The Legacy of Conflict: Aggregate Evidence from Sierra Leone

Tillman Hönig *

London School of Economics

Job Market Paper

November 2020

(click here for latest version)

Abstract

This paper studies the general equilibrium impact of civil war in Sierra Leone. First, I use an instrumental variable (IV) strategy and geographic conflict variation to estimate reduced-form effects. I show that civil war leads to affected areas having a higher share of workers in agriculture and lower worker income as a result. In order to explicitly take into account general equilibrium effects such as selective migration in response to the war, I then develop an economic geography model. The model sheds light on different mechanisms through which conflict affects aggregate income: Changes in education and firm productivity have both direct effects on income and indirect effects by changing the allocation of labour across sectors and locations. Changes in amenities also affect the spatial allocation of labour. Next, while education outcomes are observed, I leverage the structure of the model to identify unobserved firm productivities as well as amenities. I find that conflict strongly affects education and non-agricultural firm productivity while amenities and agricultural firm productivity are unaffected in the long run. Finally, I use the model to perform counterfactual simulations. In the absence of civil war, aggregate income in Sierra Leone would be 14.8% higher today. Human capital losses can account for about 1/3 of the effect.

JEL classification: D74, J24, O12, O15, O47, R1

*Contact: t.hoenig@lse.ac.uk. I am extremely grateful to Tim Besley, Gharad Bryan and Daniel Sturm for invaluable guidance and support with this project. I thank Lydia Assouad, Oriana Bandiera, Robin Burgess, Noémie Chomet, Alexandre Desbuquois, Maitreesh Ghatak, Musa Kpaka, Rachael Meager, Niclas Moneke, Gerard Padró i Miquel, Martina Zanella, Céline Zipfel, as well as seminar participants at the LSE and the RGS Econ conferences for helpful comments and discussions. I gratefully acknowledge Statistics Sierra Leone for the provision of the underlying data.

1. Introduction

Violent conflict is still a major global challenge with devastating consequences for the life of those affected. In the last decade, civil war plagued about a fifth of all countries worldwide with a global death toll of almost 900,000 people.[1] While a growing literature documents the short-run economic consequences of war, we have relatively little evidence on the persistence of these effects, in particular in terms of aggregate income. Investigating the effect on aggregate income requires taking into account general equilibrium effects of conflict. This paper provides such an analysis in the empirical context of Sierra Leone which suffered an atrocious civil war between 1991 and 2002. Specifically, this paper asks two questions. First, what is the long-run effect of the civil war on aggregate income? Second, what are the potential drivers of persistent income losses?

To answer these questions, I am the first one to develop and structurally estimate a general equilibrium model to study the effects of conflict. I find that civil war in Sierra Leone substantially reduced human capital and firm productivity in the non-agricultural sector. This has both direct implications for aggregate income as well as indirect ones by changing the sector composition of workers and the labour allocation across space. Taken together, aggregate income is 14.8% lower today as a result of the war that ended almost twenty years ago. The human capital reduction can account for about 1/3 of the effect.

This paper proceeds in four steps. First, I establish reduced-form results on the effect of conflict. Drawing mainly on household survey data from 2018, almost twenty years after the end of the war, I observe workers' income, sector choice and education as outcome variables. Conflict intensity is measured using information on households were victimised as a result of the war. Since conflict is not randomly placed in the country, I make use of an instrumental variable (IV) strategy. The heavy involvement with Liberia in the beginning of the war meant that the civil war started at Sierra Leone's south-eastern border with Liberia. This provides a convenient empirical setting to use the distance to the Liberian border as an instrument for conflict intensity. Without losing a strong first stage, I can perform the IV estimation within small districts and controlling for a range of various socio-economic and geographic characteristics which allows the instrument to create plausibly exogenous variation. While this exercise establishes that there are large income differences between more and less affected areas by conflict, these are not necessarily reflective of *aggregate* income effects because of general equilibrium effects.

[1] This is according to UCDP data from Pettersson & Öberg (2020).

Second, I present a general equilibrium economic geography model of the Sierra Leonean economy with two purposes in mind. First, the general equilibrium structure allows me to make progress on understanding and estimating aggregate income effects. Second, the model provides a theoretical framework of the mechanisms underlying the impact of conflict on the economy. Changes in the following three parameters as a result of conflict give rise to both direct and indirect effects on aggregate income: (i) human capital, (ii) firm-level productivities and (iii) amenities. The model has 151 locations which reflect chiefdoms at the lowest level of administration in Sierra Leone. It features labour mobility across these locations and two sectors. We have heterogeneous individuals with idiosyncractic sector- and destination-specific productivities as well as exogenous human capital realisations. In a Roy (1951) fashion, this will give rise to worker sorting into a destination and the agricultural or non-agricultural sector. Competing firms operate in each sector and destination to produce a unique good with firm-level productivity that is specific to the sector and destinations. Finally, destinations differ in their level of amenities. The setup is an extension of the model used by Bryan & Morten (2019) to which I introduce different sectors and a labour movement response to sector-specific human capital changes.

Third, I carry out the structural estimation of the model parameters. In particular, I can leverage the structure of the model to estimate unobserved firm-level productivities and amenities. Having identified these parameters at the sector and destination level, I can use the same IV strategy as before to estimate the causal impact of conflict on firm-level productivities and amenities.

Fourth, I provide counterfactual simulation to estimate the aggregate income effect of conflict that takes into account general equilibrium responses. The IV estimates of the effect of conflict on education, firm-level productivities and amenities inform how I reverse the conflict shock in the model for counterfactual simulations.

A key result that emerges from the reduced-form analysis is that workers living in destinations hit by conflict experience substantial income losses, even twenty years after the end of the war. An increased conflict intensity in one chiefdom vis-à-vis another leads to a drop in income by 23%. A look at the sector allocation of workers reveals an important driver behind this income loss: reverse structural transformation. Workers are 12 percentage points more likely to work in the lower-paying agricultural sector in chiefdoms experiencing one more standard deviation of conflict.

The model highlights how conflict affects income and sector allocation when people lose out on education, firms experience productivity losses in different sectors and places suffer from amenity reductions. In particular, it reveals direct effects in the locations affected by conflict and indirect effects that arise in general equilibrium when people move and market wages change. With returns to education that are estimated to be substantially higher in non-agriculture than in agriculture, education losses would naturally lead to a sector shift and income losses in both sectors. Quantitatively, this plays an important role. Increasing conflict by one standard deviation leads to a loss of schooling by 0.6 years. Using the estimates of firm-level productivities in agriculture and non-agriculture in each destination from the structural estimation, I show that conflict affects the sectors differentially. While there is no effect on agricultural productivity, non-agricultural productivity reduces by 7.7% per standard deviation of conflict. The latter coefficient is indeed insignificant, but sizeable and may simply be very imprecisely estimated. In the counterfactual analysis, I consider both a scenario with and without an effect on non-agricultural productivity. I can show that including the effect generates results that are more consistent with actually observed population movement. Naturally, a firm-level productivity loss in non-agriculture is also consistent with the sector shift into agricultural work and income losses in non-agriculture.

Beyond these direct effects, both education losses and non-agricultural productivity reductions have an important general equilibrium implication. People move in response to the shock. Workers are assumed to be mobile subject a migration cost that is specific to the level of education and the sector they work in. The migration cost is estimated to be sizeable which means that workers choosing to move to another chiefdom must be highly selected in terms of their individual productivity relative to workers staying. At the same time, migration cost for educated workers is estimated to be lower than for non-educated workers. This provides a countervailing force on population movement coming from education losses. A higher share of uneducated workers implies fewer leavers. However, again, those that decide to leave in the face of high migration cost are positively selected in terms of their individual productivity.

Since amenities are not sector-specific and are a component of utility but not of income, any amenity changes would only result in aggregate income effects through general equilibrium effects of worker movement. However, the estimated effect of conflict on amenities is small and insignificant. In the long run, it seems that amenities are not affected by the war.

General equilibrium effects imply that the observed sector allocation and income difference between more and less affected chiefdoms by conflict reflects not only the direct effect of conflict but also selective migration. In particular, with high migration cost and positive selection of those leaving high-conflict chiefdoms, the spatial divergence in income that conflict causes may overstate the aggregate income effect. To address this issue, I use the model to simulate two counterfactual scenarios that reverse conflict and all selective migration in response to conflict with it. First, I consider a full reversal of the war by reverting both the education and non-agricultural firm productivity loss. Second, to assess the quantitative importance of the channels, I reverse only the education loss. The IV estimates inform the degree to which I increase these parameters in a chiefdom as a function of the intensity of conflict it experiences.

The first scenario estimates an aggregate real income gain of 14.8% in the absence of conflict with the share of agricultural workers being 6.3 percentage points lower. In the second scenario, aggregate real income is 4.7% higher without conflict and the share of agricultural workers 1.4 percentage points lower. These results suggest that Sierra Leone is still far from full recovery of the war, even almost twenty years after it ended. A comparison of the two scenarios implies that human capital losses can only explain about 1/3 of the total effect while firm productivity losses seem to play a more important role.

While the estimate on the full war effect is large, it is still substantially lower than what the reduced-form evidence on spatial divergence might suggest. If we were to take the reduced-form estimate and calculate a country-wide weighted average by conflict intensity and chiefdom population we would arrive at a 42% aggregate income loss. This suggests that general equilibrium forces such as selective migration account for more than half the observed spatial income difference between more and less affected chiefdoms by conflict.

This paper makes two main contributions. First, to my knowledge it is the first paper to provide estimates of the long-run aggregate income effect of civil war that explicitly take into account general equilibrium forces. While early cross-country macro studies (Alesina & Perotti, 1996; Barro, 1991; Collier, 1999) show a clear negative link between conflict and aggregate economic performance, establishing causality from these correlations is difficult. A great number of institutional and economic differences between war-torn countries and countries at peace may drive the result. On the other hand, micro-empirical countries that compare more and less affected households or locations such as Miguel & Roland (2011) and Serneels & Verpoorten (2015) are similar in nature to my reduced-form analysis. Even with a solid identification strategy, they cannot account for general equilibrium effects. In order to take such effects into account, I develop an economic geography model that draws on Bryan & Morten (2019), Eaton & Kortum (2002), Hsieh et al. (2019) and Redding (2016). In particular, civil war typically leads to the reallocation of labour across space with important general equilibrium implications. This concerns not only the size of population movement but also the selection. Indeed, Davis & Weinstein (2002) and Brakman et al. (2004) find that city growth is unaffected by bombing intensity. However, this does not speak to the selection of migrants. In this study, I also find that population movement is unaffected by conflict intensity but the selection of migrants seems to play a major role in accounting for spatial differences in income as a a result of the war.[2]

The second contribution of this paper is to shed light on a channel of the long-run impact of conflict impact that has received little attention so far: reverse structural transformation, in particular as fuelled by decreases in non-agricultural firm productivity. This can capture a variety of elements that could be affected by war such as market access or electricity connection. My model features labour as the only input into production such that the firm-level productivity term would also capture any physical capital input that is immobile, such as local buildings or machines that could be destroyed in war. While only few studies consider firm or sector allocation outcomes and these are typically short-run analyses (Bozzoli et al., 2012; Camacho & Rodriguez, 2013; Collier & Duponchel, 2013), the literature on human capital effects of conflicts is large.[3] Indeed, this paper demonstrates evidence on human capital losses as a result of the war but their role in explaining persistent aggregate income effects that are driven by reverse structural transformation is shown to be limited. Depending which factors are the main drivers behind persistent consequences of conflict for economic welfare, the implications for post-war policy differ greatly. My paper suggests that restoring firm-level productivity deserves greater policy focus.

The remainder of the paper is organised as follows. Section 2 provides a brief overview of the Sierra Leonean civil war that motivates it as an empirical setting. Section 3 discusses the empirical design of this study and section 4 presents reduced-form results. In section 5, I develop the model and discuss its estimation strategy in section 6. Section 7 present the results of the model estimation, followed by the presentation and discussion of counterfactual simulations in section 8. Finally, section 9 concludes.

[2]This result is also in line with other papers that highlight the importance of selective migration in explaining spatial income differences such as Young (2013).

[3]Focusing on civil war and long-run outcomes at least 10 years after the end of war, these include Akbulut-Yuksel (2014); Akresh & De Walque (2011); Galdo (2013); Justino & Verwimp (2013); La Mattina (2018); Leon (2012); Saing et al. (2017).

2. The Civil War in Sierra Leone

Sierra Leone suffered a civil war between 1991 and 2002 that caused some 70,000 casualties, displacement of over half the population and left many people injured, maimed and raped (UNDP, 2006). While the war was extremely brutal, Sierra Leone has experienced a long period of sustained peace since it ended in early 2002. This provides an ideal setting to investigate long-run effects of the war today.

The civil war started as an insurgency by the Revolutionary United Front (RUF) under Foday Sankoh in 1991 entering the country from Liberia in the south-eastern part of the country.[4] The RUF was a small rebel group at the onset of war with the political goal of overthrowing the ruling one-party regime, led by the All People's Congress (APC) party under Joseph Saidu Momoh (Richards, 1996). Their insurgency was supported by the National Patriotic Front for Liberia (NPFL) led by Charles Taylor and involved in the ongoing Liberian civil war. In fact, the RUF had started their fighting activities in Liberia along with the NPFL when the war broke out in the neighbouring country in 1989. Foday Sankoh and Charles Taylor had met each other and worked and trained together before. The RUF remained mainly active and the fighting very much focused in the southern parts of Sierra Leone bordering Liberia between 1991 and 1995 until it eventually spread in the whole country. This involvement with Liberia means that distance to the Liberian border is highly predictive of conflict intensity and can be used as an instrument for it.

As Richards (1996) argues, the element of political grievances played a great role as a cause of the civil war. In particular young people were discontent with a patrimonial system in which a small group of patrons rules and decides on the allocation of opportunities and transfers arbitrarily. They felt disenfranchised and robbed of education and opportunities. The RUF's ideological roots lied in an idea of egalitarianism which initially helped in recruiting disenfranchised youth – however, as knowledge of the atrocities committed by the group spread, recruitment by capture became more necessary and common.

One of the atrocious features of the Sierra Leonean civil war was the extreme degree of violence against civilians, in particular all the community looting operations as well as the raping, killing and maiming that characterised the war. Quite tellingly, such operations were called "Operation Pay Yourself" or "Operation No Living Thing". These acts of violence were not only committed by the RUF, but also by the Sierra Leonean Army (SLA) throughout the war, often by so-called "Sobels" who were soldiers by day

[4]Figure 1 displays the location of Sierra Leone and Liberia in West Africa.

and rebels by night, taking on an identity under which it was more legitimate and less consequential to engage in these activities. Since such violent activities characterised how households were affected by the war, it is very convenient that the Sierra Leonean Integrated Household Survey 2011 provides direct information on how households were victimised as a result of the war. This information can be used as a measure for conflict intensity that captures the effect of the civil war well.

The opportunistic behaviour of fighters demonstrates that there was an element of "greed" to the civil war that also became increasingly prevalent in the illicit mining or smuggling of diamonds. The diamond wealth resulting from these activities helped funding the war and provided incentives to prolong it (Keen, 2005; Richards, 2004). Therefore, throughout the war, economic motives became an increasingly important motivation for rebels to engage in fighting.

Another interesting feature is the lack of ethnic or religious divisions as a key driver of war, as Bellows & Miguel (2009) point out. No ethnic group seemed to be disproportionately victimised and there seems to be no evidence that violence against a particular civil community was more pronounced if the community and the fighting group have largely differing ethnicities.

3. Empirical Design

3.1. Data

The main source of data for this paper is the Sierra Leonean Integrated Household Survey (IHS) 2011 and 2018 which are general representative individual-level survey. I use detailed questions on economic activity in 2018 to construct the following outcome variables. First, as a proxy for worker income, I use detailed household expenditure information on food and non-food items and divide this by the number of working people in a household.[5] The reason for using information on expenditure rather than income directly is that the expenditure data is much more complete and highly likely to be more reliable. It is recorded in weekly visits by enumerators during which households indicate the items they bought and at which price they did so. Such information is significantly easier to remember than providing information on different income sources over the past year in a

[5] The division by the number of workers is done to reflect worker income. However, all my IV analysis can be performed on related measures such as total household expenditures, expenditures per capita or expenditures per adult equivalent. All main findings go through in this robustness exercise. The results are available upon request.

setting where most workers are subsistence farmers or engage in small business activities without any bookkeeping. Second, the main sector that individuals work in is constructed according to the ISIC classification (see Figure 3). For all main results, a simple binary distinction between agriculture and non-agriculture (manufacturing and services) is made. Third, education outcomes are recorded directly. From information on completed grades, I use years of schooling as well as an indicator for having finished primary education.

Regarding conflict data, the 2011 survey contains a section with a number of questions on the impact of conflict on individuals and households that I make use of. As a main conflict measure, I follow Bellows & Miguel (2009) in constructing a victimisation index. This index is the share of "yes" answers to eight binary questions in the survey that cover how households were affected by the war along the following dimensions: (1) whether the household lost property or assets, (2) whether the house was burnt, (3) whether household members were killed, (4) whether relatives were killed, (5) whether household members lost limbs, (6) whether household members were molested or raped, (7) whether household members were displaced, (8) whether the war had any other effect on the household.[6] Given that the extreme degree of violence against civilians was a feature of the Sierra Leonean civil war, a victimisation index seems to be a sensible measure of the intensity of conflict. For my analysis and in line with Bellows & Miguel (2009), I aggregate the household-level victimisation experience at the chiefdom level to construct conflict measures at that level. With five chiefdoms missing in the household survey 2018 and three different ones missing in the household survey 2011, I observe 150 chiefdoms in addition to the capital Freetown in 2018. Chiefdoms are the lowest administrative level. At the next higher level, Sierra Leone is administered in 14 districts, the level of my fixed effects in the empirical analysis. The aggregation and subsequent treatment of conflict at the chiefdom level serves to capture potentially large within-chiefdom spillovers of conflict.

Furthermore, the 2018 survey contains detailed information on migration. In particular, the chiefdom of birth as well as the chiefdom of residence and the year of moving are recorded for each individual. This information will be crucial when estimating the model that requires knowledge of both origin and destination for all individuals. In addition to socio-economic control variables from the IHS data, I use data from Glennerster et al. (2013) for some geographic controls and exploit census data from 1963, 1985, 2004 and

[6]In fact, Bellows & Miguel (2009) use very similar questions in a survey carried out by the Institutional Reform and Capacity Building Project (IRCBP) to construct their victimisation index in the same way.

2015 to analyse population movements.[7] I also use geographic information to compute distance of a chiefdom centroid from the Liberian border which serves as an instrument in my IV analysis. In order to test whether this instrument correlates with pre-war characteristics, I draw on data used in Bellows & Miguel (2009) and Acemoglu et al. (2014) providing some economic outcomes before the war, including education and expenditure in 1989 as well as historic tax and trade variables. Summary statistics for all variables used are provided in Table 1.

3.2. Identification Strategy

My main specification to estimate the effect of conflict on outcomes is an instrumental variable (IV) specification using the distance to the Liberian border as an instrument for conflict. Given the inter-linkages between the civil wars in Liberia and Sierra Leone and the fact that fighting originated and focused in the border area for a long time, it seems reasonable that the distance to the Liberian border is a strong predictor for conflict levels. As Figure 4 depicts clearly, while fighting happened in the whole country throughout the whole war, the highest conflict levels are experienced by areas bordering Liberia in the South east and in the corridor between the border and the capital Freetown in the very West towards which the rebels progressed. My IV-2SLS specification is characterised by the following equations:

$$y_{ic} = \alpha_{district} + \beta\, \widehat{conflict}_c + \mathbf{X_{ic}}'\mu + \epsilon_{ic} \quad \text{(Second Stage)} \tag{1}$$

$$\widehat{conflict}_c = \hat{\alpha}^{FS}_{district} + \hat{\gamma}^{FS} distance_{ic} + \mathbf{X_{ic}}'\hat{\mu}^{FS} \quad \text{(First Stage)} \tag{2}$$

where y_{ic} is the outcome of interest for invidiual i in chiefdom c, $distance_c$ is the instrument and $\alpha_{district}$ capture district-level fixed effects. The vector of controls includes a set of socio-economic controls as well as characteristics of the land. Importantly, in all specifications, I control for distance to one of the five largest cities in Sierra Leone (Freetown, Bo, Kenema, Koidu, Makeni). These five cities are well known as urban and regional economic centres and the only large cities of the country with a population exceeding 100,000 inhabitants. Controlling for distance to these cities insures that any potential mechanical relationship between border distance and distance to large cities does not act as a confounder. The underlying reason for a potentially confounding relationship is

[7] I only have aggregate information at the chiefdom level for the 1963 and 1985 census and excerpts for the later waves of the census.

that economic development can have a highly geographic component. Growth in an area can subsequently lead to economic development around that area (Felkner & Townsend, 2011).[8]

While different regions within Sierra Leone differ not only in their distance to the Liberian border but also in terms of other characteristics that are correlated with economic outcomes, performing the analysis within districts is crucial to satisfy the exclusion restriction. The identification assumption is therefore that, within districts and conditional on the set of control variables used, distance to the Liberian border only affects outcomes after the war through inducing variation in conflict, but not through any other channel.

One potential concern with this assumption is that distance to the Liberian border is naturally related to trade. Even within the same district, chiefdoms that are located closer to the border may have been more active in trading before the war and therefore at a different level of economic development. This would lead to a violation of the identification assumption. However, I argue that such a violation is unlikely to play a major role. Trade with Liberia is only of negligible size relative to the total trade volumes of Sierra Leone before the war. Exports to Liberia as a share of total exports are less than 0.1% before the war.[9] The main trading partners of Sierra Leone are Europe and the US and trade with these partners would not go through Sierra Leone but rather through their main port in Freetown. Distance to Freetown as one of the major urban centres is controlled for in all specifications. Furthermore, I provide a test against the hypothesis of trade as a major confounder by excluding chiefdoms directly bordering Liberia and repeating my main IV analysis on worker income as an outcome. To the extent that chiefdoms with a direct border would benefit particularly from trade, this should lead to results that differ from the main results on the full sample. Table 5 provides the results which are very comparable to the main results with the full sample shown in Table 4. Finally, districts are small. Sierra Leone as a whole country is as large in area as the Netherlands and Belgium and subdivided into 13 rural districts, excluding Freetown. The average district's area is therefore only 5000 square kilometers. While location likely matters for

[8]In addition, the full set of controls contains the following variables. Socio-economic controls are household head's sex, an urban dummy, age, age squared, religion as well as household size. The land characteristics are a vector ruggedness measure, the average elevation, the share of chiefdom terrain with slope between 2-8%, 8-30% and 30-45%, the share of chiefdom terrain with coarse texture and with medium texture, as well as the share of chiefdom soil with poor drainage and with excessive drainage.

[9]UN Comtrade Database (accessed in October 2020 under https://comtrade.un.org/). The most recent year with information on trade in Sierra Leone before the war started is 1986.

trade at the country level, it is therefore unclear to what extent it would play a role within these small districts.

Another concern pertaining in particular to the sector allocation as an outcome of interest may be differences in the quality of the land that are related to the border distance. This is the main reason behind including a variety of land characteristics as a a set of control variables in the analysis. In fact, the inclusion of these controls leads to stronger IV results. Therefore, choosing the sparser and more precisely estimated specification without land controls as my preferred specification is a conservative approach which means that the results can be interpreted as a lower bound on the true effect.

Further to these considerations, I also use pre-war characteristics drawing on data used in Bellows & Miguel (2009) and Acemoglu et al. (2014) in a placebo test to check whether distance to the Liberian border has predictive power for economic outcomes before the war within districts. The results are shown in Table 2. Distance to the border within districts has no significant predictive power (at the conventional 5% level) for all outcomes. In fact, only one outcome has a significant coefficient at the 10% level which can easily arise by chance when testing nine outcomes. Given these results on observable variables, it seems very unlikely that the instrument is related to post-war outcomes through other unobservable channels.

Considering the first stage, Figure 6 graphically displays the correlation between my standardised conflict measure and distance to the Liberian border within districts. I formally test the first stage and report the results in Table 3. I include Kleibergen-Paap F statistics that allow for the cluster structure of my error term and are still comfortably above the conventional threshold of 10.[10]

For the analysis of education outcomes, the fact that education is usually obtained during a particular age provides me with additional cohort variation that I can exploit for two purposes. First, it allows for a placebo test of the identification assumption. I can split the sample and run the IV analysis separately for people who were at school age when the war started and those who were already old enough to have finished their education. I use a generous definition of school age with age 30 at the beginning of the war as cut-off point.[11] If the instruments does not satisfy the exclusion restriction in a way that is relevant for education as an outcome, this should become visible in the analysis

[10] Note that Kleibergen-Paap F statistics are equivalent to Montiel-Pflueger F statistics in the weak IV test in my case with one endogenous regressor and one instrument.

[11] The results are robust to using different cut-off points, e.g. anyone aged 18 or older at the beginning of the war.

of the old cohort. Table 9 demonstrates the impact of education by cohort. While there is a strong effect for the young cohort (columns 1 and 4), there is no significant effect for the old cohort (columns 2 and 5). This provides some reassuring evidence that the instrument is indeed orthogonal to education prior to the war.

Second, I will also estimate the following regression in the spirit of a difference-in-differences IV approach:

$$y_{ijc} = \alpha_c + \delta(\widehat{conflict_c \times coh_j}) + \zeta\ coh_j + \tilde{\mathbf{X}}'_{\mathbf{ijc}}\mu + \epsilon_{ijc} \quad (3)$$

$$(\widehat{conflict_c \times coh_j}) = \hat{\alpha}_c^{FS} + \hat{\gamma}^{FS}(distance_c \times coh_j) + \hat{\zeta}^{FS} coh_j + \tilde{\mathbf{X}}'_{\mathbf{ijc}}\hat{\mu}^{FS} \quad (4)$$

where i denotes an individual, j the cohort and c denotes the chiefdom. The variable coh_j is a cohort dummy variable indicating whether an individual was at school age during the war or not. I use the same cohort definition as before with individuals born after 1961 considered at school age and those born before 1961 considered beyond school age during the war. My main education results are robust to choosing a different cutoff year, e.g. 1973 and therefore making a distinction between individuals below or above the age of 18 when the war started. The vector of controls $\tilde{\mathbf{X}}_{\mathbf{ijc}}$ includes all socio-economic control variables that are used in the main specification. Unlike in the above IV design, the additional cohort variation allows me to control for a *chiefdom* fixed effects α_c. Therefore, any further chiefdom-level controls such as the land characteristics are captured in the fixed effect and not controlled for separately. Taking δ as the causal effect of interest, the identification assumption now becomes a weaker condition. A relationship between distance to the border and education outcomes other than through conflict may exist within chiefdoms, we only require it to be the same for school-aged and older individuals at the start of the war. This corresponds to the classic difference-in-differences (DID) assumption applied to the instrument.

4. Reduced-form IV Results

Table 4 demonstrates the results of an OLS regression and the IV specification for worker income, as proxied for by expenditures per worker, for different sets of control variables. The effect of conflict is large. An additional standard variation of conflict intensity reduced income by 24-30% twenty years after the end of the war. This implies substantially lower livelihoods of households affected by the war. The results are robust to just considering household expenditures or per capita measures of household expenditures.[12] Two observations that apply to this and further results on economic outcomes are noteworthy. First, the specification including geographic controls delivers a stronger result. This could be the case because there is some bias when not controlling for geographic features of the chiefdoms. Considering the first stage results and the large size of the effect even without geographic controls, however, this could also plausibly reflect the strength of the instrument conditional on the whole set of socio-economic and geographic controls. The relevant F statistic drops from 20.8 to 12.6 when including geographic controls. In the spirit of a cautious interpretation of the results, I would therefore consider the specification without geographic controls as the more reliable one and take it as a more reasonable estimate of the effect of civil war. Second, the IV results are stronger than the OLS results. This positive difference between IV and OLS results is consistent with a positive selection into conflict in the sense that conflict takes place in areas that are richer to begin with. In light of the fact that economic considerations played a key role in the rebels' decision to engage in conflict this is plausible. At the individual level, many young people were easily recruited by the rebel movement because engagement in looting communities was economically more attractive than alternative ways to make a living. At the collective level, the rebel movement aimed at controlling and generating revenue from diamond mines as a source of income.[13]

A large sector shift in economic activity may be a key driver behind these income effects. The IV results on the main sector of work are shown in Table 6 and graphically displayed in Figure 7. With a standard deviation increase in conflict intensity, workers are 12 ppt. more likely to work in agriculture and correspondingly less likely to engage in non-agricultural activities. Both the manufacturing and services sector are affected to a similar extent.

If returns to education are higher in the non-agricultural than in the agricultural sector, human capital loss could be the reason for this sector shift.[14] The IV estimation results of the effect of conflict on years of schooling is presented in Table 7. As conflict intensity increases by one standard deviation, individuals lose between 0.6 and 0.7 years of school-

[12] Household expenditures and expenditures per adult equivalent are reduced by 14-21%, depending on the specification. Results are not reported here, but available upon request.

[13] This may raise an endogeneity concern if the location of diamond mines at the beginning of the war is correlated with the instrument. Unfortunately, no information on the location of diamond mines in 1991 is available. Using information on their location in 2002 from Bellows & Miguel (2009), however, I can establish that the instrument is uncorrelated with distance to diamond mines in 2002. Furthermore, controlling for distance to diamond mines does not change any results in the IV analysis.

[14] Indeed, as the structural estimation of the model will reveal, returns to education in the non-agricultural sector are substantially larger than returns in agriculture.

ing. Mapping the definition of education levels in the model directly, Table 8 displays the results for the same specifications with a dummy variable for having finished primary education as an outcome. The results are in line. We see a reduction in the probability of finishing primary school by 7 percentage points.

It is reassuring to see that the effect only materialises for young people who were actually at school age when the war started while there is no significant effect for those old enough to have finished their education by that time already. Table 9 demonstrates these results. Furthermore, the DID-IV analysis in columns (3) and (6) of the same table delivers a very similar and statistically indistinguishable result from the main IV analysis.

These reduced-form results suggest that civil war may have a strong persistent effect on the livelihoods of affected households by reducing human capital and changing their sector of employment. What these estimates identify are spatial differences. The results are essentially generated by comparing more and less affected areas by the civil war. If productive resources are reallocated across space as a result of the war, however, spatial differences do not only capture the direct effect of conflict but also the reallocation of resources. In particular, we may be concerned about labour supply changing in locations when people move in response to the war. The 2011 survey indicates that around 50% of all households were displaced at some point during or after the war. Even considering migration in the long run, migration rates in the 2015 census and 2018 household survey are between 25 and 30%. It is quite plausible that the selection of migrants changes as a result of the war. For example, if more productive non-agricultural workers leave conflict areas during or after the war and stay in less affected areas by war, my estimate would also capture this type of selective migration. The fact that income is higher in less affected areas after the war is partly due to the changing type of movers to these areas. As a result, spatial income differences would overstate the aggregate income effect of conflict. In order to understand the aggregate income effect, the following part will develop a model that explicitly takes into account general equilibrium effects of the war such as selective migration. Structural estimation of the model and counterfactual simulations will allow me to estimate the aggregate income effect of the civil war in Sierra Leone.

Further to estimating the aggregate income effect, the development and estimation of the model also allows me to shed light on *unobserved* mechanisms of the effect of war. While education is observed, there are other important determinants of productivity that could be affected by the war and a driver of the sector shift and income effects. In the model, such determinants will be captured by sector-specific firm productivities. Using the structure of the model, I can estimate these objects and evaluate the impact that conflict has on them.

5. Model

5.1. Basic setup

The model is a static general equilibrium model of the Sierra Leonean economy. In order to capture basic characteristics of that economy and the key pathways of how conflict affects the economy, the model contains the following key features. First, the high migration rate in Sierra Leone with movement potentially responding to conflict motivate an economic geography setup with 151 locations (chiefdoms) and endogenous location choice. The set of locations is denoted by $\mathcal{K}$. Workers choose locations on the basis of an individual location-specific productivity and subject to migration cost. Furthermore, locations differ in aggregate firm productivities and amenities, both of which can be affected by conflict. In this way, conflict can change the spatial allocation of labour.

Second, the reduced-form results indicate that the sector allocation is an important pathway how conflict affects the economy. Therefore, the model features two sectors $S \in \{A, N\}$, agriculture and non-agriculture, and worker sorting into sectors on the basis of individual sector-specific productivities. Firm productivities are also sector-specific and govern sector choice. A differential effect of conflict on firm productivities by sectors can be one mechanism how conflict affects sector allocation. Each location has firms operating in both sectors and producing a unique agricultural and non-agricultural good in an Armington (1969) fashion. The Armington structure serves as a dispersion force. In order to keep the model as simple as possible, I abstract from trade cost and assume that goods are traded costlessly across space.[15]

Third, education and employment in the non-agricultural sector are highly correlated in Sierra Leone. This could be the result of differing returns to education across sectors. With sector-specific returns to education, an education reduction through conflict can have both direct effects on income and indirect effects through a change in the sector composition. To capture this mechanism, the model features education with differential returns by the sector an individual chooses to work in.

This structure allows me to lay out both the direct effect of conflict and the indirect

[15] Section 6 and 8 discuss how the introduction of trade cost would change the model estimation and counterfactual simulations, respectively.

effects of conflict through a change in the sector composition and spatial allocation of labour. By explicitly taking into account general equilibrium forces, I am able to estimate aggregate income effects when estimating the model in a next step. While the reduced-form results identified spatial income differences, aggregate income effects may be very different if there is a strong spatial re-allocation of labour in response to the war. Finally, estimating unobserved parameters such as firm productivities and amenities using the structure of the model allows me to consider potential mechanisms of a conflict impact on income that are not directly observed in the data.

The basic setup borrows from Bryan & Morten (2019) and contains elements from Hsieh et al. (2019), Eaton & Kortum (2002), Allen & Arkolakis (2014) and Redding (2016). I use the structure of labour mobility under movement cost as well as sector and destination choice based on individual productivity draws from the former authors. To this setting, I introduce sectors and human capital with differential returns to education by sector. Relevant to my context, this will give rise to worker sorting across sectors on the basis of both sectoral firm productivities and education.

Preferences. A continuum of individuals i exist in the economy. Each individual is born in an origin chiefdom $c \in \mathcal{K}$ with exogenous education endowment e and choose a destination chiefdom $d \in \mathcal{K}$ to live and work in. Individuals have CES preferences over all goods produced in each sector and destination. They consume quantity c_{id}^S of the good produced in sector S and destination d. Their utility is also influenced by amenities in their destination location a_d and the cost of moving to that destination and choosing their sector to work in τ_{ecd}^S. The utility function is therefore:

$$U_{id}^S = a_d(1 - \tau_{ecd}^S)\left(\sum_d (c_{id}^A)^{\frac{\sigma-1}{\sigma}} + (c_{id}^N)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{5}$$

where σ denotes the elasticity of substitution. In line with standard economic geography models amenities can be thought of to capture elements such as natural beauty, the availability of local public goods and services and the quality of housing. In the particular context of this conflict study, an additional way to think of what amenities capture would be the safety of a place. They enter multiplicatively in the utility function. Like migration cost, they will play a key role in workers' location choice. I model migration cost τ_{ecd}^S as the share of income that workers with education e lose when moving from c to d and choosing sector S. Migration cost can be thought of in several ways. One is the actual physical cost of moving away from home. Beyond that and potentially much more importantly, however, this parameter captures the cost of integration into a new community in order to be able to work there or the (dis)similarity of agricultural and non-agricultural goods produced across space. In the Sierra Leonean context, chiefdoms still have traditional chiefdom administrations and strong local governance structures. Traditional chiefs who are part of locally well-known and respected so-called "ruling families" govern many parts of public life. In this context, integration into a new community is an important factor to be able to live and work there.

This cost τ^S_{ecd} is assumed to be both education and sector-specific. The motivation behind this is threefold. First, this is the most general type of moving cost the model could allow for. I will be able to estimate this cost non-parametrically and therefore let the data decide to what extent moving cost may actually differ across sectors and education level. Second, a growing theoretical and empirical literature seems to suggest that moving cost may differ by education level.[16] Reasons cited here are, for example, a better state of information about opportunities in different places among more educated people or greater availability of valuable job matches. Third, with sector dependence this parameter will capture any friction of entering the non-agricultural sector. This may be due to the necessity of moving to a larger town within the same location for non-agricultural employment. This could also embody the fact that non-agricultural work requires connections or fixed investments for people born in rural areas where agricultural work may be the default option. Empirically, it turns out that such frictions are quite real. Given the estimated large wage differences between workers in agriculture and non-agriculture, such a friction can rationalise the relatively high number of workers in agriculture.

Goods are assumed to be costlessly traded across space which results in common prices for each good p^S_d and a common CES price index P for the entire economy across all destinations:

$$P = \left(\sum_d (p^A_d)^{1-\sigma} + (p^N_d)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{6}$$

Workers supply l^S_{id} effective units of labour in the sector and destination of their choice and are getting paid at the wage rate w^S_d. Their nominal income is therefore $m^S_{id} = w^S_d l^S_{id}$. Standard CES utility maximisation results in indirect utility as a linear function of worker

[16] See, for example, Amior (2015); Kennan & Walker (2011); Wozniak (2010).

income:

$$V_{id}^S = \frac{a_d(1-\tau_{ecd}^S)w_d^S l_{id}^S}{P} \tag{7}$$

Productivities and labour supply. Workers draw individual productivities for each sector and destination z_{id}^S from a Fréchet distribution. The average productivity draw is different for each sector and destination and it also varies by origin and education level. For workers from origin c with education level e, the average productivity draw for sector S and destination d is captured in the scale parameter q_{ecd}^S. Along with independence, this results in the following multivariate distribution from which productivities are drawn:

$$F_{ec}(\mathbf{z}) = \exp\left(-\sum_d \sum_S \left(\frac{z_{id}^S}{q_{ecd}^S}\right)^{-\theta}\right) \tag{8}$$

where $\mathbf{z}$ denotes the vector of productivity draws and the shape parameter θ governs the dispersion of productivities. A high realisation of θ means low dispersion, that is productivities for the goods produced in different destinations are close to each other. One interpretation of this would be a high degree of similarity across products produced in different destinations.

In this formulation, workers from origin c with education e are, on average, the same in the sense that they draw from the same productivity distribution. Let $q_{ecd}^S = \overline{q_{ec}}\epsilon_{ecd}^S$, that is, average productivity draws are made up of a common component $\overline{q_{ec}}$ for all workers from c with education e and some variation across sectors and destination with mean 1. Differences in the common component $\overline{q_{ec}}$ across origins and education levels capture the idea that some origins have better capabilities to produce high-productivity individuals, e.g. better quality of schooling, childcare, etc.

Apart from productivities, workers also differ in the exogenous realisation of their education level e which determines human capital in a standard Mincerian way. Let ϕ^S be the sector-specific returns to education. Human capital for workers with education e is exogenous and given by

$$h_e^S = \exp(\phi^S e) \tag{9}$$

Human capital is combined with a worker's productivity draw to determine their effective individual amount of labour supplied in the market place:

$$l_{id}^S = h_e^S z_{id}^S \tag{10}$$

Sector and destination choice. Using the expression for labour supply 10 in the indirect utility function 7, we can rewrite indirect utility as

$$V_{id}^S = \frac{v_{ecd}^S z_{id}^S}{P} \tag{11}$$

where $v_{ecd}^S := a_d(1-\tau_{ecd}^S)w_d^S h_{ed}^S$. This essentially captures the wage rate in destination-sector (d,S) adjusted for amenities, movement cost and human capital. From the distribution of productivities, it follows that indirect utility for workers from origin c with education e across sectors and destinations follows a Fréchet distribution with scale parameter $\frac{v_{ecd}^S q_{ecd}^S}{P}$. Based on this distribution, we can characterise workers' sector and destination choice. The probability that a worker with education level e from origin c chooses sector S in destination d is:

$$\pi_{ecd}^S := Pr\left[V_{id}^S > V_{id'}^{S'} \; \forall d', S'\right] = \frac{(v_{ecd}^S \epsilon_{ecd}^S)^\theta}{\sum_{d'}\sum_{S'}(v_{ecd'}^{S'}\epsilon_{ecd'}^{S'})^\theta} \tag{12}$$

This equation describes worker sorting behaviour: If the adjusted wage rate in a particular chiefdom and sector is high relative to all other chiefdoms and sectors, the chiefdom and sector are attractive to work in.

In a similar way, properties of the Fréchet distribution give rise to the following characterisation of the average productivity of workers *choosing* sector S and destination d:

$$E\left[z_i | i \in M_{ecd}^S\right] = (\pi_{ecd}^S)^{-1/\theta} q_{ecd}^S \Gamma\left(\frac{\theta-1}{\theta}\right) \tag{13}$$

where $\Gamma(\cdot)$ is the Gamma function.[17] Average productivity in a destination-sector (d,S) depends negatively on the share of workers choosing that destination d and sector S. This reflects a selection mechanism: The marginal migrant who chooses to make that move will be the one drawn from the left-most part of the distribution with the lowest productivity. This selection mechanism is displayed graphically for sector choice in Figure 2. The negative relationship between the share of workers choosing (d,S) and average productivity in that destination-sector pair giving rise to this selection mechanism is a result of alignment of comparative and absolute advantage. Graphically, this is represented by the upward sloping curves in the figure. This assumption of alignment is hard-baked

[17] Derivations of equation 12 and 13 are provided in the Appendix, section C.

into the model here by using indepdent Fréchet draws. Heckman & Honore (1990) refer to this as the standard case and it is intuitively appealing that those workers who have a comparative advantage for working in a particular destination-sector pair would also absolutely perform better there.[18]

This formulation of average productivities is useful to represent the relationship between average wages and average productivity:

$$\overline{m^S_{ecd}} = w^S_d h^S_e (\pi^S_{ecd})^{-1/\theta} q^S_{ecd} \Gamma\left(\frac{\theta - 1}{\theta}\right) \tag{14}$$

Since average productivity determines average wage in a sector S and destination d for all workers from c with education e, the same negative relationship between the share of workers and wages manifests here. The strength of this negative relationship is determined by the size of $1/\theta$. Lower productivity dispersion (high θ) will lead to a small size of $1/\theta$ and therefore little reactivity of average wages to the share of workers from c with education e choosing (d, S). This is intuitively appealing: Low productivity dispersion implies that the marginal migrant worker will be very similar to previous migrants in terms of their productivity. Equation 12 and 14 will be key estimating equations to determine parameters of the model.

Production. Production is linear in the sole input of production, labour L^S_d. Within sectors and destinations, perfectly competing firms are identical and the representative firm deploys firm-level productivity A^S_d:

$$Y^S_d = A^S_d L^S_d \tag{15}$$

Denote the set of workers with education level e from origin c choosing to live in destination d and work in in sector S by M^S_{ecd}. The labour force in a particular sector is the accumulation of individual sector-specific productivities:

$$L^S_d = \sum_c \sum_e \int_{i:i \in M^S_{ecd}} l^S_{id} \, dF_{ec}(z) \tag{16}$$

Firms are paying the wage rate w^S_d per effective unit of labour supplied. Perfect com-

[18] Lagakos & Waugh (2013) and Adao (2016) expand on this notion more formally. Lagakos & Waugh (2013) also provides an alternative multivariate characterisation of the Fréchet distribution using the Frank copula and arbitrary correlation between productivity draws. This version does *not* automatically generate alignment between comparative and absolute advantage. The aim of their analysis is indeed to allow for different cases and let the data in their global cross-country analysis decide whether they find evidence of such alignment. This turns out to be the case in their empirical analysis.

petition among firms within a sector and destination implies that prices equal marginal cost:

$$p_d^S = \frac{w_d^S}{A_d^S} \tag{17}$$

5.2. Equilibrium and Conflict Impact

Market clearing and equilibrium. Market clearing implies that total production equals total consumption for each unique sectoral good in a destination:

$$A_d^S L_d^S = C_d^S = \frac{(p_d^S)^{-\sigma}}{P^{1-\sigma}} GDP \tag{18}$$

where C_d^S denotes total consumption of the good produced in destination d and sector S and GDP denotes GDP or total income in the economy. The second equality comes from CES preferences over all goods.

Using the definition of individual labour supply 10 and the characterisation of average productivity in a destination-sector 13, we can reformulate total labour supply in a destination-sector as a function of labour movement across space π_{ecd}^S and exogenous parameters only:

$$L_d^S = \sum_c \sum_e \overline{N_{ec}} (\pi_{ecd}^S)^{1-\frac{1}{\theta}} h_e^S q_{ecd}^S \Gamma \left(\frac{\theta - 1}{\theta} \right) \tag{19}$$

where $\overline{N_{ec}}$ is the birth population in origin c with education level e. Noting that the movement probabilities π_{ecd}^S are itself a function of exogenous parameters and endogenous wage rates w_d^S, substituting equation 17 in for prices and taking the ratio of the market clearing conditions across a sector or destination, the model can be solved as the following system of equations in endogenous destination-sector labour unit wage rates:

$$\left(\frac{w_d^S}{w_f^T} \right)^{\sigma} = \left(\frac{A_d^S}{A_f^T} \right)^{\sigma - 1} \frac{L_f^T}{L_d^S} \quad \forall S, T \in \{A, N\} \; \forall d, f \in \mathcal{K} \tag{20}$$

The impact of conflict. In this economy, conflict is considered to have an impact on the following sets of parameters. First, it may affect education. The destruction of schools and killings of teachers can plausibly result in individuals losing out on education in their chiefdom of origin. If individuals have a worse realisation of education e, this will directly translate into lower human capital $h_e^S = \exp(\phi^S e)$. To what extent human capital suffers from the education reduction will be governed by the returns to education in each sector ϕ^S.

Second, conflict may affect firm-level productivities A_d^S in each sector. In the model, this parameter would essentially capture any sector-specific determinant of productivity. For example, it could entail the destruction of essential infrastructure for production in a sector. It could also capture immobile physical capital. I do not model physical capital here explicitly. To the extent that physical capital is immobile, however, this would enter the model in exactly the same way as A_d^S. Any destruction of physical capital without reconstruction in the long run would therefore be captured by a reduction of A_d^S.

Third, conflict may have an impact on the amenities of a chiefdom a_d. If fighting leads to the destruction of local public goods or a generally risky environment to live in, this would be captured by the amenities in the model.

Changes to these parameter have direct implications for income of workers and indirect implications through changing the allocation of labour across sectors and space. Figure 5 shows these relationships in a diagram. The market clearing condition 20 reveals that changes to firm-level productivities have a first-order impact on labour unit wage rates while changes to education and amenities will only affect wages through the general equilibrium channel by affecting labour supply in each sector and destination. This means that shocks to firm-level productivities have first-order impacts on income as well as sector and destination choice through their direct impact on wage rates.[19]

The effect on income is rather straightforward. Worker income is a direct function of the sector-destination wage rate w_d^S and human capital h_e^S. Hence, reductions in education and firm productivities will directly decrease income.

For the effect on sector and destination choice, it is instructive to consider the ratio of π_{ecd}^S across sectors and locations to see how conflict changes the labour allocation across sectors and space:

$$\frac{\pi_{ecd}^A}{\pi_{ecd}^N} = \left(\frac{w_d^A \exp(\phi^A e)(1-\tau_{ecd}^A)\epsilon_{ecd}^A}{w_d^N \exp(\phi^N e)(1-\tau_{ecd}^N)\epsilon_{ecd}^N}\right)^\theta \tag{21}$$

$$\frac{\pi_{ecd}^S}{\pi_{ecf}^S} = \left(\frac{a_d w_d^S (1-\tau_{ecd}^S)\epsilon_{ecc}^S}{a_c w_c^S (1-\tau_{ecc}^S)\epsilon_{ecc}^S}\right)^\theta \tag{22}$$

[19]The intuition for this positive first-oder impact of A_d^S on w_d^S is straight-forward. Consider an increase in A_d^S. As a result, marginal cost will decrease, and prices for the good will drop. Demand rises and the wage will adjust on the labour market to enable increased production.

As equation 21 reveals, education will have a direct impact on sector choice within a location to the extent that returns to education differ by sector. In particular, if $\phi^N > \phi^A$, workers are more likely to work in agriculture as conflict negatively affects their education since education losses translate into a greater human capital loss in sector N.

Through its first-order impact on wage rates w_d^S, any changes in firm-level productivities will also affect sector choice to the extent that one sector is hit harder than the other. If non-agricultural firm productivity decreases to a greater extent than agricultural firm productivity, this will lead to a decrease in the relative wage w_d^N/w_d^A with the implication that individuals will move out of sector N into sector A.

Changes in amenities do not have a direct effect on the sector composition but they directly affect location choice. Equation 22 represents the probability of leaving relative to the probability of staying at home within a given sector S. Considering a conflict effect at origin, as a_c decreases workers are more likely to leave. Similarly, if firm productivities at home go down wage rates w_d^S decrease which encourages leaving.

Education losses can lead to direct effects on location choice if moving cost differs by education level. In particular, if moving is more costly for the uneducated $\tau_{0cd}^S > \tau_{1cd}^S$ a decrease in education will encourage more stayers. This effect goes in the opposite direction of the amenity and firm productivity effect. Apart from potentially changing the number of movers, the interplay of education, firm productivity and amenity changes as a result of conflict may also change the selection of movers. How the selection changes would, for example, depend on whether and to what extent moving cost differentials between educated and uneducated people exist in the two sectors.

6. Model Estimation

In order to test directly how these different parameters of the model are affected, I need to estimate the parameters of the model. While education outcomes are observed in the data, firm-level productivities and amenities are not. The structure of the model allows for a recursive estimation strategy in four steps, three of which are broadly based on the estimation strategy used by Bryan & Morten (2019). First, I use a measure of observed income variance to identify the Fréchet shape parameter θ that captures the dispersion of individual productivities. Second, the variation of income across sectors, space and education levels identifies labour unit wage rates w_d^S, returns to education ϕ^S and the common component of the Fréchet scale parameter by origin and education $\overline{q_{ec}}$. The model definition of average income (equation 14) is used as a regression equation in this step. Third, from estimated wage rates in each location and sector w_d^S I can infer firm-level productivities A_d^S using the market clearing conditions. Fourth, I make use of the fact that location and sector choice depend on amenity and wage rate differences across space and sectors as well as migration cost (equation 12). Conditional on estimated wage rates w_d^S, observed migration flows and sector choices identify amenities a_d and migration cost τ_{ecd}^S. This section explains these steps in detail.

Step 1: Estimating productivity dispersion θ. The Fréchet distribution permits the following expression of moments of the average income:

$$\frac{var[m_{ecd}^S]}{\overline{m_{ecd}^S}^2} = \frac{\Gamma(\frac{\theta-2}{\theta})}{\Gamma(\frac{\theta-1}{\theta})^2} - 1 \tag{23}$$

To use this and further relationships from the model in estimation, I construct a dataset in which the unit of observation is a (c, e, d, S) cell on the migration matrix. Throughout, the education realisation is a simple binary indicating whether workers have primary school or not. For each origin-education pair (151 origin chiefdoms $\times$ 2 education levels), I consider outcomes in each destination-sector pair (151 destination chiefdoms $\times$ 2 sectors).[20]

The left-hand side of equation 23 is the squared coefficient of variation of income. How income varies within a destination-sector for people sharing the same origin and education level identifies the dispersion of individual productivities. The intuition is that this group for workers faces the same wage rates, human capital and average individual productivity and is selected in the same way. Therefore, the only element that can explain how their income varies is variation in their individual productivities.

Using the observed moment on the left-hand side in each (c, e, d, S) cell, I perform a general method-of-moment estimation of θ. For the second moment on the Fréchet distribution to exist, θ needs to exceed 2. In this admissible range of θ values greater than 2, the function has a unique solution, that is, I can uniquely identify the parameter.

Step 2: Estimating wage rates w_d^S, education returns ϕ^S and average individual productivities $\overline{q_{ec}}$. Taking the logarithm of equation 14 yields the following regression equation:

[20] The 151 chiefdoms consist of 146 observed chiefdoms (5 missing in the 2018 household data, 3 missing in the 2011 household data with conflict information) in all but the Western Region, the four chiefdoms in the Western Rural District and taking the capital Freetown as one chiefdom.

$$\ln \overline{m^S_{ecd}} = \underbrace{\ln \Gamma\left(\frac{\theta-1}{\theta}\right) w^S_d}_{\text{Destination-sector FE}} - \frac{1}{\theta}\ln \pi^S_{ecd} + (\phi^A - \phi^N)(e \times I^A) + \underbrace{\phi^N e + \ln \overline{q_{ec}}}_{\text{Origin-education FE}} + \ln \epsilon^S_{ecd} \quad (24)$$

where I^A is a dummy variable for the agricultural sector. Average wages for workers with origin-education (c, e) who choose destination-sector (d, S) is determined by the wage rate in that destination-sector, their human capital and the average individual productivity draw $q^S_{ecd} = \overline{q_{ec}}\epsilon^S_{ecd}$. The share of workers π^S_{ecd} with origin c and education e who choose destination d and sector S enters as a selection term: The more people make that choice, the lower the average productivity (see discussion above and Figure 2). Observing π^S_{ecd} and having estimated θ, we can fully control for this selection mechanism by plugging this information into the equation. Taking $\ln \epsilon^S_{ecd}$ with mean 0 as an error term, a regression of $\ln \overline{m^S_{ecd}} + \frac{1}{\theta}\ln \pi^S_{ecd}$ on destination-sector fixed effects, origin-education fixed effects and observed education realisation times sector choice identifies the wage rates w^S_d, origin-education component of the average individual productivity $\overline{q_{ec}}$, and differential returns to education in the two sectors $\phi^A - \phi^N$.[21] Intuitively, controlling for the selection into a destination and sector, variation in income across space, sector and education levels identifies the wage rates, average individual productivities and returns to education.[22] Using the θ estimates from Step 1, I can recover estimates of w^S_d from the destination-sector fixed effects.

Step 3: Computing labour supply L^S_d and firm productivity A^S_d. Using estimates from the previous steps and observed origin populations by education level $\overline{N_{ec}}$ allows me to compute labour supply in each destination and sector as stated in equation 19. Having computed labour supplies and drawing on my w^S_d estimates from Step 2, I only need an

[21]With sector-origin fixed effects, I will not be able to separately identify ϕ^A, ϕ^N and $\overline{q_{ec}}$. Instead, I can identify $\phi^A - \phi^N$, $\ln \overline{q_{0c}}$ for the uneducated and $\phi^N + \ln \overline{q_{1c}}$ for the educated. For all further estimations and simulations, however, this information suffices since average individual productivities $\overline{q_{ec}}$ and human capital $\exp(\phi^S e)$ only enter jointly in these calculations.

[22]Destination-sector fixed effects and origin-education fixed effects are only identified up to scale relative to each other since origins and destinations are the same locations. As a normalisation, I choose the origin-education fixed effect of Freetown to be 0 with the implication that $\exp(\phi^N)\overline{q_{1,Freetown}} = 1$. Therefore, all other origin scale parameter average estimates $\overline{q_{ec}}$ are evaluated relative to the distribution for educated people in the capital. To the extent that ϵ^S_{ecd} differs by sector on average, this will be captured in the destination-sector fixed effects and thus the estimate of w^S_d. In particular, the estimate of wage rates is scaled by the sector average of ϵ^S_{ecd}. Since this estimate is used in the following to find firm productivities A^S_d, the latter estimate is also scaled with the implication that its logarithm $\ln A^S_d$ is shifted by a constant. However, this is irrelevant for the simulations since the necessary information that informs the counterfactual simulations is the relationship between $\ln A^S_d$ and conflict. A constant shift of $\ln A^S_d$ does not change that relationship.

estimate of the elasticity of substitution σ to proceed with the identification of firm-level productivities A_d^S. I borrow the value from the literature and set $\sigma = 8$.[23]

With these values in hand, firm-level productivities are identified up to scale in the market clearing conditions 20, restated here for ease of reference:

$$\left(\frac{w_d^S}{w_f^T}\right)^{\sigma} = \left(\frac{A_d^S}{A_f^T}\right)^{\sigma-1} \frac{L_f^T}{L_d^S} \quad \forall S, T \in \{A, N\} \; \forall d, f \in \mathcal{K} \tag{25}$$

I choose to normalise $A^N_{Freetown} \equiv 1$ and can then recover A_d^S for each destination and sector one-by-one. The normalisation essentially implies that all values of A_d^S are measured relative to the non-agricultural productivity in the capital Freetown.[24]

Step 4: Estimating amenities a_d and migration cost τ^S_{ecd}. In order to estimate amenities and movement cost across sectors and destinations, I consider the share of leavers within a sector relative to stayers and the sector share differential within a location. This yields the following equations (the logarithms of equations 22 and 21):

$$\begin{aligned}\frac{\ln \pi^S_{ecd} - \ln \pi^S_{ecc}}{\theta} &= \ln a_d - \ln a_c + \ln w_d^S - \ln w_c^S \\ &\quad + \ln(1-\tau^S_{ecd}) - \ln(1-\tau^S_{ecc}) + \ln \epsilon^S_{ecd} - \ln \epsilon^S_{ecc}\end{aligned} \tag{27}$$

$$\begin{aligned}\frac{\ln \pi^A_{ecc} - \ln \pi^N_{ecc}}{\theta} &= \ln w_c^A - \ln w_c^N + (\phi^A - \phi^N)e \\ &\quad + \ln(1-\tau^A_{ecc}) - \ln(1-\tau^N_{ecc}) + \ln \epsilon^A_{ecc} - \ln \epsilon^N_{ecc}\end{aligned} \tag{28}$$

Using estimated θ, w_c^S and observed migration flows across sectors and destinations π^S_{ecd}, equation 27 identifies amenities up to scale. I treat the whole expression on the bottom involving τ and ϵ as a residual in a regression. The amenities are coefficients

[23] There are not many estimates of this elasticity of substitution. I follow Bryan & Morten (2019) and Allen & Arkolakis (2014) who, like in my model, have an elasticity of substitution of Armington goods across space within a country. Hsieh et al. (2019) and Bernard et al. (2003) use values between 3 and 4 for the substitution between goods across industries and countries, respectively. I discuss robustness of my main results to different values of σ between 2 and 10 in section 8.

[24] An introduction of trade cost in the classic iceberg format would have an implication for the estimation of A_d^S in this step. Denote the iceberg trade cost by $\eta^S_{cd} \geq 1$ as the amount of good produced in d that would need to be bought in any other location g to consume one unit of that good in g. The market clearing conditions would contain a market access term:

$$\left(\frac{w_d^S}{w_f^T}\right)^{\sigma} = \left(\frac{A_d^S}{A_f^T}\right)^{\sigma-1} \frac{\sum_g (\eta^S_{dg})^{-\sigma} GDP_g / P_g^{1-\sigma}}{\sum_g (\eta^S_{fg})^{-\sigma} GDP_g / P_g^{1-\sigma}} \frac{L_f^T}{L_d^S} \tag{26}$$

Market access is the weighted sum of real GDP in all buying locations of a good that is produced in d or f where the weights are the inverse trade cost. With trade cost, therefore, my estimates of firm productivities A_d^S are in fact a composite of firm productivities and market access. Section 8 discusses the implications for the results in counterfactual simulations and the estimation of the aggregate effects of civil war.

on the regressor $I_{d=k} - I_{c-k}$ which is a destination minus an origin dummy variable. Intuitively, controlling for wage rate differences across space, migration flows are governed by amenity differences between locations and migration cost. To the extent that migration cost is symmetric, amenity differences and migration cost can be separately identified because amenity differences will shape uni-directional flows while migration cost affects bi-directional flows. If amenities in location d are much larger than in location c, few people move from d to c while many people make the opposite move. If migration cost between the two locations is large, few people move both ways. Asymmetries in migration cost between two locations, that is $\tau^S_{ecd} \neq \tau^S_{edc}$, will be loaded onto the amenity estimate. Since amenities are the same across education levels and sectors within a location while migration cost is education- and sector-specific, the amenity estimate will be a composite of amenities and the average asymmetry in migration cost across education and sectors in that case. This only affects the interpretation of the estimate, but not any substantial results in counterfactual simulations performed below.

In the same spirit as before with firm-level productivities, I normalise everything against the capital and set $a_{Freetown} \equiv 1$. To be able to make use of information when $\pi^S_{ecd} = 0$ and deal with potential heteroskedasticity issues introducing bias in this log-specification, I follow Silva & Tenreyro (2006) and make use of Poisson pseudo-maximum likelihood (PPML) estimation for this equation.[25]

In order to uncover migration cost τ^S_{ecd}, the residuals from this regression can be computed. Focusing on the τ elements in the error term, we note that we will identify movement cost only relative to relative to the cost of staying at home within a sector. The cost to enter a particular sector may be different for the agricultural and non-agricultural sector. In order to capture systematic differences between $1 - \tau^A_{ecc}$ and $1 - \tau^N_{ecc}$, I estimate equation 28 which will have the difference between these two terms as a residual. Using the residuals from this regression and combining it with the first regression residuals, I can identify τ^S_{ecd} for both sectors relative to the cost of staying at home and working in agriculture τ^A_{ecc}.[26]

[25] The moments corresponding to the score vectors in this ML estimation are:

$$E\left[\left(\left(\frac{\pi^S_{ecd}}{\pi^S_{ecc}}\right)^{\frac{1}{\theta}} \frac{w^S_c}{w^S_d} - \exp\left(\ln a_d - \ln a_c\right)\right)\left(I_{d=k} - I_{c=k}\right)\right]$$

where $I_{d=k} - I_{c=k}$ is a destination minus origin dummy variable.

[26] The simple adjustment of the first regression residuals for non-agriculture would be subtract the residuals from the second regression since $\ln(1-\tau^N_{ecd}) - \ln(1-\tau^A_{ecc}) = \ln(1-\tau^N_{ecd}) - \ln(1-\tau^N_{ecc}) - \left(\ln(1-\tau^A_{ecd}) - \ln(1-\tau^N_{ecc})\right)$. I have abstracted from the ϵ elements here which are random varia-

7. Estimation Results

7.1. Model Parameters

Table 10 shows the estimated model parameters. The dispersion parameter θ is estimated to be 4.729. In relation to Bryan & Morten (2019) who estimate a very similar model, I find parameters that compare plausibly to theirs. Their equivalent estimate of θ is 3.18 for Indonesia and 2.69 for the US. The degree of productivity dispersion this measures can be interpreted as the degree of similarity of goods produced in different locations. It would therefore be expected that less developed countries with a smaller range of varied goods produced would have a smaller degree of dispersion which is reflected in a higher θ realisation.

The returns to education estimates show a great differential by sector. This estimate stems from a regression equation 24 with log wage as an outcome variable and can therefore be interpreted as a relative return. Returns to primary school education in non-agriculture exceed returns to education in agriculture by 14 percentage points. Considering migration cost by sector and education level, there is a significant difference by sector. It turns out to be much more costly for people to enter the non-agricultural sector and move somewhere than the agricultural sector. Relative to Bryan & Morten (2019) who find 39% migration cost in Indonesia, I find generally higher migration cost in Sierra Leone. The average across both education levels and sectors is 56%. Like in their case, I find a strong positive correlation between my estimate of τ and log distance between two places. Figure 8 displays the strong positive relationship with distance which suggests that the estimated parameters capture something real about the cost of moving. Another interesting fact about estimated moving cost is that there is a clear differential by education level in the agricultural sector. Uneducated people have to forego 19 percentage points more of their income when leaving one's origin than educated people in that sector. No such difference exists for non-agricultural workers. This could reflect a dissimilarity in crops across space whereby educated farmers are better able to adapt to different crops. Alternatively, this may also be due to an information barrier or the degree of connections to other chiefdoms.

tions with a mean of 1. Strictly speaking, when computing the residuals as described and taking the exponential, I would identify $(1-\tau^S_{ecd})\epsilon^S_{ecd}/(1-\tau^A_{ecc})\epsilon^A_{ecc}$. When using these estimates going forward in a model simulation, however, they only enter in the expression for π^S_{ecd} both in the numerator and the denominator (cf. equation 12). The origin-agriculture specific element in the denominator of the identified residual would cancel out in the expression of π^S_{ecd}. Furthermore, ϵ^S_{ecd} actually appears multiplicatively in π^S_{ecd}, so this residual is exactly what we need. Since it only captures mean one random variation, however, to fix ideas it is easier to abstract from ϵ^S_{ecd} when we conceptually think about the estimation.

Educated farmers may be better informed about crops or opportunities for agricultural work elsewhere or better connected to farmers away from home. To the extent that integration into a new community is essential in order to work there, the latter point may be particularly important.

The estimates of labour unit wage rates by sector are shown in Figure 9 relative to the distribution of observed income. In line with the income distribution, labour unit wage rates in non-agriculture are higher than in agriculture. Interestingly, their variation is also greater. Generally, there is substantial variation in wage rates. The highest wage rates differ from the lowest wage rates by a factor of more than ten. Firm-level productivities vary similarly substantially. Their distribution by sector is shown in Figure 10. Average agricultural firm-level productivity is 16% of the non-agricultural productivity in the capital and average non-agricultural firm-level productivity is 20% of non-agricultural firm-level productivity in the capital. To assess whether these estimates are capturing something real about firm-level productivities in Sierra Leone, I correlate them with observed variables related to firm-level productivity in the data. Figure 11 demonstrates the result. Agricultural productivity estimates are strongly positively correlated with access to agricultural drying or storage space. Similarly, non-agricultural productivity estimates are highly positively correlated with measures of the existence and extent of electricity connection.

Finally, Figure 12 displays the distribution of amenities across space. Compared to wage rate and firm-level productivities, amenities vary considerably less space. Relative to the capital Freetown, average amenities are 13% smaller. In a similar exercise as before, I also test the plausibility of these estimates by correlating them with observed variables in the data that plausibly reflect amenities. From information on access to eight public goods and services, I construct two public good indices.[27] I also consider phone coverage and recharge possibilities within chiefdoms. With all these measures, amenities are strongly positively correlated as Figure 13 shows.

7.2. Impact of Conflict

With these parameter estimates in hand, we can proceed to estimate the impact of conflict on firm-level productivities and amenities directly. Since conflict variation is not

[27] In particular, the public goods indices measure whether, on average, the following public goods are within 30 or 60 minute reach: (i) supply of drinking water, (ii) a food market, (iii) public transportation, (iv) a priamry school, (v) a secondary school, (vi) a health clinic, (vii) a hospital, and (viii) an all year motorable road.

exogenous, I employ the same IV identification strategy as the reduced-form analysis. In particular, the reduced-form results suggest that conflict intensity is higher in places that have higher income to begin with since the OLS results are considerably weaker than the IV results. In the context of the model, higher initial income in locations that experience conflict is plausibly related to greater productivity of firms in such places before the war. Therefore, when investigating the effect of conflict on firm productivities, a simple OLS estimate would suffer from selection bias.

Using distance to the Liberian border as an instrument for conflict intensity, I can identify the causal effect of conflict on firm productivities and amenities. Table 11 shows the results at the chiefdom level.[28] The effect of conflict on agricultural firm productivity and amenities is small and insignificant. While the effect of conflict on firm productivity in sector N is also insignificant, its size is non-negligible. An increase in conflict intensity by one standard deviation would lead to a decrease in non-agricultural firm productivity by 7.7% according to this estimate. Since this is not a trivial decrease and given that the IV estimation procedure is quite imprecise here, I am cautious not to simply interpret this as a zero effect owing to a lack of significance. In the following, I will provide counterfactual simulations both with and without considering an effect on non-agricultural firm productivity.

8. Counterfactual Simulations

Having estimated all the parameters of the model, I can perform counterfactual simulations. These simulations serve two purposes. First, by simulating away the conflict in the whole economy, I can generate a true counterfactual of what the entire Sierra Leonean economy would look like today in the absence of civil war. This allows for the estimation of *aggregate* effects. If people move in response to the war, the reduced-form analysis does not provide such a counterfactual. The reduced-form analysis IV analysis essentially considers chiefdoms that are farther away from the Liberian border a counterfactual for chiefdoms closer to the border. However, population movement in response to the war implies that any difference between such chiefdoms captures both the direct effect of conflict as well as migration. In order to generate a no-war counterfactual of the Sierra Leonean

[28]Some firm productivities cannot be estimated since no one actually works in that sector and destination. Hence, we do not have any income information for those sector-destinations. This applies to one location in agriculture and 13 locations in non-agriculture. These chiefdoms are excluded from the analysis of firm productivities here. In the simulations, the firm productivity for those sector-destinations is set equal to zero to rationalise no worker choosing these pairs.

economy, the first simulation reverts all effects that the war has on model parameters. I identified two such effects: first, a reduction in education and second, a reduction in firm productivity in the non-agricultural sector. The IV estimates inform the size of the conflict effect. Figure 14 displays how this scenario changes the underlying distribution of (un-)educated origin populations as well as non-agricultural firm productivity A_d^N. By reverting the effect of the war, chiefdoms that experienced conflict have a larger educated population and greater non-agricultural firm productivity in the simulated counterfactual.

Unlike a partial equilibrium exercise in which everything else is kept constant, the simulation will generate a new counterfactual in general equilibrium. The model will trace through both the direct implications of these parameter changes for aggregate income as well as the indirect implications by changing the allocation of labour across sectors and space. A greater share of educated people in a chiefdom leads to a greater share of workers in the non-agricultural sector in which returns to education are larger than in agriculture. In the agricultural sector in which migration cost is much lower among the educated, it also results in more leavers in locations with higher human capital. By contrast, higher firm productivity in non-agriculture attracts more stayers and in-migrants in that sector.

Second, by simulating away only the effect of conflict on education, I can assess the quantitative importance of the human capital effect of war relative to firm productivity losses. In this scenario, only the educated populations in affected chiefdoms is increased while firm productivity stays at the level at which it is everywhere. The simulation traces through both the direct effect of human capital increases on income and indirect implications by re-allocating labour across sectors and space in response to higher levels of human capital in some locations.

I find evidence of a strong effect of civil war on aggregate income in the first scenario. In the absence of the war, aggregate income in Sierra Leone would be 14.8% higher today. With lower pay in agriculture, an aggregate sector shift seems to be an important driver of this result. The economy-wide share of people in agriculture would be 6.3 percentage points lower. These are substantial long-term consequences of the civil war that suggest that the country is farm from having recovered from its violent past even twenty years after fighting ended.

There are four limitations to this analysis that qualify the result. First, the identification of this aggregate effect implicitly assumes that some chiefdoms are not at all affected by the war. The identification of how conflict affects key parameters of the model rests on within-country comparisons of chiefdoms that are closer to or farther from the Liberian border and therefore experience conflict to varying degrees. In the counterfactual analysis, chiefdoms that are very far from the border with a zero realisation on the conflict measure are assumed not to experience any effect of civil war. If those chiefdoms suffer in some way from the war, my estimate would be an underestimate of the true aggregate effect.

Second, the estimate captures the net effect of civil war in the long run after post-conflict interventions between the end of the war and 2018. Since the international community was engaged in reconstruction work after the war, the aggregate effect I present is not the pure effect of civil war but rather how the economy still suffers after taking into account reconstruction efforts. The pure effect of the war would therefore be weakly greater than my estimate.

Third, if one were to relax the assumption of costlessly tradable goods across space, the actual effect of conflict on firm productivities may be smaller and my aggregate income effect thus be overestimated. With trade cost, my estimates of firm productivities would capture a combination of these productivities and market access of a location. Market access is the weighted sum of all other locations' real GDP where the weights are inverse trade cost (cf. equation 26). Part of the conflict effect on firm productivities that I estimate could stem from a reduction in market access. If market access is lower because trade cost increased as a result of the war, my estimate of the aggregate income effect is not necessarily different but the mechanisms driving this result are – part of the effect would be due to increased trade cost rather than reduced firm productivity. However, if market access is lower because economic performance in close chiefdoms with low trade cost contracted, this contraction should not be loaded on a reduction in the firm productivity parameter. As a result, the actual effect on firm productivities would be lower and the aggregate income effect would be smaller. However, with positive trade cost market access would contaminate the firm productivity estimates for both agriculture and non-agriculture. Since I do not find any effect of conflict on agricultural firm productivity it is unlikely that the latter channel plays a major role. Therefore, any potential upward bias from assuming free trade across chiefdoms is likely to be small.

Fourth, while the aggregate income results are robust to different values of the elasticity of substitution σ, the sector shift is more sensitive to the choice of that value. Figures 20 and 21 show the aggregate income and sector shift changes for a range of σ values between 2 and 12. The sector shift is stronger the higher the value. This is intuitive since the elasticity of substitution governs the trade-off between two competing effects on the sector allocation as productivity changes. As productivity in sector N decreases relative to sector A, the non-agricultural good gets more expensive which leads to lower demand. Hence, production in that sector goes down. At the same time, lower productivity implies that the labour input needed to produce a given amount is larger. The elasticity of substitution determines the strength of the first effect. The larger the elasticity, the stronger the substitution effect and the more people will shift sector to accommodate changing demand.[29] The relevant literature does not use a value below $\sigma = 4$. Even in that case, my result on the sector shift is still sizeable with the share of agriculture increasing by 4.5 percentage points. The aggregate income increase is 13.9%, very close to my original result.

In contrast to the results in the first scenario, the differences in aggregate income and the share of agricultural workers are a lot smaller in the second simulation, albeit still sizeable. Only reverting the human capital loss resulting from war, aggregate income in the country would be 4.7% higher today and the economy-wide share of workers in agriculture 1.4 percentage points lower. This means that the human capital loss can account for about 1/3 of the total long-run effect of civil war in Sierra Leone relative to a reduction in non-agricultural firm productivity.

These results mask substantial heterogeneity. Considering the full reversal of the war, Figure 15 displays the aggregate income increases in each chiefdom with the largest gainers experiencing a 31% increase in income. While the aggregate income effect of civil war in many chiefdoms is sizeable, it is substantially smaller than what the reduced-form estimates would suggest. In Figure 16, I show the reduced-form results in each chiefdom that are estimated off spatial income differences. These results are created by taking the IV coefficient estimate of the effect of conflict intensity on worker income and scaling it by the intensity of conflict that each chiefdom experienced. These reduced-form estimates go up to 71% income differences in the highest cases and are considerably larger than the aggregate income differences throughout. The fact that spatial income differences exceed aggregate income differences suggests that selective migration in response to conflict matters a great deal.

It is therefore worth considering the nature of this migration response. Changing both human capital and non-agricultural firm productivity in the model generates no correlation between conflict intensity and population change in a chiefdom. Figure 17 displays

[29] As long as the $\sigma > 1$, the result on the sector shift only changes quantitatively. If $\sigma < 1$, the first effect is weaker than the second effect and a productivity decrease in sector N implies more people working in that sector. This is indeed Baumol's cost disease argument and shown formally in Ngai & Pissarides (2007). However, empirically, the relevant literature for the elasticity of substitution used in my context suggests values that are considerably larger than 1.

this relationship. By contrast, as demonstrated in Figure 18, changing human capital only generates a positive relationship between conflict intensity and population change in a chiefdom. The changes considered here are negative changes, that is, (partial) simulations of conflict relative to a no-war scenario. These correlations are driven by two opposing forces on population changes. One is that a drop in firm productivity encourages leaving. The other one is that a drop in human capital encourages staying since moving cost is higher for the uneducated. It seems that these two forces cancel each other off in the first scenario.

This result corresponds to the empirical reality which implies that both human capital and non-agricultural firm productivity decreases as a result of civil war seems to be the appropriate scenario to capture the full effect of conflict. Empirically, I can use the IV identification strategy and population information from 1985 before the war and 2015 after the war in each chiefdom to test what the impact of conflict on population changes is. Figure 19 depicts the relationship between the instrument and population changes. There is clearly no relationship between the two and hence no long-run effect of conflict on population changes. This result is in line with Davis & Weinstein (2002) who find that population densities in cities in Japan recovers from bombing in the long run.

However, while this suggest that the migration flows are unaffected by the war, it does not speak to the *composition* of movers that may change as a result of the war. A look at moving cost will be informative in this regard. Moving costs are higher for the uneducated, but only in the agricultural sector. For agricultural workers, it is on average 19 percentage points more costly to move if they are uneducated. For non-agricultural workers, there is no significant difference in moving cost (see Table 10). Therefore, as conflict affects education and sector composition, it will have implications for the composition of movers. Human capital losses and firm productivity losses push individuals back into agriculture and have a greater share of uneducated people as a consequence. While uneducated agricultural workers face higher moving cost and are less likely to move, non-agricultural workers leave in response to conflict. This selective migration response to the war implies that spatial income differences observed after the war overstate the aggregate income effect of the war.

9. Conclusion

This paper investigates how civil war has persistent effects on the economy in Sierra Leone. To this end, I first establish that conflict led to large spatial differences in income between areas that were more and less affected. These seem to be driven by a sector shift: People are more likely to work in agriculture as a result of the war. While I can identify these spatial differences with an identification strategy that deals with the non-random placement of conflict, they still capture both the direct effect of conflict as well as general equilibrium forces. The spatial differences are therefore not necessarily reflective of changes in aggregate income or sector composition.

In order to make progress on estimating aggregate effects, this paper develops an economic geography model. To keep track of population movement and its general equilibrium implication in response to the war, labour mobility under migration cost is a key feature of the model. Individuals can also choose their sector of work subject to a productivity draw for each sector and destination. In the model, conflict can change human capital, firm productivities and amenities. This affects aggregate income both directly and indirectly through changing the sector composition and spatial allocation of labour.

The key parameters of the model can be estimated in a simple recursive procedure. In particular, observed income and migration flows identify firm productivities in each sector and destination as well as location amenities. While education outcomes are directly observed in the data these parameters are not. Having estimated them allows me to shed light on such potential mechanisms of a conflict impact as well. I find that education and firm productivity in the non-agricultural sector are persistently and strongly affected by the war while amenities and agricultural firm productivity are not.

Finally, these results can be taken forward for counterfactual simulations of the Sierra Leonean economy in the absence of the war. A full reversal of the war scenario involves reverting both the human capital loss and non-agricultural firm productivity loss. In this scenario, Sierra Leonean aggregate income would be 14.8% higher today and the economy-wide share of workers in agriculture 6.3 percentage points lower. By contrast, if only the human capital loss is reverted, aggregate income would be 4.7% higher today and the share of agricultural workers 1.4 percentage points lower. While the latter effects are still sizeable, they also suggest that firm productivity effects of civil war play a greater role in explaining aggregate income losses than human capital loss in Sierra Leone. Identifying exactly the relevant elements for firm productivity that are potential drivers behind this and using a similar model structure to learn about the aggregate effect and mechanisms of conflict impact in other settings are left as promising avenues for future research.

References

Acemoglu, D., Reed, T., & Robinson, J. A. (2014). Chiefs: Economic development and elite control of civil society in Sierra Leone. *Journal of Political Economy*, 122(2), 319–368. [Cited on pages 9 and 11.]

Adao, R. (2016). Worker heterogeneity, wage inequality, and international trade. *Working Paper, University of Chicago*. [Cited on page 20.]

Akbulut-Yuksel, M. (2014). Children of war: The long-run effects of large-scale physical destruction and warfare on children. *Journal of Human Resources*, 49(3), 634–662. [Cited on page 5.]

Akresh, R. & De Walque, D. (2011). Armed conflict and schooling: Evidence from the 1994 Rwandan genocide. *World Bank Policy Research Working Paper*. [Cited on page 5.]

Alesina, A. & Perotti, R. (1996). Income distribution, political instability, and investment. *European Economic Review*, 40(6), 1203–1228. [Cited on page 4.]

Allen, T. & Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3), 1085–1140. [Cited on pages 16 and 26.]

Amior, M. (2015). Why are higher skilled workers more mobile geographically? the role of the job surplus. *Working Paper, LSE*. [Cited on page 17.]

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers*, 16(1), 159–178. [Cited on page 15.]

Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2), 407–443. [Cited on page 4.]

Bellows, J. & Miguel, E. (2009). War and local collective action in Sierra Leone. *Journal of Public Economics*, 93(11), 1144–1157. [Cited on pages 7, 8, 9, 11, and 13.]

Bernard, A. B., Eaton, J., Jensen, J. B., & Kortum, S. (2003). Plants and productivity in international trade. *The American Economic Review*, 93(4), 1268–1290. [Cited on page 26.]

Bozzoli, C., Brück, T., & Wald, N. (2012). Self-employment and conflict in Colombia. *Journal of Conflict Resolution*, 57(1), 117–142. [Cited on page 5.]

Brakman, S., Garretsen, H., & Schramm, M. (2004). The strategic bombing of cities in Germany in World War II and its impact on city growth. *Journal of Economic Geography*, 4(1), 201–218. [Cited on page 5.]

Bryan, G. & Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5), 2229–2268. [Cited on pages 2, 5, 16, 23, 26, and 28.]

Camacho, A. & Rodriguez, C. (2013). Firm exit and armed conflict in Colombia. *Journal of Conflict Resolution*, 57(1), 89–116. [Cited on page 5.]

Collier, P. (1999). On the economic consequences of civil war. *Oxford Economic Papers*, 51(1), 168–183. [Cited on page 4.]

Collier, P. & Duponchel, M. (2013). The economic legacy of civil war: Firm-level evidence from Sierra Leone. *Journal of Conflict Resolution*, 57(1), 65–88. [Cited on page 5.]

Davis, D. R. & Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *The American Economic Review*, 92(5), 1269–1289. [Cited on pages 5 and 34.]

Eaton, J. & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5), 1741–1779. [Cited on pages 5 and 16.]

Felkner, J. S. & Townsend, R. M. (2011). The geographic concentration of enterprise in developing countries. *The Quarterly Journal of Economics*, 126(4), 2005–2061. [Cited on page 10.]

Galdo, J. (2013). The long-run labor-market consequences of civil war: Evidence from the Shining Path in Peru. *Economic Development and Cultural Change*, 61(4), 789–823. [Cited on page 5.]

Glennerster, R., Miguel, E., & Rothenberg, A. D. (2013). Collective action in diverse Sierra Leone communities. *The Economic Journal*, 123(568), 285–316. [Cited on page 8.]

Heckman, J. J. & Honore, B. E. (1990). The empirical content of the roy model. *Econometrica*, (pp. 1121–1149). [Cited on page 20.]

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, 87(5), 1439–1474. [Cited on pages 5, 16, and 26.]

Justino, P. & Verwimp, P. (2013). Poverty dynamics, violent conflict, and convergence in Rwanda. *Review of Income and Wealth*, 59(1), 66–90. [Cited on page 5.]

Keen, D. (2005). *Conflict and Collusion in Sierra Leone*. James Currey. [Cited on page 7.]

Kennan, J. & Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1), 211–251. [Cited on page 17.]

La Mattina, G. (2018). How persistent is the effect of conflict on primary education? long-run evidence from the rwandan genocide. *Economics Letters*, 163, 32–35. [Cited on page 5.]

Lagakos, D. & Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2), 948–80. [Cited on page 20.]

Leon, G. (2012). Civil conflict and human capital accumulation: The long-term effects of political violence in Perú. *Journal of Human Resources*, 47(4), 991–1022. [Cited on page 5.]

Miguel, E. & Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, 96(1), 1–15. [Cited on page 4.]

Ngai, L. R. & Pissarides, C. A. (2007). Structural change in a multisector model of growth. *The American Economic Review*, 97(1), 429–443. [Cited on page 33.]

Pettersson, T. & Öberg, M. (2020). Organized violence, 1989-2019. *Journal of Peace Research*, 57(4). [Cited on page 1.]

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics*, 101, 148–167. [Cited on pages 5 and 16.]

Richards, P. (1996). *Fighting for the Rain Forest: War, Youth & Resources in Sierra Leone*. James Currey. [Cited on page 6.]

Richards, P. (2004). The political economy of internal conflict in Sierra Leone. *Working Paper*. [Cited on page 7.]

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2), 135–146. [Cited on page 2.]

Saing, C. H., Kazianga, H., et al. (2017). The long-term impacts of violent conflicts on human capital: US bombing and, education, earnings, health, fertility and marriage in Cambodia. *Working Paper*. [Cited on page 5.]

Serneels, P. & Verpoorten, M. (2015). The impact of armed conflict on economic performance: Evidence from Rwanda. *Journal of Conflict Resolution*, 59(4), 555–592. [Cited on page 4.]

Silva, J. S. & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, 88(4), 641–658. [Cited on page 27.]

UNDP (2006). Case study Sierra Leone: Evaluation of UNDP assistance to conflict-affected countries. *Report*. [Cited on page 6.]

Wozniak, A. (2010). Are college graduates more responsive to distant labor market opportunities? *Journal of Human Resources*, 45(4), 944–970. [Cited on page 17.]

Young, A. (2013). Inequality, the urban-rural gap, and migration. *The Quarterly Journal of Economics*, 128(4), 1727–1785. [Cited on page 5.]

Appendix

A. Tables

Table 1: Summary Statistics

	N	Mean	Std. Dev.
Individual Level			
Female dummy	14532	0.52	0.50
Urban dummy	14532	0.42	0.49
Household size	14532	6.94	3.45
Age	14532	37.9	14.0
Religion is Christianity	14532	0.21	0.41
Religion is Islam	14532	0.78	0.41
Years of schooling	14526	4.21	5.21
Finished primary school	14526	0.37	0.48
Main sector is agriculture	14532	0.54	0.50
Main sector is manufacturing	14532	0.090	0.29
Main sector is services	14532	0.36	0.48
20 day expenditures per worker (in USD)	14405	60.2	61.8
Chiefdom Level			
Conflict	163	0.21	1.01
Distance to border (km)	166	114.9	80.0
Vector ruggedness measure, (3x3 window)	166	0.37	0.081
Average Elevation (km)	166	0.17	0.15
% of chiefdom w/ slope between 2-8%	166	58.5	14.8
% of chiefdom w/ slope between 8-30%	166	19.8	15.3
% of chiefdom w/ slope between 30-45%	166	3.31	5.04
Coarse texture soils (%)	166	7.45	17.3
Medium texture soils (%)	166	89.7	20.2
Soil w/ poor drainage (%)	166	51.7	24.2
Soil w/ excessive drainage (%)	166	2.67	12.0
Minimum distance to five large towns	166	48.6	28.5
School attendance 1989	76	0.28	0.20
School enrollment 1989	76	0.30	0.20
Log p.c. expend. 1989	76	7.94	0.68
Log pop. density 1985	159	3.79	0.78
19th Cen. trading route	154	20.0	19.7
Mining permission 1930	154	0.18	0.38
Hut Tax/Area 1900	89	0.94	1.42
Hut Tax/Pop. 1900	88	0.028	0.030
Population 1963	153	13712.0	9961.6

The individual sample consists of all working individuals in 2018 who were born before the end of the war in 2001. All monetary values are in 2018 USD and the top 1% is truncated. [Back to main]

Table 2: Correlation of Instrument with Pre-war Observables

	(1)	(2)	(3)
	School attendance 1989	School enrollment 1989	Log p.c. expend. 1989
Distance to border	0.00140	0.00128	0.00673
	(0.00119)	(0.00126)	(0.00659)
N	76	76	76
R^2	0.511	0.438	0.326
District FE	✓	✓	✓
	(4)	(5)	(6)
	Log pop. density 1985	19th Cen. trading route	Mining permission 1930
Distance to border	0.00401*	-0.0652	0.00109
	(0.00235)	(0.0492)	(0.00128)
N	159	154	154
R^2	0.280	0.571	0.208
District FE	✓	✓	✓
	(7)	(8)	(9)
	Hut Tax/Area 1900	Hut Tax/Pop. 1900	Population 1963
Distance to border	0.000485	0.0000284	0.954
	(0.00629)	(0.000140)	(24.33)
N	89	88	153
R^2	0.332	0.416	0.199
District FE	✓	✓	✓

Note: All specifications include district fixed effect. Clustered standard errors at the chiefdom level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 3: First Stage

	(1) Conflict	(2) Conflict
Distance to border	-0.0117***	-0.00967***
	(0.00256)	(0.00272)
N	14482	14482
r2	0.733	0.747
F (Kleibergen-Paap)	20.79	12.61
Socio-econ. Controls	✓	✓
Land Controls		✓
District FE	✓	✓

Note: All specifications include district fixed effect. Clustered standard errors at the chiefdom level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 4: Expenditures per workers

	Log household expenditures per worker			
	(1)	(2)	(3)	(4)
Conflict	-0.0507*	-0.0495*	-0.238**	-0.309**
	(0.0301)	(0.0286)	(0.0967)	(0.138)
N	14482	14482	14482	14482
R^2	0.424	0.432	0.407	0.402
Estimation Method	OLS	OLS	IV	IV
F (Kleibergen-Paap)			20.79	12.61
Socio-econ. Controls	✓	✓	✓	✓
Land Controls		✓		✓
District FE	✓	✓	✓	✓

Note – Outcome variables: Log total expenditures per worker. Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 5: Robustness: Expenditures per workers

	Log household expenditures per worker			
	(1)	(2)	(3)	(4)
Conflict	-0.0701**	-0.0699**	-0.236***	-0.278**
	(0.0308)	(0.0296)	(0.0909)	(0.115)
N	13711	13711	13711	13711
R^2	0.426	0.434	0.412	0.415
Estimation Method	OLS	OLS	IV	IV
F (Kleibergen-Paap)			23.59	15.88
Socio-econ. Controls	✓	✓	✓	✓
Land Controls		✓		✓
District FE	✓	✓	✓	✓

Note – *The sample is restricted to chiefdoms that do not have a direct border with Liberia. This excludes 10 chiefdoms from the analysis.* Outcome variables: Log total expenditures per worker. Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 6: Sector Allocation

	Work in Agriculture				Work in Non-agriculture			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conflict	0.0415***	0.0456***	0.121***	0.170***	-0.0415***	-0.0456***	-0.121***	-0.170***
	(0.0149)	(0.0138)	(0.0422)	(0.0569)	(0.0149)	(0.0138)	(0.0422)	(0.0569)
N	14482	14482	14482	14482	14482	14482	14482	14482
R^2	0.528	0.533	0.519	0.515	0.528	0.533	0.519	0.515
Estimation Method	OLS	OLS	IV	IV	OLS	OLS	IV	IV
F (Kleibergen-Paap)			20.79	12.61			20.79	12.61
Socio-econ. Controls	✓	✓	✓	✓	✓	✓	✓	✓
Land Controls		✓		✓		✓		✓
District FE	✓	✓	✓	✓	✓	✓	✓	✓
	Work in Manufacturing				Work in Services			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Conflict	-0.0152	-0.0124	-0.0623***	-0.0789***	-0.0263**	-0.0332***	-0.0584*	-0.0909**
	(0.00923)	(0.00896)	(0.0200)	(0.0291)	(0.0124)	(0.00989)	(0.0313)	(0.0407)
N	14482	14482	14482	14482	14482	14482	14482	14482
R^2	0.0865	0.0892	0.0776	0.0733	0.416	0.423	0.415	0.419
Estimation Method	OLS	OLS	IV	IV	OLS	OLS	IV	IV
F (Kleibergen-Paap)			20.79	12.61			20.79	12.61
Socio-econ. Controls	✓	✓	✓	✓	✓	✓	✓	✓
Land Controls		✓		✓		✓		✓
District FE	✓	✓	✓	✓	✓	✓	✓	✓

Note – Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 7: Years of Schooling

	Years of Schooling			
	(1)	(2)	(3)	(4)
Conflict	-0.129	-0.177	-0.641*	-0.961**
	(0.113)	(0.114)	(0.354)	(0.463)
N	14476	14476	14476	14476
r2	0.338	0.343	0.334	0.336
Estimation Method	OLS	OLS	IV	IV
F (Kleibergen-Paap)			20.79	12.61
Socio-econ. Controls	✓	✓	✓	✓
Land Controls		✓		✓
District FE	✓	✓	✓	✓

Note – Outcome variable: Years of schooling. Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 8: Primary Education

	Primary Education			
	(1)	(2)	(3)	(4)
Conflict	-0.0141	-0.0162*	-0.0697**	-0.0928**
	(0.00958)	(0.00944)	(0.0312)	(0.0413)
N	14476	14476	14476	14476
r2	0.287	0.292	0.282	0.284
Estimation Method	OLS	OLS	IV	IV
F (Kleibergen-Paap)			20.79	12.61
Socio-econ. Controls	✓	✓	✓	✓
Land Controls		✓		✓
District FE	✓	✓	✓	✓

Note – Outcome variable: Indicator for having finished primary school. Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 9: Education Further Results

	Years of Schooling			Primary Education		
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	-0.751**	0.0807		-0.0832**	0.0178	
	(0.382)	(0.431)		(0.0349)	(0.0368)	
Conflict × 61-01			-0.511***			-0.0438**
			(0.187)			(0.0170)
N	12967	1509	14476	12967	1509	14476
r2	0.325	0.309	0.368	0.271	0.279	0.313
Estimation Method	IV	IV	DID-IV	IV	IV	DID-IV
F (Kleibergen-Paap)	20.79	12.61		20.79	12.61	
Sample	Young	Old	All	Young	Old	All
Socio-econ. Controls	✓	✓	✓	✓	✓	✓
FE Level	District	District	Chiefdom	District	District	Chiefdom

Note – Outcome variables: Years of schooling (columns 1-3) and indicator for having finished primary school (columns 4-6). Samples: 'Young' are all individuals born before 1961, 'Old' are all individuals born after 1961. Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 10: Model parameters

Parameter	Description	Value	Source
θ	Fréchet dispersion	4.729	Estimation
$\phi^A - \phi^N$	Diff. in educ. returns	0.14	Estimation
σ	Elast. of subst.	8	Literature
A_d^A	Agric. firm product.	mean: 0.155	Estimation
A_d^N	Non-agric. firm product.	mean: 0.199	Estimation
a_d	Amenities	mean: 0.872	Estimation
q_{0c}	Avg. non-educ. origin prod.	mean: 0.844	Estimation
q_{1c}	Avg. educ. origin prod.	mean: 1.048	Estimation
τ_0^A	Migr. cost non-educ. in A	mean: 0.527	Estimation
τ_0^N	Migr. cost non-educ. in N	mean: 0.693	Estimation
τ_1^A	Migr. cost educ. in A	mean: 0.340	Estimation
τ_1^N	Migr. cost educ. in N	mean: 0.689	Estimation
N_0	Non-educ. origin pop.	mean: 1204	Observed
N_1	Educ. origin pop.	mean: 425	Observed

Note – Model parameters. Normalisations: $A^N_{Freetown} = a_{Freetown} = \phi^N q_{1,Freetown} = 1$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

Table 11: Effect of conflict on A_d^S and a_d

	(1)	(2)	(3)
	$\ln A_d^A$	$\ln A_d^N$	$\ln a_d$
Conflict	0.0116	-0.0765	-0.00635
	(0.134)	(0.174)	(0.0339)
N	150	138	151
r2	0.412	0.523	0.418
Estimation Method	IV	IV	IV
F (Kleibergen-Paap)	13.11	18.18	13.70
Socio-econ. Controls	✓	✓	✓
District FE	✓	✓	✓

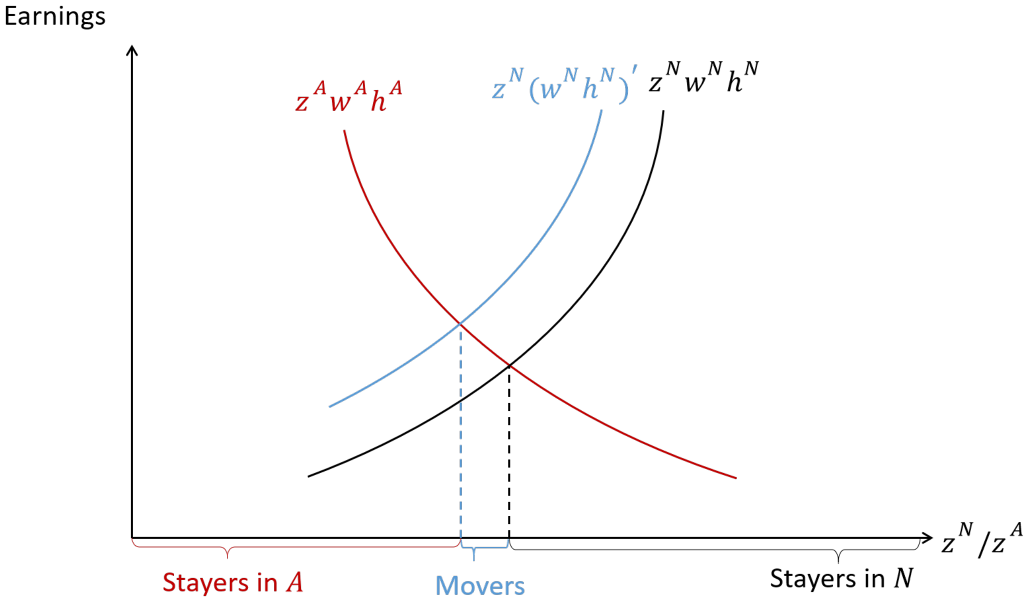
Note – Outcome variables: Log firm-level productivity in agriculture and non-agriculture (columns (1) and (2)) and log amenities (column (3)) at chiefdom level. Conflict standardised and instrumented with distance to border. Clustered standard errors at the chiefdom level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ [Back to main]

B. Figures

Figure 1: Map of Sierra Leone in West Africa

Note – Freetown is the capital of Sierra Leone. Source: Open Street Map. [Back to main]

Figure 2: Worker selection



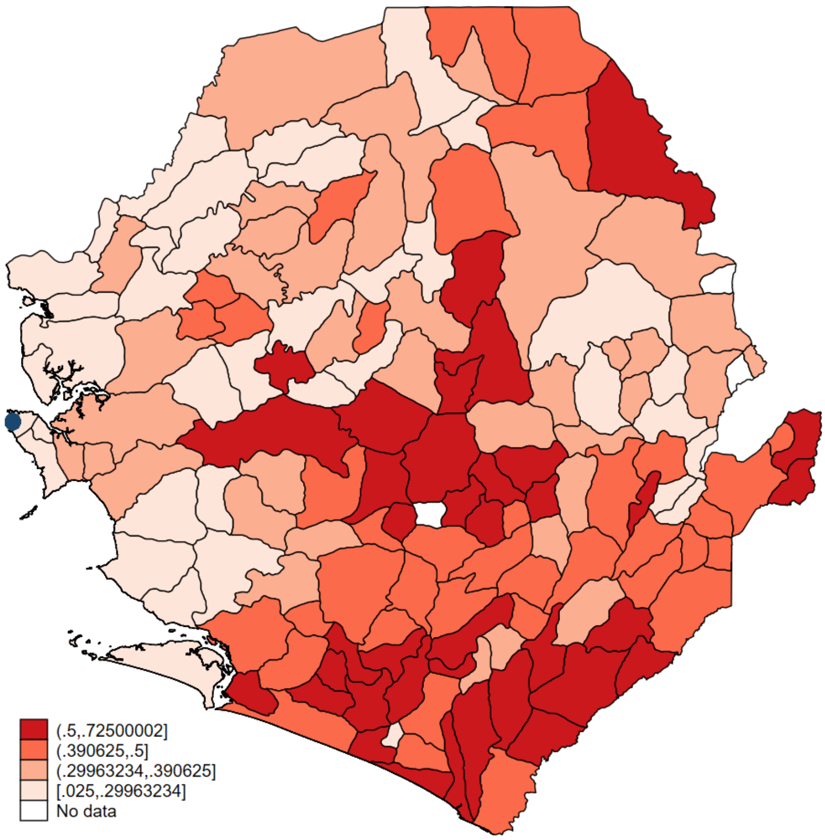
This graph shows the sector choice response within a location as wages change. A higher wage in non-agriculture implies more people choosing to work in this sector. As a result of comparative and absolute advantage being aligned (both curves upward sloping), the average productivity of the movers is lower than that of the stayers in N. Therefore, overall average productivity in that sector must go down. [Back to main]

Figure 3: ISIC Sector Classification

Section	Divisions	Description	
A	01–03	Agriculture, forestry and fishing	Agriculture
B	05–09	Mining and quarrying	Manufacturing
C	10–33	Manufacturing	
D	35	Electricity, gas, steam and air conditioning supply	
E	36–39	Water supply; sewerage, waste management and remediation activities	
F	41–43	Construction	
G	45–47	Wholesale and retail trade; repair of motor vehicles and motorcycles	Services
H	49–53	Transportation and storage	
I	55–56	Accommodation and food service activities	
J	58–63	Information and communication	
K	64–66	Financial and insurance activities	
L	68	Real estate activities	
M	69–75	Professional, scientific and technical activities	
N	77–82	Administrative and support service activities	
O	84	Public administration and defence; compulsory social security	
P	85	Education	
Q	86–88	Human health and social work activities	
R	90–93	Arts, entertainment and recreation	
S	94–96	Other service activities	
T	97–98	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use	
U	99	Activities of extraterritorial organizations and bodies	

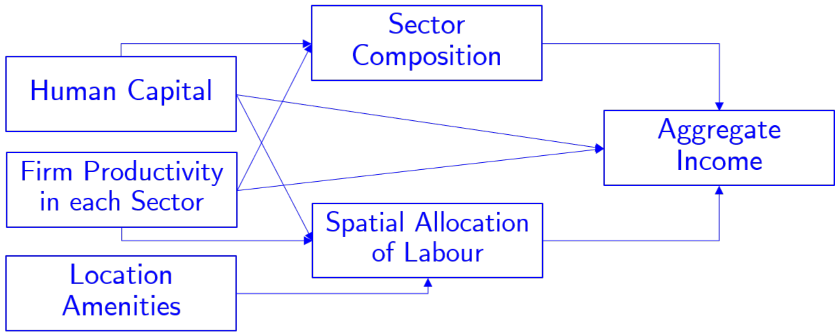
Note – From information on the ISIC section of their main professional activity, workers are classified to work in one of the three sectors according to the above figure. [Back to main]

Figure 4: Variation of the Victimisation Index



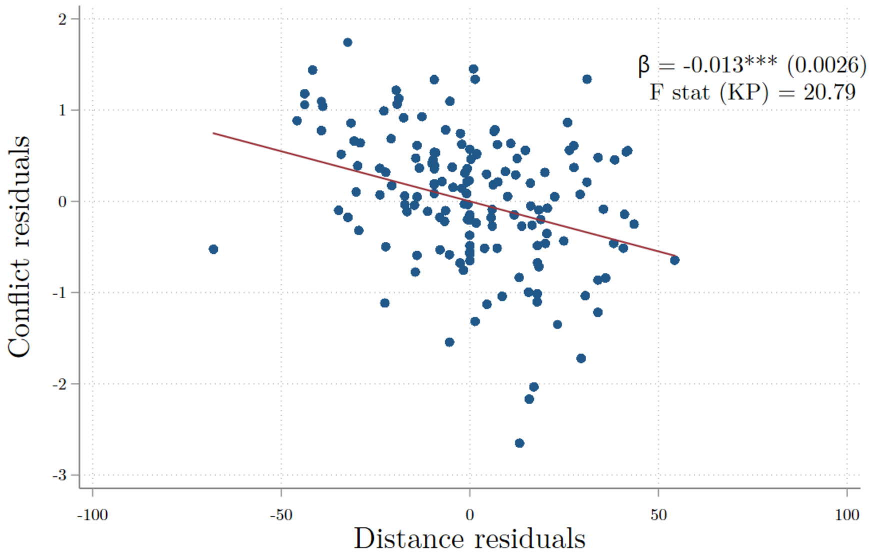
Note – The victimisation index is explained in section 3.1. The realisation is displayed for all chiefdoms of Sierra Leone. The south-eastern border on the map is the border with Liberia. The blue dot in the western part of the country marks the capital Freetown. Missing conflict information in three chiefdoms. [Back to main]

Figure 5: Key Mechanisms in the Model



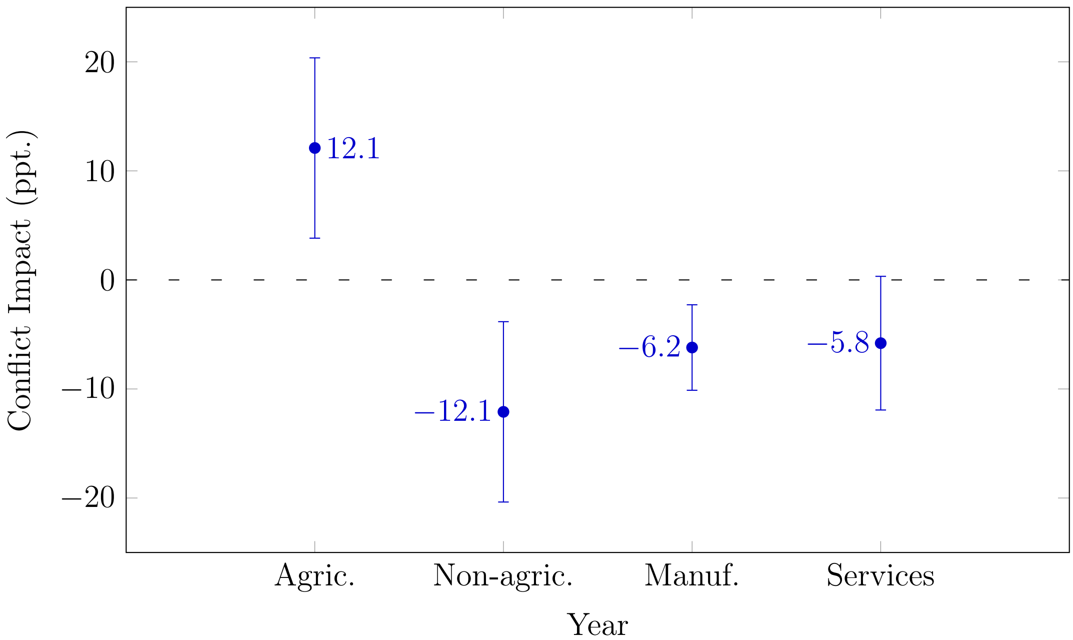
Note – This diagram depicts the key mechanisms how conflict affects income both directly and indirectly in the model. [Back to main]

Figure 6: First Stage Within Districts



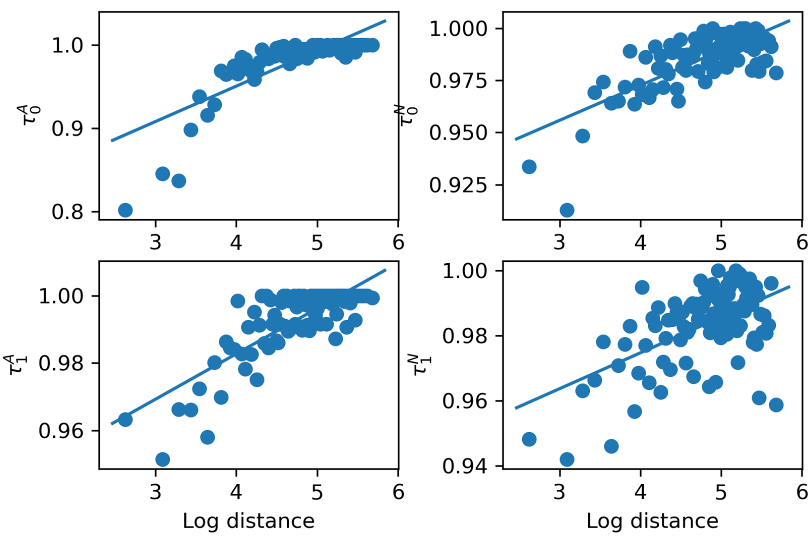
Note – The conflict measure is standardised and distance is measured in km. The red line presents an OLS line of best fit. [Back to main]

Figure 7: Sector Shift



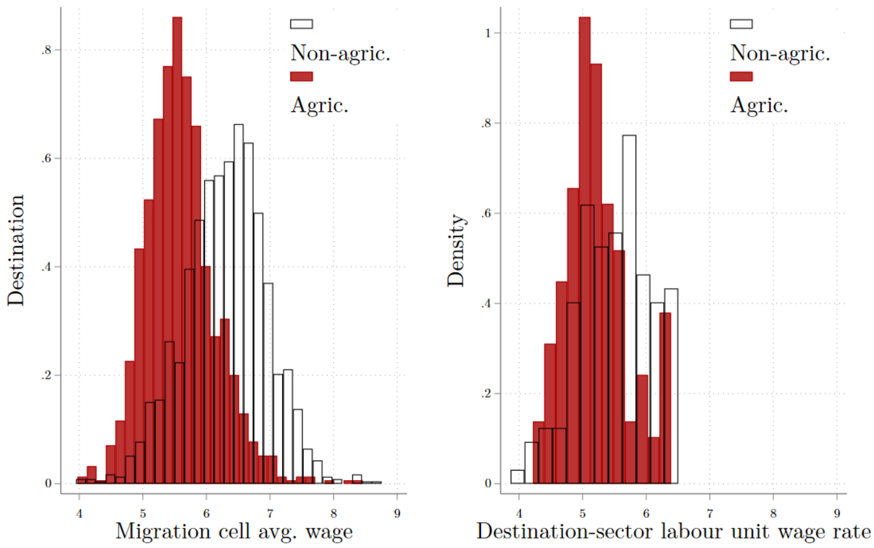
Note – IV coefficients depicted for specification with socio-economic controls, district fixed effects and minimum distance to large cities controlled for. 95% confidence intervals based on clustered standard errors at the chiefdom level. [Back to main]

Figure 8: Migration cost and distance



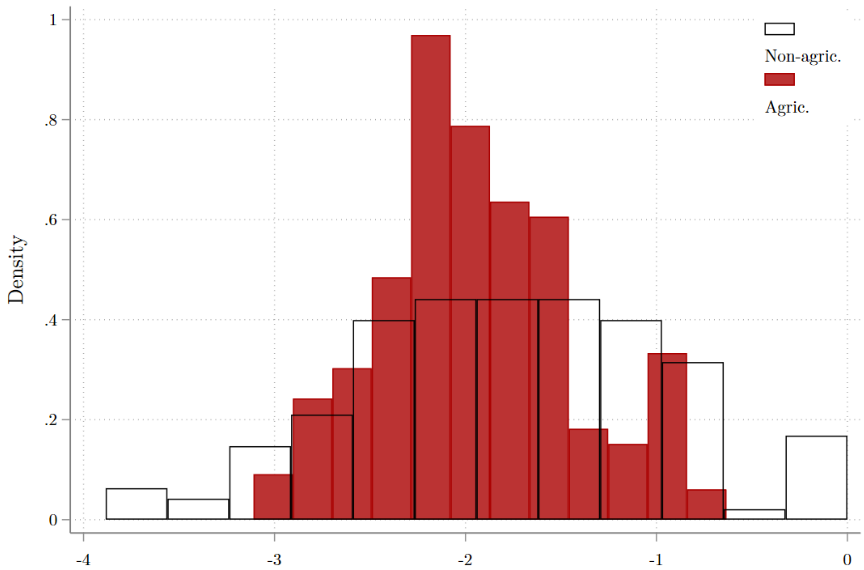
Note – Binscatter of the relationship between migration cost (by primary school education and sector) and log distance between locations. The line is the OLS line of best fit. [Back to main]

Figure 9: Distribution of labour unit wage rates



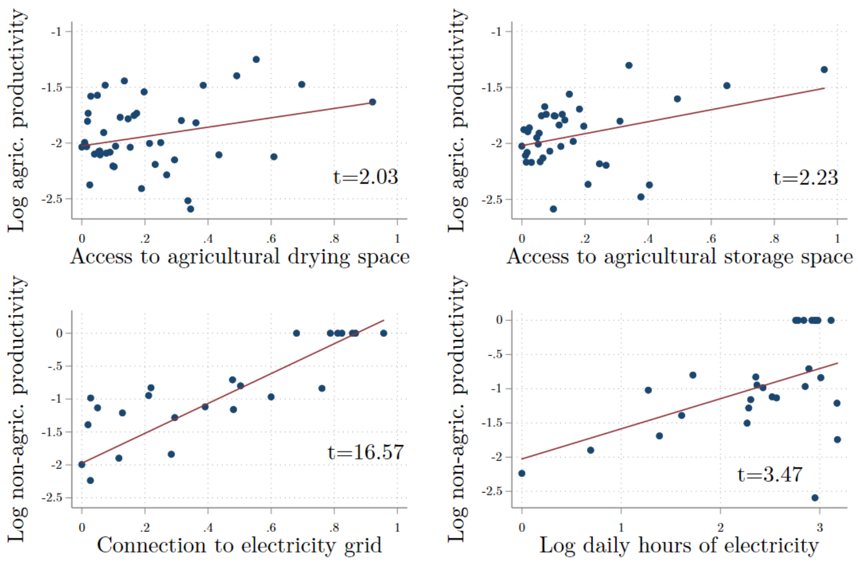
Note – Histograms of observed log income in the data $\ln M^S_{ecd}$ and estimated log labour unit wage rates $\ln w^S_d$ by sector. [Back to main]

Figure 10: Distribution of firm-level productivities



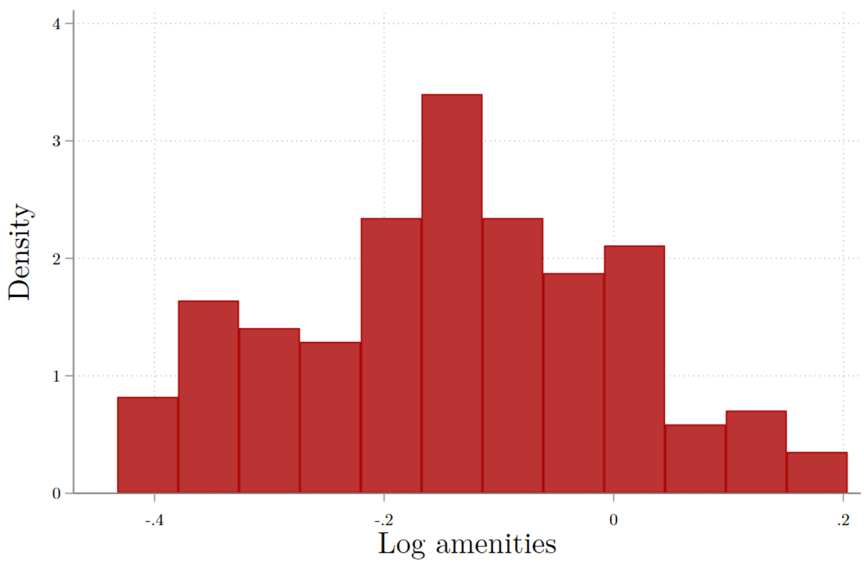
Note – Histograms of log firm-level productivities by sector. [Back to main]

Figure 11: Correlates with firm-level productivities



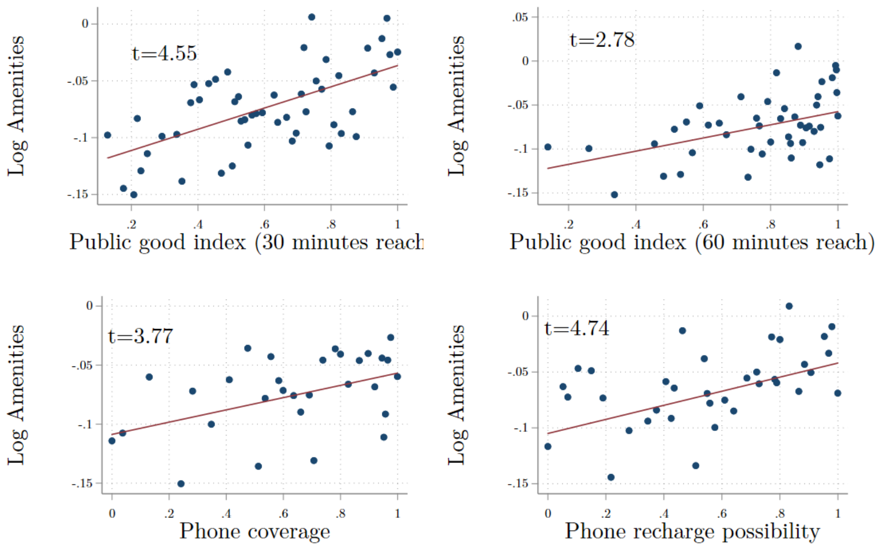
Note – Bin scatter plots with line of best fit of the relationship between log (non-)agricultural firm-level productivity $\ln A_d^A$ ($\ln A_d^N$) and various outcomes on the x-axis. [Back to main]

Figure 12: Distribution of amenities



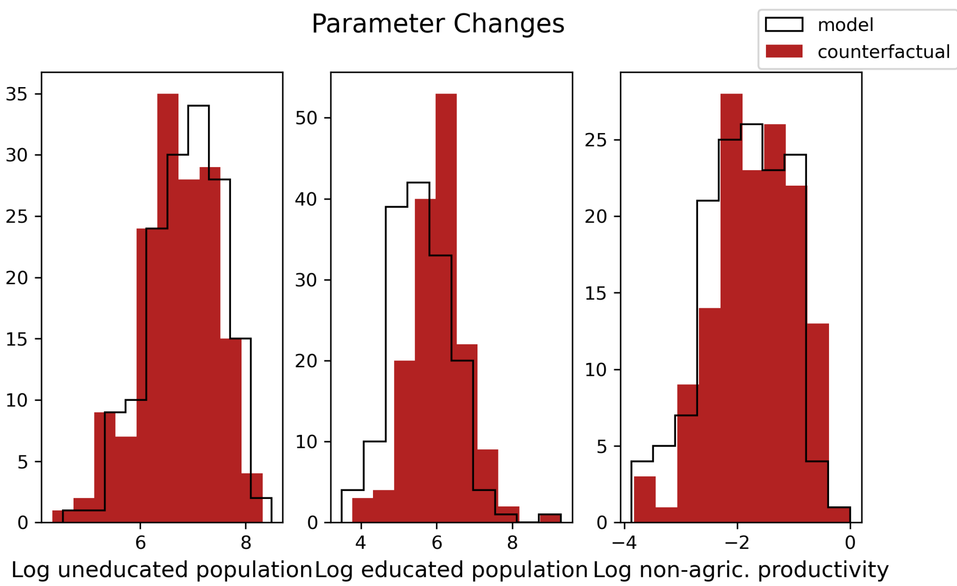
Note – Histograms of log amenities. [Back to main]

Figure 13: Correlates with amenities



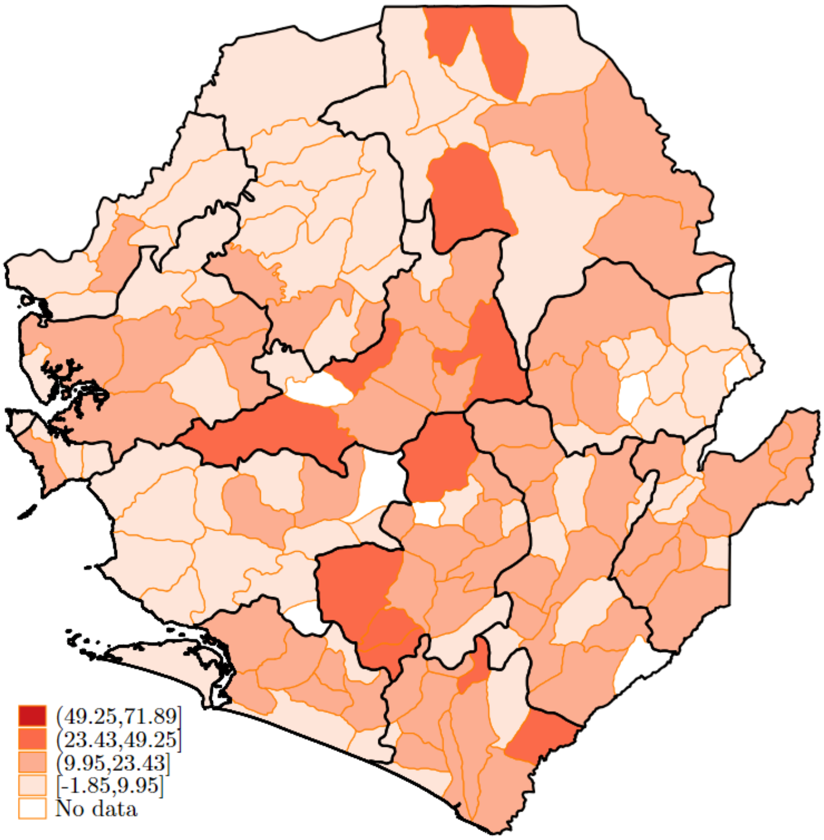
Note – Bin scatter plots with line of best fit of the relationship between log amenities $\ln a_d$ and various outcomes on the x-axis. The public good index is the average of eight binary variables indicating whether the following public goods/services are within 30 (60) minutes reach: (i) supply of drinking water, (ii) food market, (iii) public transportation, (iv) primary school, (v) secondary school, (vi) health clinic, (vii) hospital, (viii) all year motorable road [Back to main]

Figure 14: Parameter changes



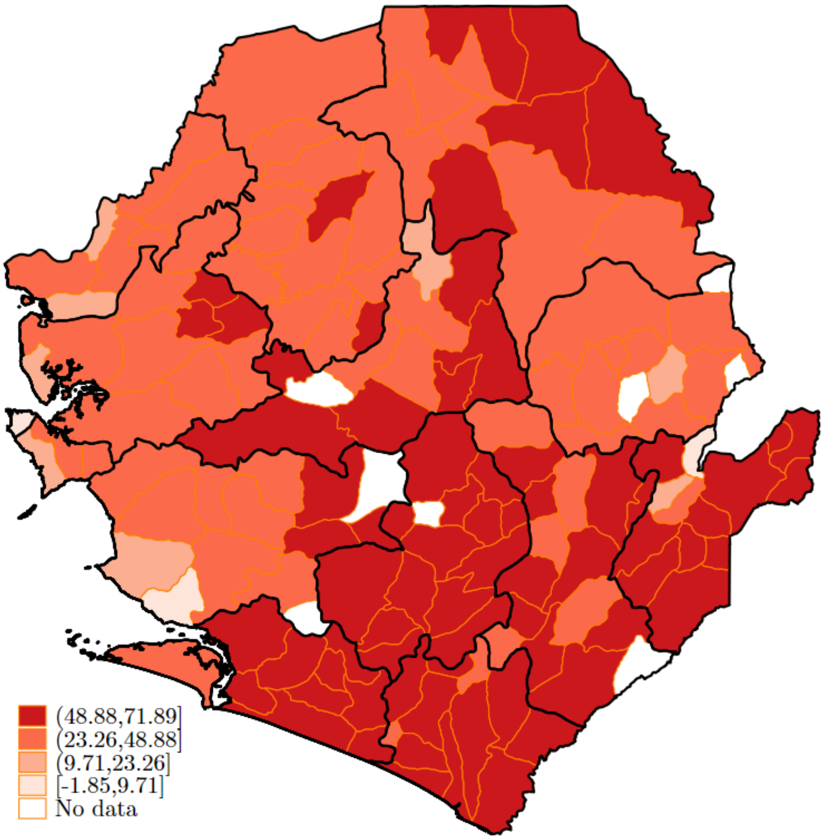
Note – Distributions of origin populations of individuals with and without primary school education as well as the parameter A_d^S (in logarithms). The black histogram displays the original distributions and the red histogram displays the counterfactual distributions. [Back to main]

Figure 15: Aggregate Income Changes



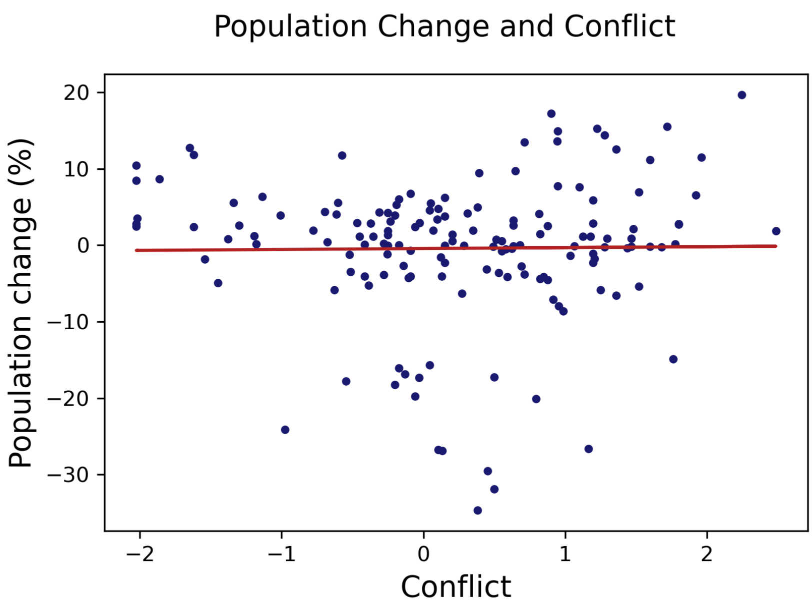
Note – Absolutve value of aggregate income changes (in %) in each chiefdom as a result of the civil war relative to a no-war scenario. Both human capital and non-agricultural firm productivity losses considered. Estimates stem from model simulation. Missing data: 5 chiefdoms unobserved in households survey 2018; 3 chiefdoms without conflict information. [Back to main]

Figure 16: Reduced-form Income Effect



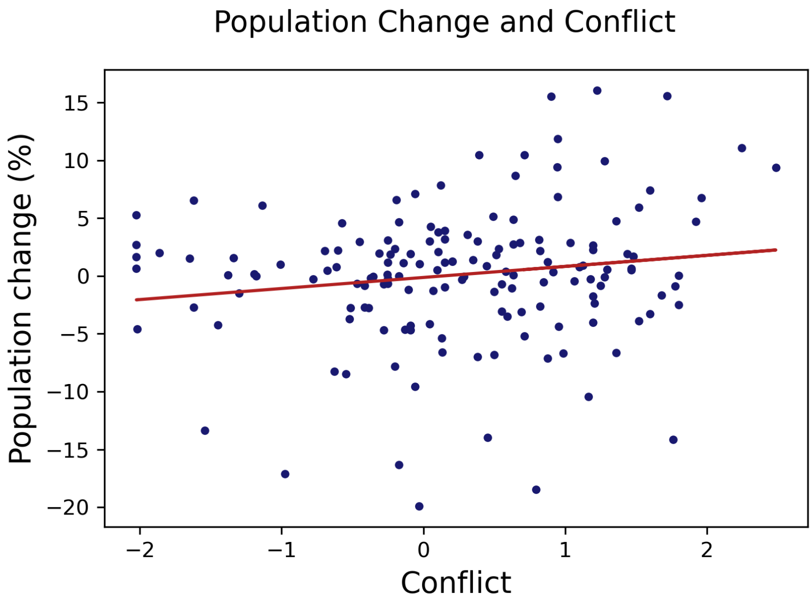
Note – Absolute value of the reduced-form income effect scaled by the conflict intensity in each chiefdom. Estimates are spatial income differences and stem from the IV estimation. Missing data: 5 chiefdoms unobserved in households survey 2018; 3 chiefdoms without conflict information. [Back to main]

Figure 17: Population Change and Conflict in First Scenario



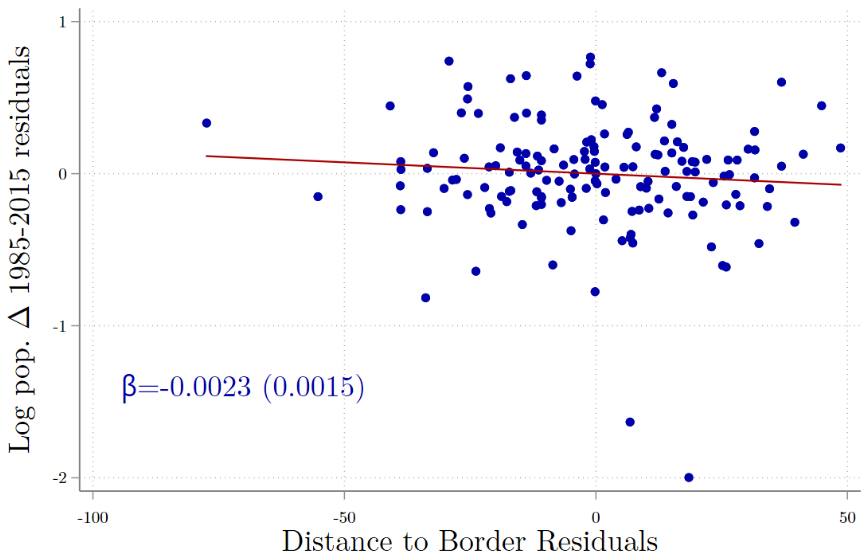
Note – Relationship between population changes in a full war scenario (both human capital and non-agricultural firm productivity loss) and conflict intensity in a chiefdom. Estimates stem from a model simulation. [Back to main]

Figure 18: Population Change and Conflict in Second Scenario



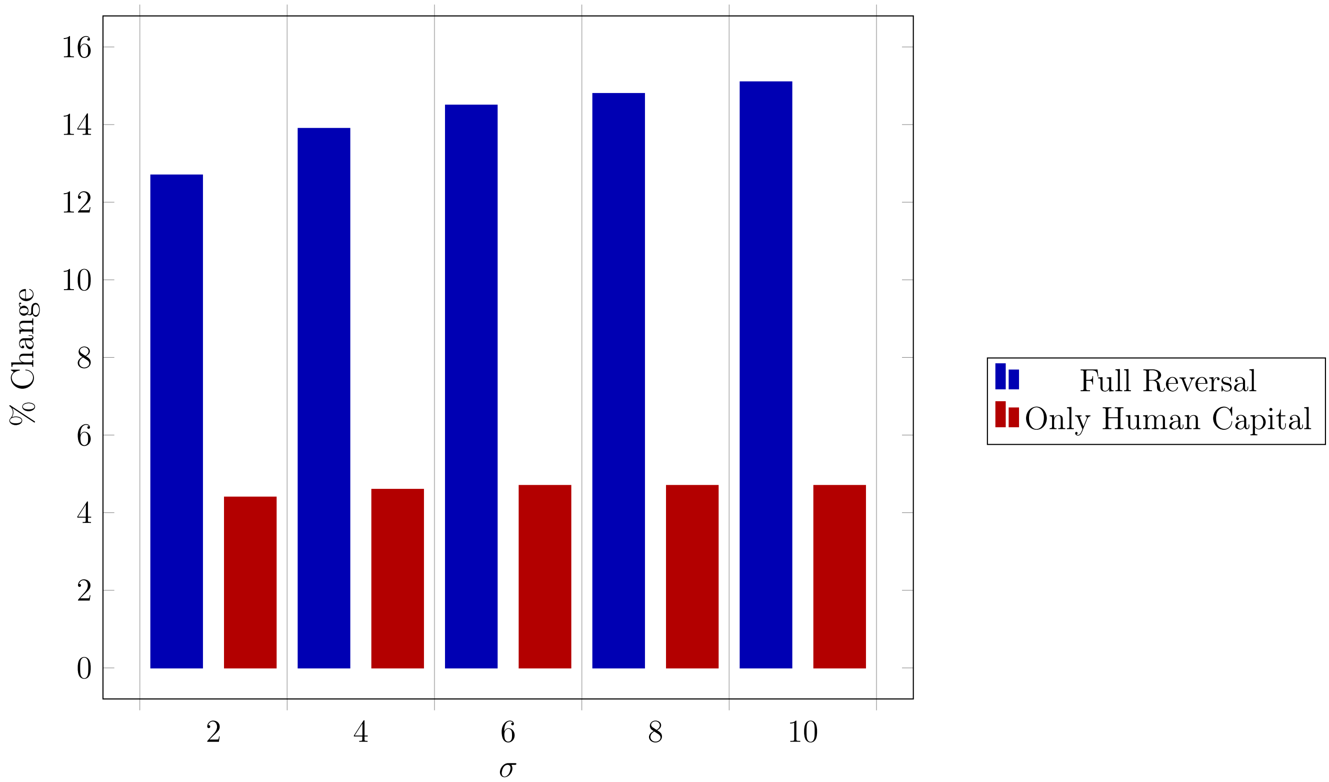
Note – Relationship between population changes in response to human capital loss only and conflict intensity in a chiefdom. Estimates stem from a model simulation. [Back to main]

Figure 19: Population Change and Distance to Border



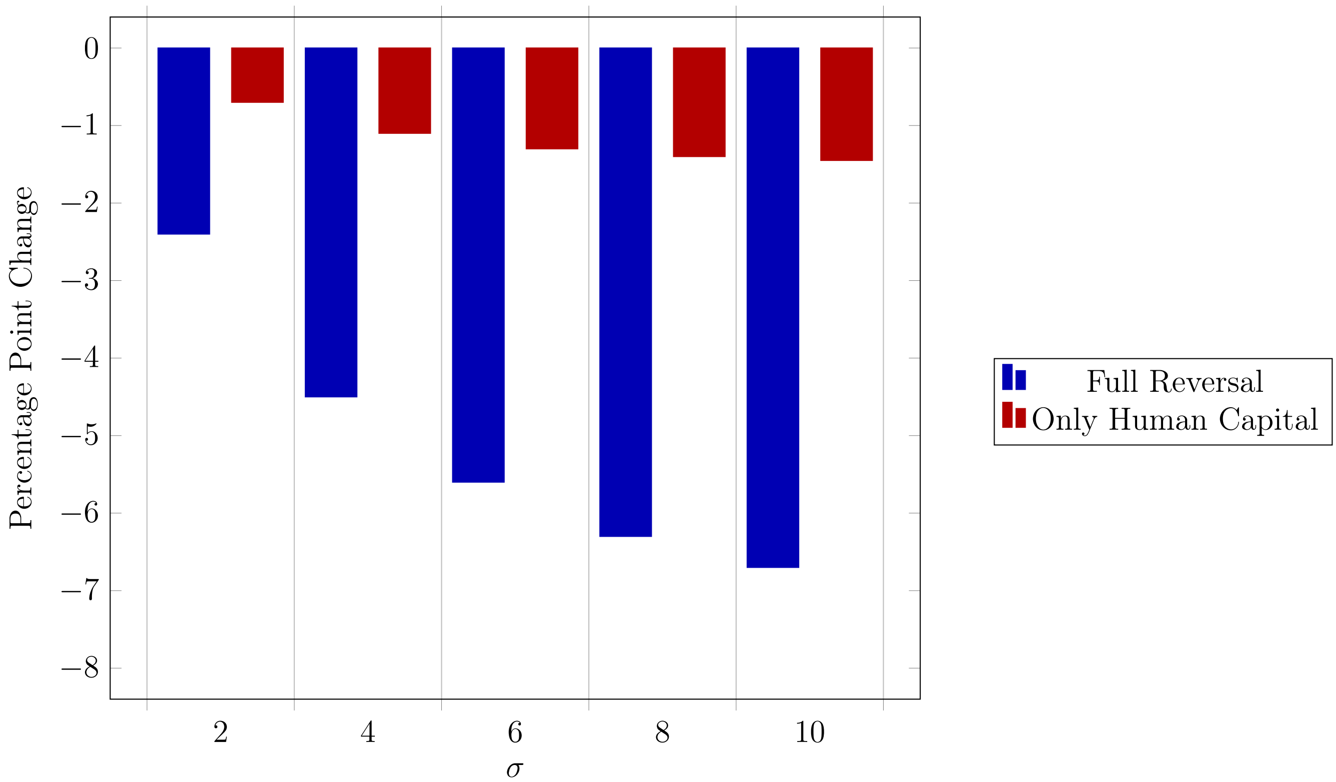
Note – Relationship between observed population changes between 1985 and 2015 and distance to border (as instrument for chiefdom) in a chiefdom. The stright line presents an OLS line of best fit. β is the regression coefficient. [Back to main]

Figure 20: Robustness to different σ values: Aggregate Income



Note – Aggregate income changes for different values of the elasticity of substitution σ. [Back to main]

Figure 21: Robustness to different σ values: Employment in Agriculture



Note – Changes in the economy-wide size of the agricultural sector for different values of the elasticity of substitution σ. [Back to main]

C. Proofs

Proof of equation 12. Recall that individual productivity draws z_{id}^S for destination d and sector S are independently Fréchet distributed:

$$F_{ec}(z_{id}^S) = \exp\left(-\left(\frac{z_{id}^S}{q_{ecd}^S}\right)^{-\theta}\right) \tag{29}$$

Consider the indirect utility distribution of workers *choosing* destination d and sector S, that is, the distribution of $\overline{V}_{id}^S := V_{id}^S | i \in M_{ecd}^S = \max_{d',S'}\{V_{id'}^{S'}\}$. Using the fact that $V_{id}^S = \frac{v_{ecd}^S z_{id}^S}{P}$, we have

$$\begin{aligned} F_{\overline{V}_{id}^S}(u) &= Pr\left[V_{id'}^{S'} < u \;\forall d', S'\right] \\ &= \exp\left(-\sum_{d'}\sum_{S'}\left(\frac{uP}{v_{ecd}^S q_{ecd}^S}\right)^{-\theta}\right) \\ &= \exp\left(-u^{-\theta}\sum_{d'}\sum_{S'}\left(\frac{v_{ecd}^S q_{ecd}^S}{P}\right)^{\theta}\right) \\ &= \exp\left(-\left(\frac{u}{\left(\sum_{d'}\sum_{S'}\left(\frac{v_{ecd}^S q_{ecd}^S}{P}\right)^{\theta}\right)^{\frac{1}{\theta}}}\right)^{-\theta}\right) \end{aligned} \tag{30}$$

Therefore, the distribution of $\overline{V}_{id}^S$ is Fréchet itself with scale parameter

$$t_{ec} := \left(\sum_{d'}\sum_{S'}\left(\frac{v_{ecd}^S q_{ecd}^S}{P}\right)^{\theta}\right)^{\frac{1}{\theta}} \tag{31}$$

The probability that someone chooses d and sector S is

$$\begin{aligned} \pi_{ecd}^S &= Pr\left[V_{id}^S > V_{id'}^{S'} \;\forall d', S' \neq d, S\right] \\ &= \int_0^\infty Pr\left[V_{id'}^{S'} < u \;\forall d', S' \neq d, S\right] dF_{V_{id}^S}(u) \\ &= \int_0^\infty \frac{F_{\overline{V}_{id}^S}(u)}{F_{V_{id}^S}(u)} \theta u^{-1-\theta}\left(\frac{v_{ecd}^S q_{ecd}^S}{P}\right)^{\theta} F_{V_{id}^S}(u) du \\ &= \left(\frac{v_{ecd}^S q_{ecd}^S}{P}\right)^{\theta} t_{ec}^{-\theta}\int_0^\infty \underbrace{\theta u^{-1-\theta} t_{ec}^{\theta} F_{\overline{V}_{id}^S}(u) du}_{f_{\overline{V}_{id}^S}(u)} \\ &= \frac{(v_{ecd}^S \epsilon_{ecd}^S)^\theta}{\sum_{d'}\sum_{S'}(v_{ecd'}^{S'}\epsilon_{ecd'}^{S'})^\theta} \end{aligned} \tag{32}$$

where the last step makes use of the decomposition of $q_{ecd}^S = \overline{q_{ec}}\epsilon_{ecd}^S$ into an origin-education average and remaining variation.

Proof of equation 13. Using the distribution of $\overline{V}_{id}^S$, we have that $z_i | i \in M_{ecd}^S = \frac{\overline{V}_{id}^S P}{v_{ecd}^S}$ also follows a Fréchet distribution with scale parameter

$$\frac{t_{ec}P}{v_{ecd}^{S}} = \left(\frac{\sum_{d'}\sum_{S'}\left(v_{ecd}^{S}q_{ecd}^{S}\right)^{\theta}}{\left(v_{ecd}^{S}\right)^{\theta}}\right)^{\frac{1}{\theta}} \tag{33}$$

It follows that

$$\begin{aligned} E\left[z_i | i \in M_{ecd}^{S}\right] &= \frac{t_{ec}P}{v_{ecd}^{S}}\Gamma\left(\frac{\theta-1}{\theta}\right) \\ &= \left(\frac{\sum_{d'}\sum_{S'}(v_{ecd'}^{S'}q_{ecd'}^{S'})^{\theta}}{\left(v_{ecd}^{S}\right)^{\theta}}\right)^{\frac{1}{\theta}}\Gamma\left(\frac{\theta-1}{\theta}\right) \\ &= (\pi_{ecd}^{S})^{-1/\theta}q_{ecd}^{S}\Gamma\left(\frac{\theta-1}{\theta}\right) \end{aligned} \tag{34}$$