the PR pool also experiences a particularly high marginalization—as this cohort simultaneously transits into the labor market, where it experiences a bottleneck effect. Thus, the fact that the youth cohorts which transit into the pool of PR simulaneously transit into the labor market, make youth bulges especially important to changes in the risk of civil conflict. It is also clear that as long as a large cohort remains within the pool of PR, the risk level itself remains high (Figure A.II). However, the risk level changes only when an exceptionally large cohort enters and exits the interval $[in, out]$. Similarly, the empirical analysis below looks at changes in the risk of conflict, since it only exploits variation over time (fixed effects).

In turn, Eq. (2) does not relate to the age composition of recruits that participate in conflicts. This composition depends on the relative size of each age cohort within the age range 15–44 and its respective marginalization rate. Hence, males aged 15-20 might not even constitute the majority of rebel fighters. In this context, youth bulges do not explain the duration of civil conflicts which is also rejected by the data. The literature even indicates that larger insurgencies might even be of shorter duration (Cunningham, Gleditsch, and Salehyan, 2009). Hence, the sustainability of conflict (duration) does not appear to depend on an additional supply of young males. However, I show below that not only civil conflict, but also the onset of low-level social violence (SCAD) is affected by youth bulges.

Youth Bulge Measures

Assume that a country i possesses the equilibrium capacity in economic and political institutions to absorb $\overline{X(T_{in})}_i$ of male youth, where the term youth population signifies the age cohort that transits into the labor market (pool of PR). Hence—given a positive youth population shock—the size of the (marginalized) excess youth labor can be written as $X(T_{in})_{i,t} - \overline{X(T_{in})}_i = \Psi > 0$. Thus, the positive deviation from the equilibrium number of male youth Ψ transits into the pool of potential recruits for insurgencies, which, in turn, increases the risk of conflict. In terms of a linear model, with onset$_{i,t}$, the probability of the onset of civil conflict, this can be formulated as follows:

$$\text{onset}_{i,t} = \alpha\,(X(T_{in})_{i,t} - \overline{X(T_{in})}_i) + \beta\, x_{i,t} = \alpha\,\Psi + \beta\, x_{i,t}, \tag{3}$$

where $x_{i,t}$ are additional factors like, such as the population structure or political institutions. Now let $x_{i,t}$ move to the error term $u_{i,t}$ which yields

$$\text{onset}_{i,t} = \alpha\,(X(T_{in})_{i,t} - \overline{X(T_{in})}_i) + u_{i,t} = \alpha\,\Psi + u_{i,t}. \tag{4}$$

However, this simply represents the fixed effects regression model in Eq. (5), assuming that the country-specific mean corresponds to the equilibrium number of male youth. Now, if it is possible to identify these country-specific fluctuations around the equilibrium ($E(X(T_{in})_{i,t}u_{i,t}) = 0$), then we can estimate its effect on the risk of civil conflict, holding everything else constant—including labor market conditions, population structure and political institutions. It is important to stress that this does not imply that countries with a generally larger population are more prone to civil conflict. However, countries with larger positive deviations from their respective youth-population country mean are more prone to civil conflict.

In turn, a scaled youth-bulge measure (e.g., the male youth population relative to total population) makes the seperation between effects of youth population fluctuations and changes in population structure difficult. For example, define the youth share $S_{i,t} = \frac{X(T_{in})_{i,t}}{P_{i,t}}$, with $P_{i,t}$ the total population. Assume $cov(X(T_{in})_{i,t}, O_{i,t}) > 0$, with $O_{i,t} = \text{onset}_{i,t}$ from above and where cov is the covariance. It then follows that $cov(S_{i,t}, O_{i,t}) > 0$, since $cov(X(T_{in})_{i,t}, S_{i,t}) > 0$.

On the other hand, assume $cov(P_{i,t}, O_{i,t}|X(T_{in})_{i,t}) \leq 0$ and it then follows[10] that $cov(S_{i,t}, O_{i,t}) \leq 0$, since $cov(P_{i,t}, S_{i,t}|X(T_{in})_{i,t}) < 0$. Thus, the share declines owing to an increase in the residual

[10]For example, the HIV epidemic in SSA led to a severe decline in the population aged $30 - -39$ (United Nations, 2003). This drop in the adult population even led to a decline in the number of marginalized youth through better employment opportunities—that is $cov(S_{i,t}, O_{i,t}) < 0$ (Coulibaly, 2004).

population—total population minus youth population—which might stem from a birth-rate surge.

Thus, an increasing share might increase the risk of civil conflict or not, which depends on the underlying mechanism of $\dot{S}_{i,t} = \frac{\partial S_{i,t}}{\partial t}$. In this case, it is not possible to separate the effect of youth pressure from other population factors. However, by using absolute numbers $X(T_{in})_{i,t} - \bar{X}_i$ as described above, allows me to hold $P_{i,t}|X(T_{in})_{i,t}$ constant if identification of $X(T_{in})_{i,t}$ is possible.

Nevertheless, by employing a 2SLS framework to identify the youth population size $X(T_{in})_{i,t}$, it is still possible to separate the effects; in particular, it is possible to hold $P_{i,t}$ constant within the youth share. In the next section, I show that using the share yields similar results.[11]

3 Data

My sample period for SSA covers the period from 1960 to 2009 (this being a much longer time frame than is used elsewhere in this field). The conflict data with at least 25 battle deaths is from the PRIO/Upsalla database, where I include only civil conflicts of type 3 and 4 that are fought in a country's own territory, this being closer to the context of youth bulges.[12] I code the onset of a civil conflict as one if there is a conflict in progress in the current year, but no conflict in the previous year. All years with no conflicts are coded as zero. All conflict years that follow on a preceding conflict year are coded missing. This corresponds to the onset definition of Themnér and Wallensteen (2012). Substantial differences exist, however, in coding the onset of conflict (Miguel, Satyanath, and Sergenti, 2004; Fearon and Laitin, 2003). My results are robust to changes in definition, as shown in the Appendix A.I.

To show that youth bulges also induce low-level social violence, I employ the Social Conflict in the Africa Database (SCAD) from Hendrix and Salehyan (2013). I code the onset of social violence for a given year as one if at least one violent event begins in that year (start year). All other years are coded as zero. Since birth-year rainfall-induced youth bulges are mostly important to rural and remote areas—due to rainfed agriculture—I drop all events that can be explicitly attributed to urban areas.

The evaluation of population effects through instrumentalizing their birth-year rainfall events requires a considerable lead time in the rainfall data (up to 30-50 years). Thus, a sufficiently long time series is needed which excludes all satellite-based datasets. My base rainfall data is drawn from the GPCC Full Data Reanalysis (Version 6) of Schneider et al. (2011) under the auspices of the World Meteorological Organization. It is a gridded monthly mean precipitation database at the 1°x1°grid level using spatial interpolation and starts from 1901. It is the most accurate

[11] By employing the instrument $z_{i,t}$ which satisfies $cov(z_{i,t}, X(T_{in})_{i,t}) \neq 0$ and $cov(z_{i,t}, u_{i,t}) = 0 \Rightarrow cov(z_{i,t}, P_{i,t}|X(T_{in})_{i,t}) = 0$, we have $cov(z_{i,t}S_{i,t}) = cov(z_{i,t}, X(T_{in})_{i,t}) \neq 0$.

[12] In the PRIO data, sideA and location coincide for type 3 and 4.

product within the GPCC product class, where the GPCC dataset contains the largest monthly precipitation database in the world. It is superior to most other available datasets on precipitation, both in terms of station density and quality control (Schneider et al., 2013). Further, it builds the in situ component for the widely used GPCP dataset. To derive country-specific yearly rainfall values, I follow Miguel and Satyanath (2011).[13] For comparison, the correlation between the GPCC-based rainfall data used here and the GPCP rainfall data (satellite based) used in Miguel and Satyanath (2011) is 0.94 for the period 1981-2009 (obs=1019). To show that the results are not driven by a particular precipitation dataset, I also use the "Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (V 3.01)" (Delaware) dataset from Matsuura and Willmott (2012) and the 1948-2010 "PRECipitation REConstruction over Land" (PREC/L) dataset (Chen et al., 2002).[14]

The data on the age distributions and total population are from *World Population Prospects (2010 Revision)*. These data are decomposed into five-year age groups, such that it is possible, for example, to observe fluctuations in the age group 15–19. Real GDP per capita and real openness are drawn from the Penn World Tables. The data on the live birth rate, the infant mortality rate, foreign aid in current U.S. dollars and urbanization is from the World Development Indicators. The democracy score (revised combined Polity score) is from the Polity IV dataset. The net agricultural production per capita index and the net agricultural capital stock in constant prices are from the FAO.

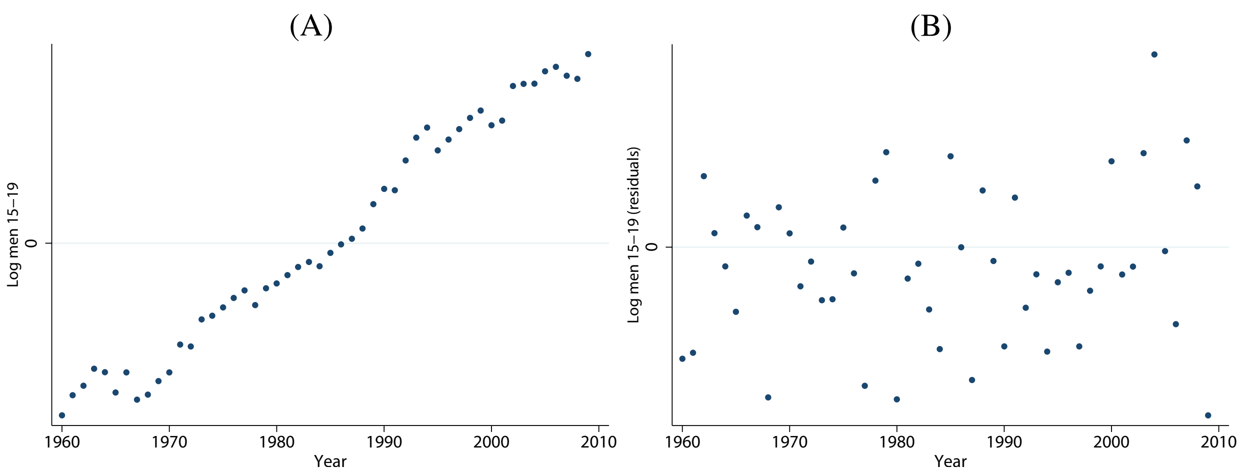

Figure I: (A) SSA average of the log men15–19 deviation from its mean. (B) SSA average of the log men15–19 residuals (mean deviation) purged by country-specific time trends, country fixed effects and time fixed effects.

According to Section 2, I expect that the cohort size fluctuations in the male population aged

[13]Basically, I use the rainfall values of grids that have the grid center inside the country polygon from GADM http:www.gadm.org.

[14]The PREC/L precipitation data is provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http:www.esrl.noaa.govpsd.

15–19 are crucial for the onset of civil conflict. In particular, it is suspected that positive deviations from the country mean (shocks) in this age group trigger conflict. Figure I displays the development of the average across the SSA countries from 1960 to 2009 relative to the mean. It is obvious that the data is not trend stationary. Thus, we find positive deviation from the mean after the mid-1980s and negative deviation before that. To remedy this problem, I control for country-specific time trends, time fixed effects and country fixed effects. Figure I shows the average purged data series which now fluctuates evenly around the mean. Thus, by controlling for these time effects in the regressions below, the youth population data becomes trend stationary. This problem is further mitigated by the fact that I use the trend stationary and exogenous rainfall data as a source of variation in my instrumental variable setup.

Generally, it is not feasible to employ growth rates of the male population aged 15–19, due to the fact that positive growth rates do not necessarily reflect a *positive*—and hence exceptional—deviation from the country-specific mean. That is, we might observe positive growth rates, even though the youth population size is below its country mean and recovering from a preceding negative shock.

4 Empirical Strategy

The setup described here for the case of the onset of civil conflict in SSA applies similarly to the onset of low-level social violence. It also applies to the cohort size aged 40–44, m40–44$_{i,t}$, which the discussion above associates with a moderating effect on the risk of conflict.

From the literature, it is to be expected that the critical age cohort $X(T_{in})$, Eq. (2), can be found within the population aged 15–19. My equation of interest, with log m15–19$_{i,t}$ the size of the male youth population aged 15–19 for country i in year t in log, is therefore

$$\begin{aligned} \text{onset}_{i,t} &= \alpha_i + \gamma_i\, t + T_t + \theta \log \text{m15–19}_{i,t} + u_{i,t} \\ &= \gamma_i\, t + T_t + \theta\, (\log \text{m15–19}_{i,t} - \overline{\log \text{m15–19}}_i) + (u_{i,t} - \bar{u}_i), \end{aligned} \tag{5}$$

where onset$_{i,t}$ is the binary variable onset of civil conflict; $\gamma_i\, t$ and T_t are country-specific time trends and time fixed effects, respectively. Including country-fixed effects in the panel setup allows me to rewrite the model as a mean deviation process, where $\overline{\log \text{m15–19}}_i$ is the country-specific mean in log. It is important to stress that Eq. (5) is an approximation in the sense that the true age of the cohort(s) that transits into the pool of PR, $X(T_{in})$, is not known ex-ante. It might be that on average the transition age is located at, e.g., 17 or that the transition period covers the whole range of cohorts aged 15–19; that is, $X(T_{in}) \subseteq$ m15–19$_{i,t}$. However, because the population data covers only five-year age groups, it is not possible to observe each age cohort separately.

Since identification might not be possible in Eq. (5), I employ birth-year rainfall variation as an instrument for the cohort size. To test if the rainfall shocks in the birth year affect the size of the respective cohorts of interest, and because these cohorts cover 5 years, I regress the log number of males aged 15–19 (log m15–19$_{i,t}$) on the log rainfall levels of the relevant birth years (log rain$_{i,t-15}$ to log rain$_{i,t-19}$),

$$\text{log m15–19}_{i,t} = \alpha_i + \gamma_i\, t + T_t + \beta_{\{t-15,..,t-19\}} \,\text{log rain}_{i,\{t-15,..,t-19\}} + v_{i,t}, \qquad (6)$$

with $\{t-15, .., t-19\}$ capturing the sequence from $t-17$ to $t-19$; $v_{i,t}$ being an idiosyncratic error term; and β capturing the effect of the birth year-specific rainfall on the cohort size.

Likewise, rainfall is expected to affect the onset of civil conflict with a lag of 15 to 19 years, since the critical age of transition (T_{in}) from school (childhood) to the labor market and into the pool of PR is between 15–19 (Section 2), and the cohort sizes are governed by birth annual rainfall. Hence, I use a similar setup to Eq. (6), to test the interdependence of lagged rainfall and civil conflict, but now using the onset of civil conflict as the dependent variable,

$$\text{onset}_{i,t} = \alpha_i + \sigma_i\, t + T_t + \lambda_{\{t-15,..,t-19\}} \,\text{log rain}_{i,\{t-15,..,t-19\}} + e_{i,t}, \qquad (7)$$

where e$_{i,t}$ is an idiosyncratic error term, and where λ captures the effect of past rain shocks on the onset of civil conflict. Similarly to (5), it is not known ex-ante which particular lagged rainfall event in the range $t-15$ to $t-19$ affects civil conflict. From the theory above, I expect that the birth-year rainfall event that corresponds to the age cohort $X(T_{in})$ has a significant impact on the onset of conflict. However, the particular rainfall event or group of events is not known ex-ante.

This setup then allows me to employ a 2SLS framework, using rainfall in the birth year as an instrument. It facilitates determining the true effect of youth bulges on the onset of civil war under the assumption that rainfall in the birth year of the critical age cohort does not affect civil conflict through another channel. Thus, the estimation procedure requires that rain shock effects disperse fast enough. Given that my analysis focuses on the short-run effects on the respective variables, this seems plausible. Furthermore, since the five-year age cohort is affected by more than one year of rainfall, I am able to test the exclusion restriction.

5 Main Results - Youth Bulges and Civil Conflict

The central focus of this paper concerns youth bulge effects, which constitute most of the discussion below. However, as outlined above, I also expect there to be a moderating effect on conflict by large cohorts aged 40–44; that is, when these cohorts transit out of the pool of PR. I present estimates on the latter effect in Section 5.4. Before discussing the main regression results, I discuss several nonparametric estimates of youth cohort size and the onset of civil conflict. Subsequently, I present the results of the first-stage and instrumental variable estimations.

5.1 Nonparametric Estimations and Scatterplots

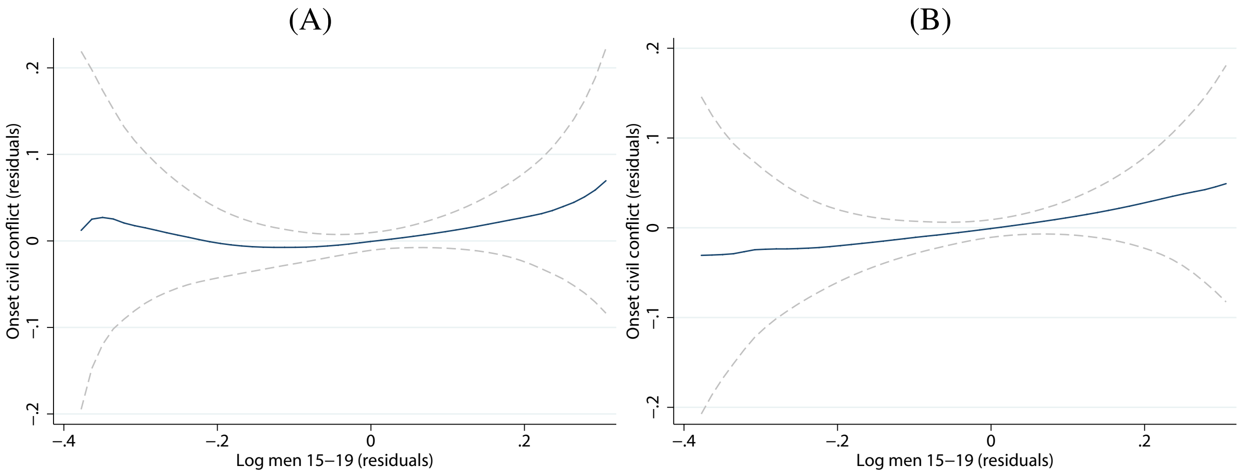

Figure II: (A) SSA sample of log men 15–19 on onset of civil conflict. (B) SSA sample, excluding the observation for Rwanda in 1996. Nonparametric local polynomial estimates and 95-percent confidence interval (dashed). Residuals are purged by time trend, country-specific time trends, country fixed effects and time fixed effects.

Figure II displays nonparametric local polynomial estimations of the residuals of the male cohort size aged 15–19 on the residuals of the onset of civil conflict.[15] Figure II(A) shows that the relationship is negative at first and then becomes positive; that is, in contrast to the monotonic and positive relationship which was expected. However, the estimation precision at the left-end is extremely low, owing to sparse observations. Additionally, the negative relationship is purely driven by the observation for the onset of civil conflict in Rwanda in 1996, which is highlighted in Figure A.I in the Appendix. Dropping the observation results in a strictly positive relationship between the youth male cohort size and the onset of civil conflict, as depicted in Figure II(B). However, dropping this particular outlier from the regressions below does not affect the results and is subsequently *not* excluded from the sample.

This first glance at the data indicates that, in fact, youth bulges and civil conflict are positively interrelated. Additional support is highlighted in Figure A.I. The bar charts to the left and right of the line through the origin (x-axis) indicate that onset events are on average associated with positive residuals of the male population size aged 15–19 (mean at 0). Thus, in 60 percent of all

[15]The residuals are obtained by purging the raw data of country-specific time trends, time fixed effects and country fixed effects.

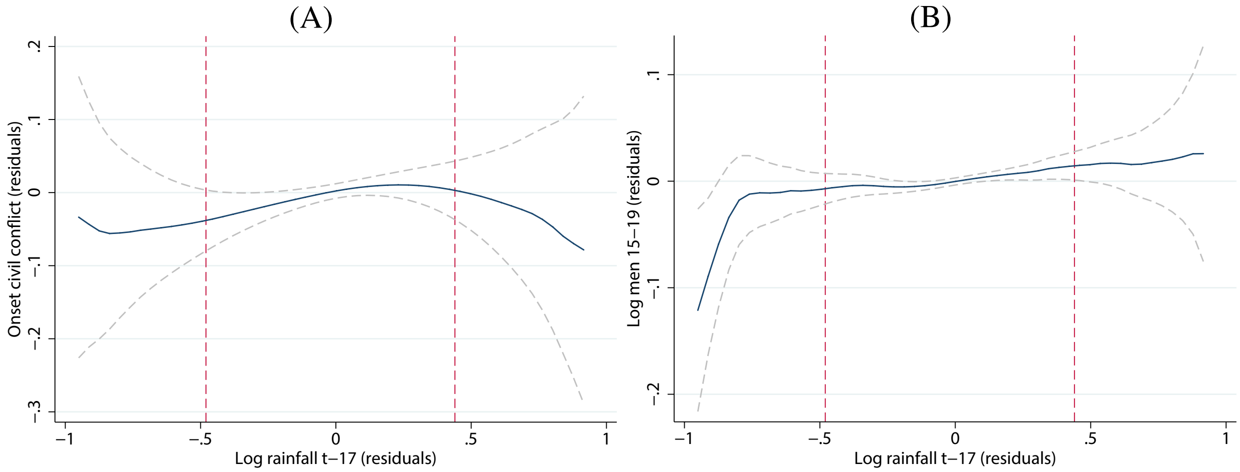

Figure III: (A) SSA sample of log rainfall $t-17$ on onset of civil conflict. (B) SSA sample of log men 15–19 on onset of civil conflict. Nonparametric local polynomial estimates and 95-percent confidence interval (dashed). The vertical lines mark the top and bottom percentile of the rainfall data. Residuals are purged by time trend, country-specific time trends, country fixed effects and time fixed effects.

cases, the onset of civil war falls within a period where the male youth population is above the mean.

Next, Figure III depicts the nonparametric estimates of the relationship between lag $t-17$ rainfall and the onset of civil conflict, and lag $t-17$ rainfall and males aged 15–19 in Figure III(A) and Figure III(B), respectively. The figure also includes vertical lines which indicate the demarcation of the bottom and top percentile of the rainfall data. To the left and the right of the vertical lines, the estimates become imprecise, owing to limited observations. Ignoring these outliers, the plots indicate a mostly positive and monotonic relationship for both the onset of civil conflict and the male youth population. In order to assert that the estimates are not driven by the outliers in the top and bottom percentile of the rainfall data, I also run regressions excluding these observations.

5.2 Rainfall, Youth Bulges and Civil Conflict

In this subsection, I discuss the direct effects of the respective birth year rainfall events on the cohort size of males aged 15–19 and the onset of civil conflict. Then, I discuss the identification strategy. It is important to stress that this analysis is concerned with the question of whether large youth population sizes induce fluctuations in the risk of civil conflict over time. It does not, however, answer the question as to which particular age cohorts are actually engaged in civil conflict—the population composition of the rebel armies.

Table I: Civil Conflict, Youth Bulge and Rainfall (OLS)

Dependent Variable:	Log Men15–19$_t$			Onset Civil Conflict$_t$		
	(1)[a]	(2)	(3)	(4)[b]	(5)	(6)
Log rain$_{t-15}$	0.030* (0.016)			-0.009 (0.029)		
Log rain$_{t-16}$	0.028** (0.014)			-0.003 (0.030)		
Log rain$_{t-17}$	0.040*** (0.014)	0.040** (0.016)		0.055** (0.022)	0.054** (0.020)	
Log rain$_{t-18}$	0.026* (0.014)			0.015 (0.032)		
Log rain$_{t-19}$	0.016 (0.016)			0.058** (0.024)		
Log $\overline{\text{rain}}_{\{t-17,..,t-19\}}$			0.083** (0.041)			0.135*** (0.048)
Obs.	1754	1754	1754	1754	1754	1754
RMSE	0.064	0.065	0.064	0.206	0.206	0.205

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses. $\overline{\text{rain}}_{\{t-17,..,t-19\}}$ is the average rainfall for the lags $t-17$ to $t-19$.
[a]Coefficients of lag rainfall $t-15$ to $t-19$ are the same (p-value 0.54).
[b]Coefficients of lag rainfall $t-17$ to $t-19$ are the same (p-value 0.45).

First-Stage and Reduced-Form Estimates

Table I, columns 1-3, show the results of lagged rainfall levels on the cohort size of young men aged 15–19. I report least square estimates with robust (Eicker-Huber-White) standard errors clustered at the country and year level.

I expect that rainfall shocks at birth affect cohort sizes. Since the dependent variable is a cohort spanning five years, there are five potential rainfall events, each affecting annual birth cohorts which, in addition, overlap.[16] Except for the lag $t-19$, the lag rainfall levels enter at least significant at the ten-percent level. However, I cannot reject the hypothesis (p-value 0.54) that all birth year rainfall events have the same coefficient size. Thus, the data reveals that these birth year rainfall events contribute in a similar way to fluctuations in the cohort size of males aged 15–19. Due to the results discussed shortly, I focus on the lag $t-17$ (column 2). However, the estimates do not rely on this particular lag, and I also present results using the average rainfall from $t-17$ to $t-19$ (column 3).

Next, I turn to the supposition that civil conflict is more likely when large cohorts of young males, on average, transit into the labor market and into the pool of PR. Following this argument, it is to be expected that rainfall shocks in the range between $t-15$ and $t-19$ are likely to have an impact on the likelihood of civil conflict starting. It is, however, ex-ante not clear which particular

[16]That is, infants born in summer are still infants in the subsequent year.

Table II: Onset of Civil Conflict and Rainfall Lags - Youth Bulge

Rainfall lags:	t+1	t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9
Onset conflict$_t$	0.021 (0.020)	0.018 (0.022)	0.016 (0.027)	0.038 (0.028)	-0.037 (0.028)	0.002 (0.042)	-0.005 (0.039)	-0.000 (0.036)	-0.057 (0.036)	-0.025 (0.040)	0.029 (0.023)
	t-10	t-11	t-12	t-13	t-14	t-15	t-16	t-17	t-18	t-19	t-20
	0.000 (0.043)	-0.006 (0.039)	-0.034 (0.021)	-0.021 (0.022)	0.028 (0.030)	-0.005 (0.026)	-0.011 (0.030)	0.048** (0.021)	0.015 (0.034)	0.056* (0.028)	-0.029 (0.032)
	t-21	t-22	t-23	t-24	t-25	t-26	t-27	t-28	t-29		
	0.015 (0.037)	-0.020 (0.035)	0.001 (0.039)	0.032 (0.034)	0.014 (0.040)	-0.032 (0.027)	0.003 (0.040)	-0.025 (0.026)	0.046 (0.035)		

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table shows the regression results of the onset of civil conflict on the rainfall lags $t+1$ to $t-29$. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses. 1,754 Observations.

age cohorts out of the age range 15–19 actually induce changes in the likelihood of conflict. In particular, it is not known at which age youth cohorts transit into the pool of PR. Hence, it is necessary to identify the cohorts which represent $X(T_{in})$ in Eq. (2). Column 4, in fact, depicts a significant and positive impact of rainfall shocks on the onset of civil conflict with a lag of between 17 and 19 years. However, I cannot reject the hypothesis (p-value 0.45) that the $t-17$ to $t-19$ rainfall events enter with a similar coefficient size. Thus, the data suggest that the relevant group, which triggers—by increasing the pool of PR ($X(T_{in})$)—the onset of civil conflict, is aged 17–19. Again, columns 5-6 show the isolated effect of the $t-17$ lag and of the average lag from $t-17$ to $t-19$. Further, Table A.I, columns 1-2, show that these findings are robust to a dynamic panel specification for both OLS and system GMM.

It is important to assure that these effects are not driven by some cross-correlation with other lag rainfall events. To test for such a possibility, I include further sets of lagged rainfall levels ($t+1$ to $t-29$) in Table II. This does not alter the estimates of the coefficients $t-17$ to $t-19$. None of the additional lags enter significantly at the ten-percent level. My estimates suggest that rainfall shocks in SSA affect the onset of civil conflict with an average lag of 17–19 years. The argument above relates this finding to a youth bulge effect, because the same lag induces fluctuations in the relevant male age cohort, which, as a consequence, induces fluctuations in the pool of PR. Thus, the data suggests that the full set 15–19 does not transit into the pool of PR, but that only—on average within this sample—the age group 17-19 does. Employing the rainfall levels $t-17$ allows me then to identify the part of the population 15–19 which is affected by this particular rainfall event.

Next, Table A.II in the Appendix depicts the results from regressions of the $t-15$ to $t-19$ rainfall levels on various contemporaneous covariates. The exercise shows that these lagged rainfall levels do not influence the risk of civil conflict through potential contemporaneous confounders—for example, dynamic effects of lower levels of investment during a low rainfall year fifteen years ago. The results highlight that potential contemporaneous rainfall effects on economic and social outcomes—which might be related to the risk of conflict—disperse fast enough such that they do not explain variation in these outcomes after 15 years. This is plausible, since there is no reason to believe that changes in variables like GDP per capita or (real) capital only affect the onset of civil conflict exactly with a lag time of 17–19 years, but have no immediate impact. By means of Figure A.II, this translates to the question why a real capital bulge should affect the risk of conflict only after seventeen years. In particular, the advancing age of real capital generally does not increase the risk of conflict.[17] In turn, fluctuations in birth cohort sizes matter—in my argument—only when this specific population cohort is aged 17–19 years, the age at which youth can be drafted into rebel armies. Furthermore, I cannot imagine any other factor that changes its risk attribute—with regard to conflict—with advancing age, except population cohorts.

Identification of the Youth Cohort Size

A fact that is important for this paper's argument is that contemporaneous rainfall acts exclusively on the size of the respective birth cohort and not on other cohort sizes. Under this assumption, birth year rainfall variation identifies the effects of their specific birth cohort sizes, but not population structure or total population variations.

To test this assumption, I include the log of the remaining population—that is total population minus males aged 15–19—in the setup in Table III.[18] Even though the coefficient sizes of the rainfall events decline, the overall picture remains the same. Most birth year rainfall lags in column 1 enter significantly at the five-percent level. Again, I cannot reject the hypothesis that all rainfall lags contribute in the same way to the cohort size variation.

To avoid a misspecification error through a potential bias in the remaining population coefficient, I employ a 2SLS framework in column 2. I use the average rainfall over ten years from $t-2$ to $t-11$ as instrument for the remaining population. This is possible, since the remaining population includes the age cohorts 2-11 which are highly correlated with their respective birth year rainfall ($t-2$ to $t-11$). The various rainfall instruments are orthogonal to each other (Table A.II). Furthermore, this ten-year rainfall average is uncorrelated with the age cohort 15–19 (Table A.III). Table III, column 2, shows that misspecification is not a problem, since the effects of the birth year rainfall events are comparable to column 1.[19]

In columns 3-7, I move the remaining population, preceding age cohort and subsequent age cohort to the left-hand side (LHS) of the regression equation. The estimates show that neither the remaining population, nor the preceding or subsequent age cohort sizes are related to the $t-15$ to

[17] The same holds true for infrastructure and the terms of trade.

[18] Using total population instead does not affect the result. However, employing the remaining population seems more intuitive.

[19] The loss in significance stems from the orthogonality of the respective instrument with the other population groups.

Table III: Identification of Male Youth Population

Dependent Variable:	Log men15–19$_t$		Log remaining population$_t$			Log pop. 20–24$_t$	Log pop. 10-14$_t$
	OLS (1)[a]	2SLS (2)[b]	OLS (3)	2SLS (4)[c]	2SLS (5)[d]	OLS (6)	OLS (7)
Log rain$_{t-15}$	0.023*	0.023*	0.007			-0.009	0.012
	(0.013)	(0.014)	(0.012)			(0.011)	(0.011)
Log rain$_{t-16}$	0.027**	0.027**	-0.006			-0.010	0.004
	(0.011)	(0.012)	(0.011)			(0.010)	(0.008)
Log rain$_{t-17}$	0.034***	0.034**	-0.000	-0.002	-0.001	-0.004	0.004
	(0.011)	(0.014)	(0.009)	(0.007)	(0.011)	(0.010)	(0.005)
Log rain$_{t-18}$	0.026**	0.026**	-0.008		-0.009	0.001	-0.001
	(0.013)	(0.013)	(0.007)		(0.010)	(0.009)	(0.005)
Log rain$_{t-19}$	0.006	0.006	0.017		0.016	0.005	0.001
	(0.014)	(0.021)	(0.019)		(0.022)	(0.008)	(0.005)
Log remaining population$_t$	0.440***	0.428					
	(0.092)	(0.497)					
Log men15–19$_t$			0.320***	0.359	0.349	0.584***	0.671***
			(0.069)	(0.282)	(0.375)	(0.107)	(0.089)
Obs.	1754	1754	1754	1754	1754	1754	1754
RMSE	0.059	0.059	0.051	0.051	0.051	0.056	0.051

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Remaining population is Total population − men15_19. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses. RMSE is the root mean square error.

[a] Coefficients of lag rainfall $t-15$ to $t-19$ are the same (p-value 0.14).

[b] 2SLS estimates with $\overline{\text{rain}}_{\{t-2,..,t-11\}}$ as first stage instrument for remaining population. Please see main text.

[c] 2SLS estimates: Log men15–19$_t$ instrumented by Log rainfall levels $t-15$, $t-16$, $t-18$ and $t-19$. The Hansen test statistic indicates that the instruments are valid (p-value 0.62).

[d] 2SLS estimates: Log men15–19$_t$ instrumented by Log rainfall levels $t-15$ and $t-16$. The Hansen test statistic indicates that the instruments are valid (p-value 0.30).

$t-19$ rainfall after controlling for males 15–19. Conditioning on this age cohort is necessary, since the age cohorts are correlated. The 2SLS estimates in column 5, where I instrument men 15–19 with some of the birth year rainfall lags, show that misspecification is not an issue. Table A.III, columns 1-2, in the Appendix depict the 2SLS estimates with similar results for the population aged 20-24 and 10-14. The table also shows that the birth year rainfall events from $t-15$ to $t-19$ have no significant effect on the cohort size of males aged 40–44.

The data strongly supports that it is possible to identify cohort sizes with the respective birth year rainfall events. That is, birth rainfall variation allows me to isolate and separate effects of a specific age cohort from effects of total population and other age cohorts (e.g., to separate youth bulge and mid-life moderation effects). Additional support for my identification strategy is presented in Section 8.

5.3 Instrumental Variable Approach

In this subsection, I present the 2SLS estimates and subsequently discuss economic and social conditions that hamper the youth bulge effect. Even though the main identification strategy in this section is through the lagged rainfall $t-17$, the results also hold for different instrument specifications (Section 8).

Youth Bulges and Civil Conflict

Table IV depicts the 2SLS results, where I instrument the size of the male youth cohort aged 15–19 with the $t-17$ lag rainfall. I report the robust standard errors clustered at the country and year level in parentheses (). For the second-stage estimates, I report the p-values of the Anderson-Rubin Chi-squared test statistic in the square brackets [] which are robust to weak instruments.[20] Since the first-stage F-statistic of $H_0 : \beta_{t-17} = 0$ is about 6.0, the Anderson-Rubin p-values are the appropriate choice (Andrews and Stock, 2005).[21] The 2SLS estimates and the Anderson-Rubin test statistic require that the model is identified. Hence, I also report the first-stage F-statistics of the excluded instruments.

Overall, the first-stage F-statistics indicates that the model is identified—with p-values generally smaller than five percent. The results in Table IV, column 1, depict a broad support for the youth bulge hypothesis. Increasing the size of the male cohort aged 15–19 by one percent induces an increase in the likelihood of the onset of civil conflict by 1.4 percentage points. I reason that rainfall at birth affects the birth cohort size. Hence, by affecting the size of the male youth cohort that transits into the pool of PR, the $t-17$ rainfall lag causes fluctuations in the likelihood of the onset of civil conflict. Since rainfall is exogenous, I conclude that youth bulges have a causal effect on the onset of civil conflict. This does not imply, however, that violence is nonexistent in later age groups. It merely states that this age group—by triggering fluctuations in the stock of potential recruits (PR)—makes rebel recruitment more feasible and, hence, induces changes in the likelihood of civil conflict. Intuitively, age cohorts cross the age-threshold to become recruitable into a rebel army only once.

I turn now to regression setups including additional covariates that are usually found in the relevant literature. Adding the future, contemporaneous and past log rainfall in columns 2 does not affect the size of the youth bulge coefficient or its significance. The rainfall levels all enter nonsignificant, which implies that contemporaneous rainfall does not appear to affect the onset of civil conflict in my sample, which also holds true for rainfall growth rates (not shown). Thus, the results of Miguel, Satyanath, and Sergenti (2004) do not translate directly to my setup. Also real

[20] The standard p-values tend towards *one* in the case of weak instruments (Andrews and Stock, 2005).

[21] Andrews and Stock (2005) consider an F-statistic smaller than 10.3 as evidence for a weak instrument.

Table IV: 2SLS Results- Youth Bulge and Civil Conflict

Dependent Variable:	Onset Civil Conflict$_t$						
	(1)	(2)	(3)^a	(4)	(5)	(6)	(7)^b
Log men15–19$_t$	1.366*** [0.005]	1.385*** [0.006]	1.477*** [0.004]	1.634** [0.014]	2.269*** [0.001]		1.192** [0.050]
Youth share$_t$						77.289*** [0.001]	
Log rain$_{t+1}$		0.007 (0.024)	0.000 (0.029)	-0.003 (0.031)	0.014 (0.032)	0.024 (0.047)	-0.009 (0.036)
Log rain$_t$		0.023 (0.028)		0.028 (0.034)	0.029 (0.040)	0.086 (0.057)	0.006 (0.033)
Log rain$_{t-1}$		0.020 (0.034)	0.007 (0.039)	-0.001 (0.036)	0.021 (0.036)	0.100** (0.039)	-0.017 (0.031)
Log GDP p. capita$_t$				-0.129 (0.111)	-0.183 (0.150)	-0.160 (0.150)	0.016 (0.026)
Log agricultural output p. capita$_t$			0.278 [0.413]				
ΔPolity IV$_t$					-0.008 (0.005)	-0.013** (0.006)	-0.009** (0.004)
Log real openness$_t$					0.031 (0.060)	0.102 (0.086)	-0.029 (0.038)
Urbanization$_t$					0.005 (0.007)	0.007 (0.014)	-0.006** (0.003)
	First stage regression: Log men15–19$_t$ (youth share in column 6)						
Log rain$_{t-17}$	0.040** (0.016)	0.040** (0.017)	0.038** (0.017)	0.035** (0.014)	0.031** (0.015)	0.001** (0.000)	0.044** (0.019)
	First stage regression: Log agricultural output p. capita$_t$						
Log rain$_t$			0.074** (0.031)				
Obs.	1754	1754	1639	1523	1523	1523	1523
RMSE	0.222	0.223	0.224	0.231	0.245	0.306	0.235
F-test excl. IV	6.033	5.871	5.58/4.35	6.261	4.169	4.722	5.446

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 2SLS regression with log rain$_{t-17}$ as an instrument for log men15–19 (youth share in column 6) in the first stage. FE estimator regressions in all columns with country dummies, country-specific time trends (except column 6), time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket[] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. The F-test is the first-stage test statistic of the excluded instrument. RMSE is the root mean square error. Youth Share is $\frac{\text{men15-19}}{\text{total population}}$. Agricultural output is measured as the net production value index per capita from the FAO. Δ Polity IV is the change of the Polity2 score from the Polity IV database. Real openness is from the PWT. Urbanization is the share of urban population in total.

[a]In column 3, 2SLS estimates with log rain$_t$ as instrument for agricultural output and log rain$_{t-17}$ as instrument for men15–19 in the first stage. F-test excl. IV in column 3 first gives the value for log rain$_t$ and then log rain$_{t-17}$.

[b]In column 7, I drop the country-specific time trends.

GDP per capita is not significantly related to contemporaneous rainfall in my sample.

However, column 3 (first-stage) shows that agricultural production value per capita is positively and significantly related to contemporaneous rainfall. This suggests that the noise from other sectors, like manufacturing and services, override the signal of the rainfall variation—through agricultural output—in the aggregated GDP per capita. It also highlights the link between rainfall and the agricultural sector in rural rainfed SSA which is important to my argument of birth year rainfall variation, agricultural income and infant mortality. Column 3 also shows the second-stage estimates, where I use the contemporaneous rainfall as an instrument for agricultural output and the $t-17$ rainfall lag as an instrument for the youth measure. The coefficient of agricultural output is nonsignificant, but the youth bulge effect enters again significant and negative with a value of 1.5.

In column 4, I add real GDP per capita which enters nonsignificant. In the regression reported in columns 5, I also include a democracy measure, urbanization and real openness to trade. The variables do not affect the outcome of the youth bulge effect, but even increase its coefficient, owing partly to a sample size reduction.

To show that the results do not depend on a particular specification, I employ in column 6 the youth share—male population aged 15–19 relative to total population—as LHS variable. Here the youth bulge coefficient again enters significantly and positively as before. However, it is important to stress that identification takes place through the male population aged 15–19. Thus, in this case, I hold total population constant—conditional on the male youth population—as discussed in Subsection 5.2. A one-percentage point increase yields an increase in the risk of civil conflict by 0.77 percentage points.[22] In column 7, I drop the country-specific time trends, which does not alter the size or significance of the youth bulge coefficient.

Next, I present additional robustness checks in Appendix A.I. In column 3, I exclude the top and bottom percentile of the lagged rainfall data, to show that the results are not driven by outliers. The estimate of the youth bulge coefficient is of similar magnitude and remains significantly at the five-percent level. Since both female and male cohort sizes are affected by birth rainfall variation, the $t-17$ lag rainfall also induces fluctuations in youth female cohorts aged 15–19. Hence, the true youth bulge effect might stem from young women. However, the literature clearly rejects such a hypothesis, as do the results in column 4. To show that the results do not rely on a particular onset specification, Table A.I, columns 5-8, show estimates, where I use the onset definitions of Miguel, Satyanath, and Sergenti (2004) and Fearon and Laitin (2003), respectively. The coefficients are in size and significance similar to the base results above, which highlights the robustness of the estimates.

Finally, Table A.IV shows the result with the full control variable specification—including the subsequent age cohort size. Even though the quantitative and qualitative effect of male youth on civil conflict does not change, the results must be interpreted with care. This is because some of the additional control variables might induce a misspecification error.[23]

[22]Note that this does not represent elasticity though, as compared to the previous columns in Table IV.

[23]It has also to be noted, that population density and total population are equivalent in the country-fixed effects

Table V: 2SLS Results - Youth Bulge and Interactions

Dependent Variable:	Onset Civil Conflict$_t$						
	(1)		(2)		(3)		(4)
Log m15–19$_t$	4.709*** [0.008]		4.629** [0.010]		4.004*** [0.002]		1.883*** [0.005]
Log rain$_t$	2.479*** (0.876)	Log agricul. output$_t$	3.955* (2.186)	Log foreign aid$_t$	0.561* (0.303)	Urbanization$_t$	0.067* (0.033)
Log m15–19 x log rain$_t$	-0.507*** (0.171)	Log m15-19 x agriculture$_t$	-0.743* (0.408)	Log m15-19 x log aid$_t$	-0.107* (0.058)	Log m15-19 x urbanization$_t$	-0.016* (0.008)
Obs.	1754		1639		1666		1754
RMSE	0.259		0.249		0.272		0.230
F-test excl. IV	20.417		4.511		3.907		7.058

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. m15–19 is the abbreviation for men15–19. 2SLS estimates with log rain$_{t-17}$ as instrument for men15-19 in the first stage. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket [] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. Agricultural output is measured as the net production value index per capita from the FAO. Urbanization is the share of the urban population in total. The F-test is the first-stage test statistic of the excluded instrument. RMSE is the root mean square error.

Youth Bulge Interactions

In this subsection, I turn to the question of which socioeconomic and political conditions mitigate or facilitate the impact of youth bulges on the onset of civil conflict. Since cohort sizes are more or less fixed over time—except in the case of rare events like famine[24]—such an offsetting effect must work through the cohort-specific marginalization rate (Eq. (2)).

In particular, the discussion in Section 2 suggests that insurgencies draw mostly on marginalized males in the rural hinterland—which depends heavily on rainfed agriculture. Hence, it is to be expected that young men facing better economic conditions are less likely to partake in violent behavior and are more difficult to recruit into rebel armies. Thus, in years of good economic conditions the influx of an exceptionally large cohort, $(X(T_{in}))$, might be offset by a decline in the marginalization rate which leads to an overall smaller pool of PR, despite this youth bulge.

Using contemporaneous rainfall as a proxy for agricultural output allows me to derive estimates which are free of endogeneity problems, since both, $t - 17$ and t, rainfall events are exogenous to civil conflict and are orthogonal to each other. Table V, column 1, shows that both contemporaneous rainfall and its interaction term with male youth aged 15–19 enter highly significantly. The youth bulge effect now increases to 4.7, and the size of the interaction term coefficient is -0.5 and negative. Thus, the data suggest that increasing income in rural agricultural areas hampers the risk of civil conflict significantly. This suggests an opportunity cost effect in the context of youth

specification as long as country area does not change.

[24] However, such exceptional events such as heavy droughts resulting in famine usually affect all population classes (Brückner, 2010)

bulges. In years with better economic perspectives, male youth are less likely to join insurgencies. This mitigating effect, however, is not limited to the youth population, since better economic conditions most likely reduce the marginalization rate across all age groups within the pool of PR—age 17–44. On the other hand, lifting SSA from rainfed agricultural dependence might help to reduce the risk of negative economic shocks and hence reduce the impact of youth bulges.

Furthermore, despite the positive sign of the contemporaneous rainfall coefficient, the total rainfall effect—evaluated at the mean—is negative, which implies a lower risk of civil conflict in years with better agricultural conditions. For example, increasing contemporaneous rainfall by one percent reduces the overall risk of conflict by about 0.8 percentage points. This corroborates the findings of Miguel, Satyanath, and Sergenti (2004) and implies that economic conditions (in the context of civil conflict) and youth population size are co-dependent.

Column 2 shows the results for a more direct measure of agricultural production, by using a per capita agricultural output (net) index from the FAO. The index and the interaction term enter significantly at the ten-percent level and yield similar results as before.

Column 3 includes log foreign aid received by each country and its interaction term with the male youth population. This latter coefficient is negative and significant at the ten-percent level. This fits into the argument outlined just yet. Increasing foreign aid at least partly helps to lift youth population out of the extreme poverty that is still found in rural areas—often dominated by subsistence-oriented agriculture. Thus, as the population becomes less vulnerable economically, recruitment costs increase and hence help to offset the risk of conflict associated with youth bulges. I conclude that, generally, factors that reduce poverty and improve economic perspectives reduce the risk of civil conflict.

In turn, evaluating the aid effect with regard to its own mean and the mean of men aged 15–19 yields a positive net impact. Thus, even though foreign aid hampers the youth bulge risk, its overall effect increases the risk of civil conflict. This is in line with the findings of Nunn and Qian (2012) who show that United-States food aid increases the risk of conflict, as looted food aid can boost insurgencies. However, my findings indicate a more subtle relationship between foreign aid and civil conflict, since aid also improves the economic perspectives of the youth population, which in turn reduces the likelihood of conflict.

Next, I turn to urbanization. Here, both the share of urban population and its interaction term with log men 15–19 enter significantly at the ten-percent level. The results show that the youth bulge effect diminishes as the share of the urbanized population increases. Thus, high urbanization might act as a protection against civil conflicts induced by youth bulges. This seems plausible, since insurgencies in SSA are predominantly rural-based, owing to the lack of government control in the hinterland (Herbst, 2000).[25] Hence, higher urbanization leads to the geographical separation

[25]This is mirrored in the circumstance that insurgencies usually draw their followers from areas outside urban

between potential recruits (urban) and insurgency hideaway (rural). Furthermore, more than 70 percent of the youths in SSA live in rural areas where the most severe poverty is located and where the youth face poor economic prospects and deprivation (World Bank, 2009; Ravallion, Chen, and Sangraula, 2007; Gordon et al., 2003; United Nations, 2005).[26] Thus, urbanization makes it more difficult for insurgency leaders to recruit, since urbanization counters extreme deprivation.

The significant interaction terms do imply another important finding: Onsets of civil conflict are more likely when several adverse factors coincide; that is, exceptionally large youth cohorts or low income years are per se not a good predictor of conflict. However, when such events occur together—e.g., low incomes and a large youth population—the likelihood of civil conflict increases.

Finally, the data does not support the idea that better democratic conditions or more openness to trade hamper the youth bulge effect (regressions not shown). I also do not find significant interaction effects between youth bulges and GDP per capita, which can be attributed to the fact that the latter does not explicitly capture agricultural income (regressions not shown).

Youth Bulge and Youth Dent

Following the argument in Section 2, positive shocks in the male youth-population size leads to an excess supply of labor and marginalization. In this subsection, I shortly discuss whether only positive shocks in the size of the male youth population (youth bulge) drive the risk of conflict or whether negative shocks in the size of the male youth population (youth dents) might also *reduce* the risk of conflict. Table A.V in the Appendix presents the results where the sample is splitted according to the country-respective mean values of the male youth population. Since the sample-split leads to a weaker first-stage results—owing to variation loss—I drop in some specifications the country-specific time trends, as this allows me to derive more precise first-stage estimates.

The results in Table A.V show that youth dents do not contribute towards explaning the pattern of civil conflict—in particular, they do not reduce the risk of conflict occurring. In turn, the results above are driven by youth bulges which increase the risk of civil conflict. That is in line with the theoretical considerations that indicate that exceptionally large youth cohorts—which cannot be absorbed by the prevailing economic system—lead to excess labor and marginalization. This, in turn, results in a higher risk of conflict. Factors that increase economic perspectives—such as higher agricultural income—reduce marginalization which hampers the risk of civil conflict attributable to youth bulges. The finding further suggests that the marginalization effect—through the bottleneck effect on the labor market—dominates the population-size effect in the context of youth bulges.

developments (Sommers, 2003).

[26] Among others, malnutrition is more acute in rural areas (Uthman and Aremu, 2008; Gordon et al., 2003).

Table VI: Onset Civil Conflict and Rainfall Lags - Moderation

Rainfall lags:	t-40	t-41	t-42	t-43	t-44	t-45	t-46	t-47	t-48	t-49	t-50
Onset conflict$_t$	-0.010 (0.032)	0.029 (0.035)	0.012 (0.012)	-0.047* (0.027)	-0.048* (0.024)	0.032 (0.039)	-0.005 (0.026)	-0.049 (0.029)	0.028 (0.027)	-0.017 (0.034)	-0.009 (0.038)

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table shows the regression results of the onset of civil conflict on the rainfall lags $t-40$ to $t-50$. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses. 1,754 Observations.

5.4 Mid-Life Moderation and Conflict

The theoretical argument above relates large cohorts in their mid-40s—which transit out of the pool of PR—to a decline in the risk of civil conflict. In particular, when an exceptionally large cohort exits the pool of PR, the size of this pool shrinks, which in turn reduces the risk of conflict. Assuming that birth year rainfall drives cohort size, it is to be expected that lagged rainfall variation in a window of $t-40$ to $t-50$ negatively affects conflict. Table VI shows the estimates of lagged rainfall $t-40$ to $t-50$ on the risk of civil conflict. In fact, the rainfall lags $t-43$ and $t-44$ display negative regression coefficients that border the five-percent significance-level with a joined F-test statistic of 4.8.

Table A.VI, columns 1-2, show that these two lagged rainfall levels are uncorrelated with males aged 15–19, but significantly positively correlated with the population size of males aged 40–44. This result is robust to the inclusion of various rainfall levels in column 3. Consequently, by using the lagged rainfall levels as instruments, the estimates of the 2SLS setup yield a significant and negative effect of the cohort size of males aged 40–44 on the risk of civil conflict (column 5) with a first-stage F-test statistic of 5.31. The point estimates imply a decline in the likelihood of conflict by 0.8 percentage points when the population aged 40–44 increases by one percent. This moderation effect that occurs when large cohorts exiting the pool of PR is smaller than its reversed youth bulge effect—where large cohorts enter the PR pool. The reason might be a decline in the cohort-specific marginalization rate over time—e.g., better employment conditions through human capital accumulation.

Column 6 shows the results when I include both, males aged 15–19 and 40–44. Importantly, the birth year rainfall variation for both cohorts identifies their respective cohorts. The results are comparable to column 5 and Table IV, column 1. Comparing the coefficients shows that the youth bulge effects dominates the moderation effect.

6 Youth Bulges and Low-Level Social Violence

In this short section, I present evidence that youth bulges also induce low-level social conflict. This section, then, serves two purposes: Youth bulges trigger different levels of violence and conflict in

SSA; the relationship between the birth-year rainfall levels of males aged 15–19 and the onset of conflict is not limited to the PRIO database.

Because of the loss in variation, owing to the severe reduction along the time dimension—a loss of 30 years—I am not able to control for country-specific time trends in this setup.[27] Table A.VII in the Appendix shows the results.[28] Column 1 depicts—at least at the ten-percent level—significant coefficients of the $t-14$ and $t-17$ lag rainfall levels. However, I cannot reject the hypothesis that the rainfall coefficients $t-14$ to $t-17$ contribute in an equivalent way to the explanation of low-level social violence (p-value 0.92). Here then, the critical age is confined to 14–17 years. This indicates that the respective youth cohorts, $X(T_{in})$, who influence the onset of low-level civil violence—such as demonstrations and riots—exhibit a somewhat younger age structure. Hence, exceptionally large youth cohorts contribute to the onset of social violence at an earlier age than in the case of full-fledged civil conflicts.

Using the birth year rainfall level $t-17$ as instrument allows me to evaluate the youth bulge effect on the onset of low-level social violence in column 3. The result is qualitatively similar to the case of civil conflict, but larger. Increasing the male youth population aged 15–19 by one percent increases the risk of social conflict by about 4 percentage points. This implies that youth bulges translate more quickly into low-level violence than into civil conflicts. Hence, the data suggests that the threshold for joining riots might be lower than the threshold for joining insurgencies. Another interpretation might be that the organization of civil conflicts is not always feasible even though a sufficiently large pool of PR exists. That again would highlight the fact that several driving factors have to emerge simultaneously in order to ignite such a rare event as a civil conflict.

In turn, the younger age structure cannot be attributed to the sample restriction towards a later time period, since column 4 still associates the onset of civil conflict with the rainfall levels $t-17$ and $t-19$.

7 Contemporaneous Rainfall and Infant Mortality

The proposition that contemporaneous rainfall affects the birth cohort size is a central proposition of this paper. Existing literature relates rural income fluctuations in the birth year to infant mortality (Devereux, 2009; Baird, Friedman, and Schady, 2011; Rice et al., 2000).

Table VII confirms these findings by linking agricultural output to infant mortality and the number of surviving infants. I employ contemporaneous rainfall as an instrument for the agricultural output. The F-statistic of the excluded instrument indicates that the identification assumption

[27] This does not appear to be a structural problem, since the reduced form effect persists.

[28] I do not present the results of the alternative rainfall levels, in order to save space. The PREC/L and Delaware rainfall data estimates yield quantitatively and qualitatively the same results. (Tables are available on request from the author.)

Table VII: 2SLS Results - Rainfall, Birth Rate and Infant Mortality until 1995

Dependent Variable:	Infant Mortality Rate	Log Infant Survivals	Log Men0–1		
	(1)	(2)	(3)		(4)
Log agricultural output$_t$	-0.014**	0.324**	0.299***	Log rain$_t$	0.027**
	[0.042]	[0.011]	[0.005]		(0.010)
Obs.	1063	1063	1063		1063
RMSE	0.005	0.052	0.065		0.061
F-test excl. IV	4.850	4.850	4.850		

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 2SLS estimates with log rain$_t$ as instrument for agricultural output in the first stage. In all columns, rain$_t + 2$ to rain$_t - 2$ are included which enter nonsignificant at the ten-percent level. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country level in parentheses (). The square bracket[] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. The F-test is the first stage test statistic of the excluded instrument. RMSE is the root mean square error. Infant mortality rate is the number of infant deaths per 1,000 live births. Mortality to live birth rate is the share of absolute infant mortalities to absolute live births. Infant survival rate is the difference between the absolute number of live births and infant mortality. Agricultural output is measured as the net production value index per capita from the FAO.

is satisfied. In particular, in the relevant period until 1995, rainfall is positively and significantly related to agricultural output.[29]

Column 1 reports a coefficient of -0.014, which implies a decrease in the infant mortality rate (mortalities per 1,000 live births) by 1.4 percentage points, if agricultural output per capita increases by one percent. In turn, column 2 depicts the elasticity between agricultural output and the number of surviving infants—live births minus mortalities in a given year. Here the results imply a 0.3 percent increase in the number of surviving infants for each percent increase in agricultural output.

In addition, columns 3 and 4 show the point estimates when the log of males aged 0 to 1 years is the dependent variable. Consistent with the argument that rainfall induces fluctuations in cohort sizes through agricultural income, the results indicate a positive and significant relationship between contemporaneous rainfall, agricultural output and the respective age cohort. Hence, Table VII supports the argument that rainfall affects the birth cohort sizes by governing the infant mortality rate through agricultural income.

8 Consistency Tests

In this section, I discuss different rainfall datasets and present additional evidence on the validity of my instrumental variable strategy in the context of youth bulges and civil conflict.

[29] 1995 is the last relevant birth cohort in my sample, since I look at male youth aged 15–19.

Full Set of Birth Year Rainfall Events and Alternative Rainfall Data

I now turn to the different instrument specifications in Table A.VIII, to ensure that the results do not depend on a particular rainfall lag or rainfall dataset. In all columns, I control for $t+1$, t, and $t-1$ rainfall levels, which enter nonsignificantly (coefficients not reported). In columns 3, 4, 6, and 8, I do report clustered standard errors at the country level, instead of the usual country and year level.[30] First, I concentrate on the rainfall levels from $t-17$ to $t-19$ which all contribute in a similar way to the explanation of civil conflict according to Section 5.2. That is, the data places the relevant age of transition into the pool of PR, $X(T_{in})$, within the age group 17–19. Columns 1 and 2 report the results for the average $t-17$ to $t-19$ rainfall lag and the individual lags as instruments, respectively. The second-stage estimates yield quantitatively and qualitatively equivalent results as before with a first-stage F-test statistic bordering the five-percent level (column 1) or better (column 2).[31] Furthermore, the Hansen-test statistic indicates that the instruments are valid.

To show that the result does not depend on the inclusion of the $t-17$ lag, column 3 depicts the result with the $t-18$ and $t-19$ rainfall lags as instruments. Here, the first-stage F-test statistic is somewhat weak. Importantly, however, the second-stage yields coefficient sizes and significance values which are close to the base results in Table IV. Thus, all rainfall lags—that affect civil conflict and male population (17 to 19)—consistently suggest similar sized youth bulge effects. There is another point to stress here. By using these rainfall lags, I identify only the part (the population aged 17–19) of the quinquennial cohort (15 to 19) that contributes significantly to the onset of conflict—the population cohorts that transit into the pool of PR. Thus, using the full set of birth year rainfall levels in column 4 yields strong first-stage estimates, but the coefficient of the youth bulge becomes smaller. This is because my instruments now also identify the younger cohorts—included in the quinquennial cohort—and which are not related to civil conflict: the age group 15–16. Consequently, by using the additional birth year rainfall events as instruments leads to a drop of the Hansen-test statistic. Nevertheless, the coefficients cannot be distinguished in a statistical sense.

Using the average rainfall from $t-17$ to $t-19$, or the $t-17$ rainfall event from the Delaware rainfall data, or the PREC/L dataset in columns 5-8, yields again basically identical youth bulge effects and similar first-stage results.[32] However, using the PREC/L datasets implies a loss of seven years of observations.

Given these results, I conclude that the findings do not rely on a particular dataset or on a particular rainfall level.

[30]This is because the number of instruments relative to cluster sizes does not allow me to calculate the robust covariance. As a result, the first-stage F-test yields implausibly high values.

[31]Note that for some setups the test statistic is based on (5,40) degrees of freedom.

[32]The first-stage F-test statistics are all below the five-percent levels.

Table VIII: Instrument Test

Dependent Variable:	Log Rain$_{t-13}$	
	OLS	IV
	(1)	(2)[a]
Log men15–19$_t$	0.184**	-0.442
	(0.077)	[0.590]
Obs.	1754	1754

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. FE estimator regressions in all columns with country dummies, country-specific time-trends and time dummies and robust standard errors clustered at the country level in parentheses (). The square bracket[] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments.

[a] In column 2, the instrument is the $t-17$ rainfall level.

Instrumental Variable Discussion

Section 5.2 highlights a significant relationship between lagged $t-17$ rainfall and the size of the male cohort aged 15–19, and the preceding subsection shows a significant relationship between contemporaneous rainfall and infant mortality. The findings confirm that birth year rainfall is a plausible instrument for cohort size. For further tests of the validity of the instrument, I consider additional properties of the instrumental variable. Table A.II shows the bivariate correlation between the lag $t-17$ rainfall and the control variables included in my study. The respective p-values are all insignificant, which gives some indication that the instrument is orthogonal—for these given confounders—to the error term of the regression equation. Further evidence is given by the fact that a direct significant link exists between lag $t-17$ rainfall and the onset of civil conflict, which also indicates orthogonality.

Next, I perform a falsification test following von Hinke Kessler Scholder et al. (2013). The test requires that the 2SLS estimates are capable of separating causal effects from naive correlation. Since birth year $t-13$ rainfall is a significant driver of the cohort size aged 10 to 14, and due to the between cohort correlation, the OLS estimates in Table VIII, column 1, show a positive and significant relationship between the lag $t-13$ rainfall and the youth cohort size aged 15–19.[33] This effect, however, is not causal, but stems from the omitted confounder (cohort aged 10 to 14). In turn, the 2SLS result shows a severe decline in the size of the coefficient and a fairly nonsignificant relationship. These findings indicate that the instrument performs well.

[33] In fact, each cohort up until age 15 to 19 can be associated with a specific and significant birth year rainfall level.

9 Conclusion

Youth bulges (exceptional surges in the male youth population) are identified as a major correlate of civil conflicts, both in theoretical and empirical work across the sciences. Evidence of a causal relationship is still missing, due to omitted variable bias and endogeneity. To address these issues, I employ rainfall variation in the birth years of the respective age cohorts as an instrument for their size.

In the case of SSA, my estimates reveal that the effect of the male youth population aged 15–19 on the risk of civil conflict occuring has a substantial effect. Increasing this population group by one percent induces an increase in the risk of 1.4 percentage points. The size of the youth bulge effect depends on the prevailing economic conditions. Better economic perspectives in rural SSA—in particular, high contemporaneous rainfall levels and, as consequence, higher agricultural income—hamper the risk of civil conflict attributable to youth bulges. These findings, then, suggest that the emergence of civil conflict is more likely in years when several adverse factors coincide; e.g., a large youth population combined with low agricultural output. The data, further, suggest that the rate of urbanization mitigates the effect of youth bulges on the onset of civil conflict. In order to replicate my results, I show that youth bulges also increase the risk of low-level social violence in SSA. The estimates here show an even stronger effect. Increasing male youth population aged 15–19 by one percent increases the risk of the onset of low-level social conflict by 4 percentage points.

I present evidence that exceptionally large male cohorts in their mid-40s trigger a decline in the risk of conflict. Increasing this population group by one percent results in a decline in the risk of conflict by 0.8 percentage points.

I argue that infant mortality rates governed the cohort sizes in SSA during the sample period. In turn, exogenous shocks (rainfall) induce fluctuations in infant mortality rates—and hence cohort sizes—through available rural household income in these rain-fed farming systems. Youth bulges lead to changes in the likelihood of civil conflict by increasing the pool of potential recruits. That is, when cohorts reach the critical age where they can be drafted into rebel armies, large youth cohorts shift the size of the equilibrium pool of potential recruits upwards. This takes place along an extensive (cohort size) and intensive (marginalization) margin. The latter is attributed to a bottleneck effect on the labor market. Similarly, when large cohorts exit the pool of PR when age 40–45, the likelihood of conflict declines again.

I back the plausibility of my instrument by identifying contemporaneous rainfall as a driving factor of agricultural output and infant mortality in SSA. Instrumenting agricultural output with contemporaneous rainfall shows that agricultural output is closely related to infant mortality in this region.

References

Andrews, Donald W.K. and James H. Stock. 2005. "Inference with Weak Instruments." *NBER Technical Working Papers 0313* .

Baird, Sarah, Jed Friedman, and Norbert Schady. 2011. "Aggregate Income Shocks and Infant Mortality in the Developing World." *The Review of Economics and Statistics* 93(3):847–856.

Bhalotra, Sonia. 2010. "Fatal fluctuations: Cyclicality in infant mortality in India." *Journal of Development Economics* 93(1):7–19.

Blattman, Christopher and Edward Miguel. 2010. "Civil War." *Journal of Economic Literature* 48(1):3–57.

Brückner, Markus. 2010. "Population Size and Civil Conflict Risk: Is there a Causal Link?" *Economic Journal* 120(544):535–550.

Burke, Marshall B., Edward Miguel, Shanker Satyanathd, John A. Dykemae, and David B. Lobell. 2009. "Warming increases the risk of civil war in Africa." *Proceedings of the National Academy of Sciences USA* 106:20670–20674.

Chandrasekhar, C. P., Jayati Ghosh, and Anamitra Roychowdhury. 2006. "The 'Demographic Dividend' and Young India's Economic Future." *Economic and Political Weekly* 41(49):5055–5064.

Chen, Mingyue, Pingping Xie, John E. Janowiak, and Phillip A. Arkin. 2002. "Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations." *Journal of Hydrometeorology* 3(3):249–266.

Ciccone, Antonio. 2011. "Economic Shocks and Civil Conflict: A Comment." *American Economic Journal: Applied Economics* 3(4):215–27.

Collier, Paul and Anke Hoeffler. 2002. "On the Incidence of Civil War in Africa." *Journal of Conflict Resolution* 46(1):13–28.

———. 2004. "Greed and grievance in civil war." *Oxford Economic Papers* 56(4):563–595.

Collier, Paul, Anke Hoeffler, and Dominic Rohner. 2009. "Beyond greed and grievance: Feasibility and civil war." *Oxford Economic Papers* 61(1):1–27.

Coulibaly, Ibrahima. 2004. "The impact of HIV/AIDS on the labour force in Sub-Saharan Africa: A preliminary assessment." *International Labour Organization: Research and policy analysis* .

CTA. 2012. "Increasing rural employment in sub-Saharan Africa." *CTA Policy Brief* 4.

Cunningham, David E., Kristian Skrede Gleditsch, and Idean Salehyan. 2009. "It Takes Two. A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution* 53(4):570–597.

Devereux, Stephen. 2009. "Seasonality and Social Protection in Africa." *FAC Working Paper No. SP07* .

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Garcia, Marito and Jean Fares, editors. 2008. *Youth in Africa´s Labor Market*. The International Bank for Reconstruction and Development / The World Bank.

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.

Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World*. University of California Press.

———. 2010. "The New Population Bomb." *Foreign Affairs* URL `http://www.foreignaffairs.com/articles/65735/jack-a-goldstone/the-new-population-bomb`.

Gordon, David, Shailen Nandy, Christina Pantazis, Simon Pemberton, and Peter Townsend. 2003. *Child poverty in the developing world*. The Policy Press.

Hendrix, Cullen S. and Idean Salehyan. 2013. "Social Conflict in Africa Database (SCAD)." *www.scaddata.org* .

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press.

———. 2004. "African Militaries and Rebellion: The Political Economy of Threat and Combat Effectiveness." *Journal of Peace Research* 41(3):357–369.

Hsiang, Solomon M., Kyle C. Meng, and Mark A. Cane. 2011. "Civil conflicts are associated with the global climate." *Nature* 476:438–441.

Matsuura, Kenji and Cort J. Willmott. 2012. *Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (V 3.01)*. Center for Climatic Research, Department of Geography, University of Delaware. URL `http://climate.geog.udel.edu/~climate/html_pages/archive.html`.

Miguel, Edward and Shanker Satyanath. 2011. "Re-examining Economic Shocks and Civil Conflict." *American Economic Journal: Applied Economics* 3(4):228–232.

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.

Murray, C J L, G King, A D Lopez, N Tomijima, and E G Krug. 2002. "Armed conflict as a public health problem." *British Medical Journal* 324:346–349.

Nunn, Nathan and Nancy Qian. 2012. "Aiding Conflict: The Impact of U.S. Food Aid on Civil War." *NBER Working Papers 17794, National Bureau of Economic Research, Inc.* .

Okumu, Wafula and Augustine Ikelegbe, editors. 2010. *Militias, Rebels and Islamist Militants, Human Insecurity and State Crisis*. Institute for Security Studies.

Ravallion, Martin, Shaohua Chen, and Prem Sangraula. 2007. "New Evidence on the Urbanization of Global Poverty." *Population and Development Review* 33(4):667–701.

Rice, Amy L., Lisa Sacco, Adnan Hyder, and Robert E. Black. 2000. "Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries." *Bulletin of the World Health Organization* 78:1207–1221.

Schneider, Udo, Andreas Becker, Peter Finger, Anja Meyer-Christoffer, Bruno Rudolf, and Markus Ziese. 2011. "GPCC Full Data Reanalysis Version 6.0 at 1.0°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data." .

Schneider, Udo, Andreas Becker, Peter Finger, Anja Meyer-Christoffer, Markus Ziese, and Bruno Rudolf. 2013. "GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle." *Theoretical and Applied Climatology* :1–26.

Smith, Stephen W. 2011. "Youth in Africa: Rebels without a Cause but not without Effect." *SAIS Review* 31(2):97–110.

Sommers, Marc. 2003. "War, Urbanization and Africa´s Youth at Risk: Understanding and Addressing Future Challenges." *Basic Education and Policy Support (BEPS) Activity and Creative Associates International* .

Themnér, Lotta and Peter Wallensteen. 2012. “Armed Conflicts, 1946-2011.” *Journal of Peace Research* 49:565–575.

UNICEF. 2012. “Levels & Trends in Child Mortality.” URL http://www.unicef.org/videoaudio/PDFs/UNICEF_2012_child_mortality_for_web_0904.pdf.

United Nations. 2003. “The impact of HIV/AIDS on mortality.” *Workshop on HIV/AIDS and Adult Mortality in Developing Countries. UN/POP/MORT/2003/14* .

———. 2005. “World Youth Report 2005: Young People Today, and in 2015.” URL http://social.un.org/index/WorldYouthReport/2005.aspx.

Urdal, Hendrik. 2006. “A Clash of Generations? Youth Bulges and Political Violence.” *International Studies Quarterly* 50(3):607–629.

———. 2012. “A clash of generations? Youth bulges and political violence.” *United Nations Department of Economic and Social Affairs, Population Division Expert Papers Series* .

Uthman, OA and O Aremu. 2008. “Malnutrition among women in sub-Saharan Africa: Rural-urban disparity.” *Rural Remote Health* 8(2).

von Hinke Kessler Scholder, Stephanie, George Davey Smith, Debbie A. Lawlor, Carol Propper, and Frank Windmeijer. 2013. “Child height, health and human capital: Evidence using genetic markers.” *European Economic Review* 57(C):1–22.

World Bank. 2008. “World Development Report 2008: Agriculture for Development.” URL http://go.worldbank.org/2DNNMCBGI0.

———. 2009. “Youth and Unemployment in Africa: The Potential, the Problem, the Promise.” Online. URL http://siteresources.worldbank.org/EXTSTATINAFR/Resources/ADI-200809-essay-EN.pdf.

———. 2011. “World Development Report 2011: Conflict, Security, and Development.”

Appendices

A Tables and Figures

Table A.I: Robustness - Dynamics and Different Onset Specifications

Dependent Variable:	Onset Civil Conflict$_t$				Onset(B)^c Civil Conflict$_t$		Onset(C)^d Civil Conflict$_t$	
	OLS (1)	sysGMM (2)	2SLS (3)^a	2SLS (4)^b	OLS (5)	2SLS (6)	OLS (7)	2SLS (8)
Onset Civil Conflict$_{t-1}$	-0.186*** (0.037)	-0.124** (0.051)						
Log rain$_{t-17}$	0.054** (0.022)	0.046** (0.020)			0.052** (0.022)		0.051*** (0.019)	
Log men15–19$_t$			1.613** [0.012]	5.471*** [0.010]		1.467*** [0.010]		1.389*** [0.005]
Obs.	1669	1669	1721	1754	1678	1678	2050	2050
RMSE	0.201		0.228	0.229	0.206	0.216	0.198	0.210
F-test excl. IV			5.022	2.036		5.960		5.291

Note: ** $p < 0.05$, *** $p < 0.01$. sysGMM refers to system GMM estimates. FE estimator regressions in all columns with state dummies, time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket[] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. RMSE is the root mean square error. The F-test is the first-stage test statistic of the excluded instrument.

[a] In column 3, I dropped the upper and lower percentile of the $t - 17$ rainfall data, as described in Section 5.1.
[d] In column 4, I partial out the log female15-19 population.
[c] Onset(B) is coded following Miguel, Satyanath, and Sergenti (2004), where the onset of peace is coded as missing.
[d] Onset (C) is coded following Fearon and Laitin (2003), where all non-conflict onset observations are coded as 0.

Table A.II: Correlation Coefficients - Instrument and Covariates

Dependent Variable (t):	Log GDP p. Capita (1)	Urbanization (2)	Democracy (Polity IV) (3)	Log Rain (4)	Log Foreign Aid (5)	Log $\overline{\text{rain}}_{\{t-2,..,t-11\}}$ (6)	Log Capital Stock (7)
Log rain$_{t-15}$	0.044	-0.238	-0.486	0.040	-0.019	-0.030	0.012
	(0.042)	(0.417)	(0.345)	(0.048)	(0.124)	(0.018)	(0.020)
Log rain$_{t-16}$	0.037	-0.041	0.062	-0.033	0.133	-0.039	0.014
	(0.048)	(0.420)	(0.260)	(0.041)	(0.145)	(0.016)	(0.020)
Log rain$_{t-17}$	0.048	0.210	0.229	-0.012	0.102	-0.038	0.017
	(0.043)	(0.437)	(0.272)	(0.026)	(0.117)	(0.016)	(0.017)
Log rain$_{t-18}$	0.033	0.449	-0.261	-0.067	0.255	-0.036	0.011
	(0.036)	(0.353)	(0.540)	(0.049)	(0.176)	(0.019)	(0.019)
Log rain$_{t-19}$	-0.009	0.451	0.089	0.063	0.515	-0.049	-0.008
	(0.034)	(0.365)	(0.345)	(0.038)	(0.264)	(0.023)	(0.022)
Obs.	1677	1754	1546	1754	1666	1754	1071

Note: * $p < 0.10$, ** $p < 0.05$. FE estimator regressions in all columns with country dummies, country-specific time trends and time dummies and robust standard errors clustered at the country and year level in parentheses (). Urbanization is the share of urban population in total. Foreign Aid is the total aid received in current US dollars. Capital stock is the real capital stock of the agricultural sector. $\overline{\text{rain}}_{\{t-2,..,t-11\}}$ is the average rainfall over a ten-year period from $t-2$ to $t-11$.

Table A.III: Additional Results-Identification Strategy

Dependent Variable:	Log men 15–19$_t$	Log pop. 20–24$_t$	Log pop. 10–14$_t$	Log pop. 25–29$_t$		Log pop. 40–44$_t$	
	OLS (1)	2SLS (2)^a	2SLS (3)^a	OLS (4)	2SLS (5)^a	OLS (6)	2SLS (7)^a
Log $\overline{\text{rain}}_{\{t-2,..,t-11\}}$	0.028						
	(0.097)						
Log rain$_{t-15}$				0.001		0.002	
				(0.019)		(0.015)	
Log rain$_{t-16}$				-0.002		-0.001	
				(0.018)		(0.013)	
Log rain$_{t-17}$		0.009	-0.007	-0.001	-0.001	-0.003	-0.004
		(0.013)	(0.007)	(0.017)	(0.019)	(0.012)	(0.012)
Log rain$_{t-18}$		0.009	-0.008	-0.004	-0.003	-0.010	-0.010
		(0.011)	(0.008)	(0.017)	(0.016)	(0.017)	(0.019)
Log rain$_{t-19}$		0.010	-0.004	0.001	0.001	-0.011	-0.011
		(0.010)	(0.009)	(0.015)	(0.014)	(0.020)	(0.021)
Log men15-19$_t$		0.253	0.946	0.258	0.247	0.554***	0.572
		[0.720]	[0.174]	(0.182)	[0.898]	(0.153)	[0.616]
Obs.	1754	1754	1754	1754	1754	1754	1754
RMSE	0.065	0.060	0.054	0.089	0.089	0.104	0.104

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket[] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. RMSE is the root mean square error. $\overline{\text{rain}}_{\{t-2,..,t-11\}}$ is the average rainfall over a ten-year period from $t-2$ to $t-11$.
^{a}2SLS estimates with lagged rainfall levels $t-15$ and $t-16$ as first-stage instrument.

Table A.IV: 2SLS Results: Full Control Variables Specification

Explanatory Variables:	Log men15–19$_t$	Log rain$_t$	Log rain$_{t+1}$	Log pop. density$_t$	Log GDP p. capita$_t$	ΔPolity IV$_t$	Log real openness$_t$	Log remaining population$_t$	Log pop20–24
Onset conflict$_t$	3.815***	0.057	0.030	-1.841	-0.208	-0.010*	0.082	0.233	-1.082*
	[0.001]	(0.041)	(0.023)	(1.755)	(0.203)	(0.006)	(0.070)	(0.377)	(0.638)

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 2SLS regression with log rain$_{t-17}$ as instrument for log men15–19 in the first stage. Observations: 1,520. F-test of excluded instrument: 3.2. FE estimator regressions in all columns with country dummies, country-specific time trends (except column 6), time dummies and robust standard errors clustered at country level in parentheses (). The square bracket[] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. The F-test is the first-stage test statistic of the excluded instrument. Δ Polity IV is the change of the Polity2 score from the Polity IV database. Real openness is from the PWT.

Table A.V: 2SLS Results - Sample Split Youth Bulge

Dependent Variable:	Onset Civil Conflict$_t$				
	OLS	OLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
	Sample A: Above country-specific mean				
Log rain$_{t-17}$	0.123***	0.128***			
	(0.045)	(0.035)			
Log men15–19$_t$			8.526***	3.635***	4.849***
			[0.003]	[0.000]	[0.000]
Log rain$_t$					2.078**
					(0.935)
Log men15–19 x log rain$_t$					-0.399**
					(0.167)
Obs.	764	764	764	764	764
F-test excl. IV			1.509	4.626	12.430
	Sample B: Below country-specific mean				
Log rain$_{t-17}$	0.004	0.002			
	(0.015)	(0.014)			
Log men15-19$_t$			0.448	0.046	-0.025
			[0.787]	[0.889]	[0.981]
Log rain$_t$					-0.038
					(0.602)
Log men15-19 x log rain$_t$					0.006
					(0.138)
Obs.	990	990	990	990	990
F-test excl. IV			1.761	4.029	3.713
Country-specific time trends	YES	NO	YES	NO	NO

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Sample A includes all observations with the men15–19 population above the country specific mean. Sample B vice versa. FE estimator regressions in all columns with country dummies, time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket [] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. The F-test is the first-stage test statistic of the excluded instrument. RMSE is the root mean square error.

Table A.VI: 2SLS Results - Mid-Life Moderation and Civil Conflict

Dependent Variable:	Log men15–19$_t$	Log men40–44$_t$		Onset Civil Conflict$_t$		
	OLS	OLS	OLS	OLS	2SLS	2SLS
	(1)	(2)	(3)[a]	(4)	(5)	(6)
Log rain$_{t-43}$	0.005	0.065**	0.066**	-0.048*		
	(0.012)	(0.030)	(0.029)	(0.026)		
Log rain$_{t-44}$	0.004	0.063*	0.063*	-0.054*		
	(0.014)	(0.032)	(0.032)	(0.028)		
Log men40–44$_t$					-0.790***	-0.887***
					[0.000]	[0.000]
Log men15–19$_t$						1.484***
						[0.000]
First stage regression Log men40–44$_t$						
Log rain$_{t-43}$					0.065**	0.069**
					(0.030)	(0.030)
Log rain$_{t-44}$					0.063*	0.064*
					(0.032)	(0.032)
Log rain$_{t-17}$						0.024
						(0.019)
Log rain$_{t-18}$						0.013
						(0.021)
First stage regression Log men15–19$_t$						
Log rain$_{t-43}$						0.010
						(0.013)
Log rain$_{t-44}$						0.006
						(0.014)
Log rain$_{t-17}$						0.041**
						(0.016)
Log rain$_{t-18}$						0.028*
						(0.015)
Obs.	1754	1754	1754	1754	1754	1754
RMSE	0.065	0.109	0.109	0.206	0.221	0.230

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. FE estimator regressions in all columns with country dummies, country-specific time trends, time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket [] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. RMSE is the root mean square error.
[a]Regression model includes lagged rainfall levels $t+1$ to $t-1$ as controls (coefficients not reported).

Table A.VII: Youth Bulge and Low-Level Social Violence

Dependent Variable:	Onset Low-Level Social Violence$_t$			Onset Civil Conflict$_t$
	OLS	OLS	2SLS	OLS
	(1)^a	(2)	(3)^a	(4)
Log men15–19$_t$			4.025***	
			[0.004]	
Log rain$_{t-14}$	0.154*			0.038
	(0.082)			(0.049)
Log rain$_{t-15}$	0.110			-0.077
	(0.072)			(0.046)
Log rain$_{t-16}$	0.106			0.047
	(0.087)			(0.039)
Log rain$_{t-17}$	0.166**	0.185**		0.152***
	(0.063)	(0.067)		(0.042)
Log rain$_{t-18}$	0.056			-0.071
	(0.066)			(0.046)
Log rain$_{t-19}$	0.053			0.106*
	(0.067)			(0.062)
Obs.	820	820	820	662
RMSE	0.400	0.401	0.483	0.244

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In all columns I control for t and $t+1$ rainfall levels (nonsignificant). FE estimator regressions in all columns with country dummies, time dummies and robust standard errors clustered at country and year level in parentheses (). The square bracket [] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. RMSE is the root mean square error. Low-level social violence is from the Social Conflict in Africa Database (SCAD) and excludes all explicit urban events.

^{a} 2SLS regression with log rain$_{t-17}$ as instrument for log men15–19 in the first stage (F-test excluded instrument 3.5).

Table A.VIII: 2SLS Robustness- Youth Bulge and Civil Conflict

Dependent Variable:	Onset Civil Conflict$_t$							
	GPCC Rain Data				Delaware Rain Data		PREC/L Rain Data	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Men15-19$_t$	1.651***	1.369**	1.388**	0.739**	1.724**	1.716**	1.267*	1.365**
	[0.003]	[0.048]	[0.011]	[0.039]	[0.013]	[0.024]	[0.078]	[0.013]
	First stage: Log men15-19							
Log $\overline{\text{rain}}_{\{t-17,..,t-19\}}$	0.083*							
	(0.042)							
Log rain$_{t-15}$				0.030**				
				(0.014)				
Log rain$_{t-16}$				0.029**				
				(0.012)				
Log rain$_{t-17}$			0.040***	0.040***	0.035*	0.035**	0.036*	0.039**
			(0.014)	(0.013)	(0.018)	(0.016)	(0.019)	(0.018)
Log rain$_{t-18}$		0.027*	0.027**	0.026**		0.019		0.022
		(0.015)	(0.012)	(0.012)		(0.013)		(0.018)
Log rain$_{t-19}$		0.015	0.016	0.015		0.008		0.017
		(0.016)	(0.014)	(0.014)		(0.012)		(0.013)
Obs.	1754	1754	1754	1754	1754	1754	1553	1476
RMSE	0.230	0.222	0.223	0.210	0.232	0.232	0.223	0.223
F-test	4.027	2.007	3.373	3.026	3.721	3.450	3.535	4.295
Hansen p-value		0.129	0.253	0.107		0.362		0.274

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. FE estimator regressions in all columns with country dummies, country-specific time-trends, time dummies and robust standard errors clustered at country and year level—except in column 3, 4, 6, 8: country level—in parentheses (). The square bracket [] gives the p-value of the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. All regressions include log rainfall $t+1$, t, $t-1$ which are not reported and enter insignificant. The F-test is the first stage test statistic of the excluded instrument. RMSE is the root mean square error. The rainfall datasets are described in the main text. $\overline{\text{rain}}_{\{t-17,..,t-19\}}$ is the average rainfall for the lags $t-17$ to $t-19$.

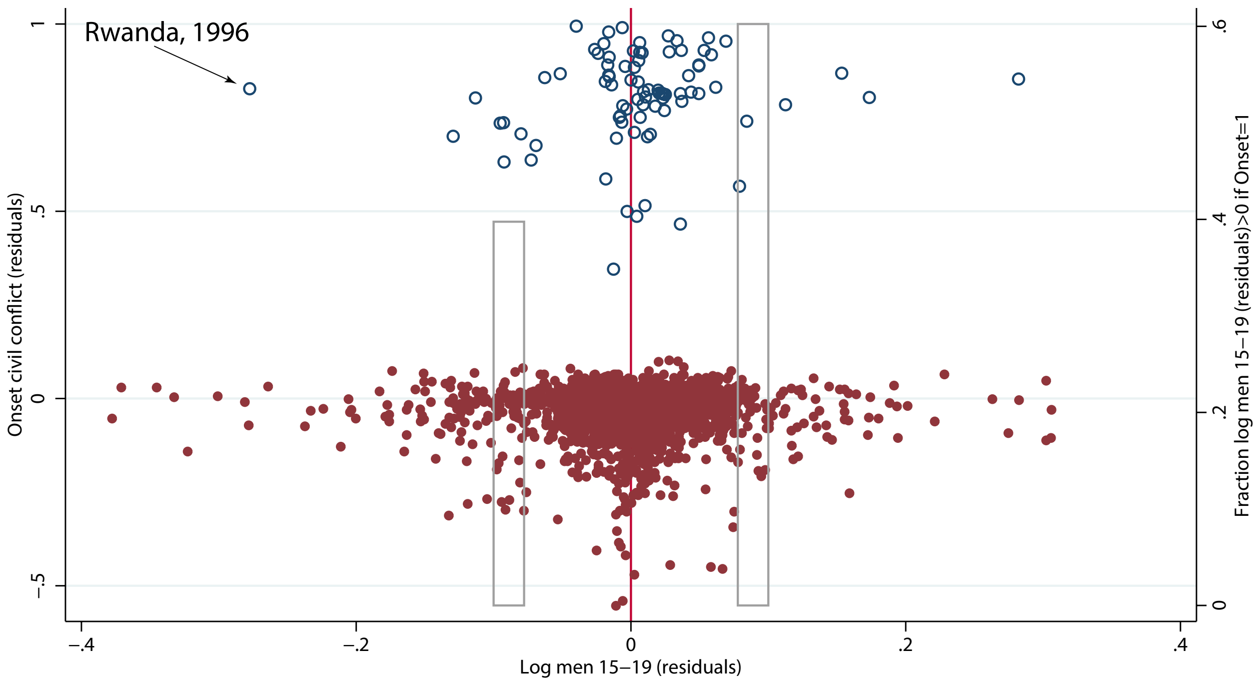

Figure A.I: SSA full sample of log men15–19 and onset of civil conflict. Hollow circles represent civil conflict onset events. Filled dots are no civil conflict events. Bar charts represent fraction of observations of log men 15–19 below and above the sample mean (at 0) given that civil conflict onset (onset=1) is observed. Residuals are purged by time trend, country-specific time trends, country fixed effects and time fixed effects.

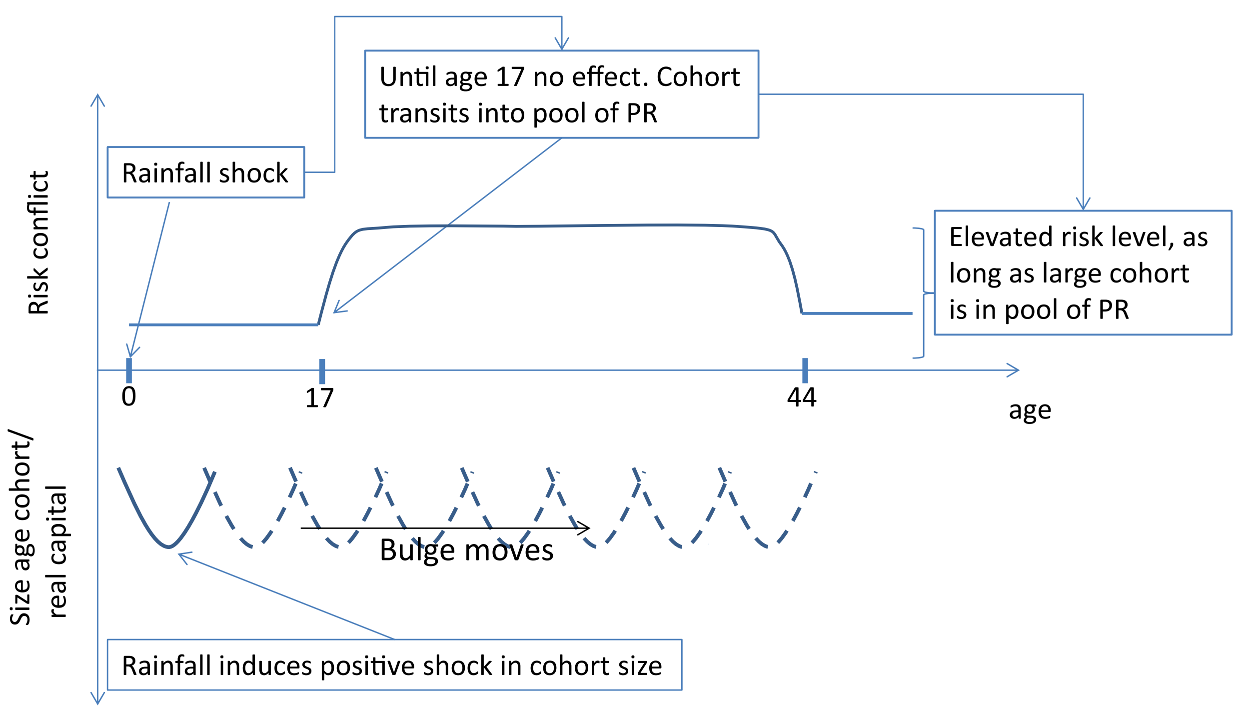

Figure A.II: Bulges and Risk of Civil Conflict