WPS7273

Policy Research Working Paper 7273

Infrastructure in Conflict-Prone and Fragile Environments

Evidence from the Democratic Republic of Congo

Rubaba Ali
A. Federico Barra
Claudia N. Berg
Richard Damania
John D. Nash
Jason Russ

WORLD BANK GROUP
Agriculture Global Practice Group
May 2015

Policy Research Working Paper 7273

Abstract

In conflict-prone situations, access to markets is necessary to restore economic growth and generate the preconditions for peace and reconstruction. Hence, the rehabilitation of damaged transport infrastructure has emerged as an overarching investment priority among donors and governments. This paper brings together two distinct strands of literature on the effects of conflict on welfare and on the economic impact of transport infrastructure. The theoretical model explores how transport infrastructure affects conflict incidence and welfare when selection into rebel groups is endogenous. The implications of the model are tested with data from the Democratic Republic of Congo. The analysis addresses the problems of the endogeneity of transport costs and conflict using a novel set of instrumental variables. For transport costs, a new instrument is developed, the "natural-historical path," which measures the most efficient travel route to a market, taking into account topography, land cover, and historical caravan routes. Recognizing the imprecision in measuring the geographic impacts of conflict, the analysis develops a spatial kernel density function to proxy for the incidence of conflict. To account for its endogeneity, it is instrumented with ethnic fractionalization and distance to the eastern border. A variety of indicators of well-being are used: a wealth index, a poverty index, and local gross domestic product. The results suggest that, in most situations, reducing transport costs has the expected beneficial impacts on all the measures of welfare. However, when there is intense conflict, improvements in infrastructure may not have the anticipated benefits. The results suggest the need for more nuanced strategies that take into account varying circumstances and consider actions that jointly target governance with construction activities.

This paper is a product of the Agriculture Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at rdamania@worldbank.org.

Infrastructure in Conflict-Prone and Fragile Environments: Evidence from the Democratic Republic of Congo

Rubaba Ali[1]*
A. Federico Barra**
Claudia N. Berg***
Richard Damania**
John D. Nash**
Jason Russ***

JEL Codes: O1, I3, L9

Keywords: Transport, Conflict, DRC, nightlight data, kernel density, ethnic fractionalization, natural historical path

[1] Authors in alphabetical order *University of Maryland, ** World Bank,*** George Washington University.

Infrastructure in Conflict-Prone and Fragile Environments: Evidence from the Democratic Republic of Congo

1. Introduction

The rehabilitation of damaged transport infrastructure is often a high priority for governments and donors in post-conflict and conflict prone fragile states. The justification seems compelling - improved connectivity can rekindle economic activity, revive fragile economies, and spur economic growth that could stave off future conflict. For example, in the conflict-afflicted Democratic Republic of Congo (DRC), spending on transport infrastructure was approximately US$230 million per year during the mid-2000s, and increased to US$275 million per year in 2008 and 2009 (Pushak and Briceño-Garmendia, 2011). The investment in road transport infrastructure in DRC will continue to be high as the World Bank, European Commission, African Development Bank, Belgium, DFID, Japan, the Republic of Korea, and Canada together donated US$1.19 billion, for roads (African Development Bank 2013). This pattern is evident in other conflict-prone economies too. The New Partnership for Africa's Development (NEPAD) has proposals for 9 highways across the continent, at an estimated cost of US$ 4.2 billion (Review 2003), all of which pass through fragile states (as defined by the OECD), and in most cases, the portions of the roads that need the most rehabilitation lie within these countries. In Afghanistan, the U.S. Agency for International Development (USAID), provided over US$1.8 billion between 2002 and 2007 to reconstruct roads (GAO 2008), while the US Department of Defense has allocated about US$300 million from the Commander's Emergency Response Program (CERP) funds for roads. The emphasis on spending for transport infrastructure raises questions about whether these funds are efficiently allocated, and whether they are effective in fragile states. In this paper, we study this much neglected question using data from DRC acknowledging that transport and conflict are just two of many other determinants of wealth, growth and poverty.

The DRC provides an apt case study on the effects of infrastructure in the context of conflict. The history of DRC has been characterized by frequent conflict, international exploitation, and economic stagnation. Since Sir Henry Stanley and King Leopold II of Belgium first drew the Congo Free State's borders in 1877, international and regional powers have competed for the country's resources and wealth. Upon achieving independence in 1960, after undergoing one of the darkest colonial periods in history, the country was again pillaged, this time by domestic forces, under the reign of President Mobutu. Since the overthrow of Mobutu in 1996, DRC has been in a nearly constant state of conflict and civil war, mainly in the mineral-rich eastern part of the country. Given its violent history, its network of roads—where they exist—have fallen into a dreadful state of disrepair. Constant conflict, poor governance, and lack of infrastructure have left DRC one of the poorest countries in the world, with the average Congolese resident living on less than $US0.75 per day. And yet, geography and natural endowments give DRC the potential to become one of the richest countries in the region. It is estimated that DRC's unexploited mineral wealth stands at US$24 trillion (UNEP 2011). Additionally, with over 22.5 million hectares of uncultivated, unprotected, non-forested fertile land, it is suggested that DRC has the potential to become the breadbasket of Africa, and feed over 1 billion people (Deininger and Byerlee 2011).

Harnessing the growth potential of these endowments is not without challenges. With violence having declined since the end of the second Congolese civil war in 2003, DRC now seems poised to make investments to spur future economic growth. However, the state of DRC's road infrastructure is severely deficient even by the standards of other low-income countries (see table 1). It is striking that only four provincial capitals out of ten can be reached by road from the national capital, Kinshasa. Improvement in road infrastructure will undoubtedly need to include significant road improvement and construction projects. However, with conflict still erupting in parts of the country, this raises the question of whether improvements in transport infrastructure will bring benefits to local economies. There is also the possibility that such interventions may have perverse effects and provide violent militias easier access to vulnerable and remote communities, or tempt subsistence farmers to join the militias. This paper attempts to shed light on these questions, by studying the interaction between conflict within DRC and transport infrastructure.

Table 1.

Indicator	Units	Low-Income Country Average	DRC
Paved Road Density	km/1000 km^2 of land	16	1
Unpaved Road Density	km/1000 km^2 of land	68	14
Paved Road Traffic	Average daily traffic	1,028	257
Unpaved Road Traffic	Average daily traffic	55	20
Perceived Transport Quality	% firms identifying as major business constraint	23	30

Source: The Democratic Republic of Congo's Infrastructure: A Continental Perspective, March 2010

In order to motivate and contextualize our empirical analysis, we first present a simple theoretical model to analyze how transport costs and conflict are interlinked and how they directly and jointly affect welfare indicators. Together with a rigorous estimation strategy, our analysis adds to the literature by examining the effect of transportation costs and conflict on a variety of measures of well-being in DRC, captured by a wealth index, the probability of being multi-dimensionally poor (as defined by an index of living standard, education and health dimensions), and an estimate for local GDP which uses data from nighttime lights. We test for heterogeneous effects of transportation costs, conflict and their combined effect on these indicators.

Using DRC's Africa Infrastructure Country Diagnostic road data[2] and GIS road network data on both trunk and rural roads, we develop a new data set of travel costs. While our measure of transport costs is the most accurate possible given available data, it is still endogenous due to the non-random placement of rural infrastructure. That is, roads tend to be placed near developed or high economic potential areas. Similarly, conflict is endogenous as it is closely related to wealth: conflict negatively affects wealth, while by the same token low levels of income can trigger incidences of conflict. We address these sources of bias following an instrumental variables strategy.

We instrument for transport cost by constructing a novel instrumental variable, which we term the natural-historical path, which measures the walking time taken to reach markets using the natural path (i.e. the shortest route given local geography and land cover), as well as historical caravan routes which were used to transport slaves and ivory. To instrument for conflict, we use a measure of the level of local social fractionalization, as theory predicts that polarization generates higher levels of conflict (Esteban and Ray 1999). While we believe that our instrumentation methodology is sound, and represents a significant improvement over the current literature, we also conduct robustness checks under the assumption that our instruments do not perfectly satisfy the exclusion restriction assumptions. Using Conley Bounds, we demonstrate that our results all remain consistent when the exclusion restriction assumption is relaxed for each of our instruments.

Overall, we find higher transportation costs have a significantly negative impact on wealth and a significantly positive impact on the probability of being multi-dimensionally poor. We also find that the location of conflict is important in determining its effect. Conflict near households has a strongly negative impact on a household's wealth, and conflict near markets has a large, positive impact on the probability of being multi-dimensionally poor. More significantly, we find that when there is high conflict near both the household and market, households farther away from the market are likely to have higher welfare indicators. In such cases improved transport infrastructure does little to improve, and may indeed worsen, well-being as distance and remoteness can presumably provide at least partial sanctuary from conflict.

[2] http://www.infrastructureafrica.org/library/doc/597/democratic-republic-congo-roads

The rest of the paper is organized as follows. Section 2 provides a brief overview of the related literature. Section 3 presents our theoretical framework. Section 4 describes our data sources and key variables constructed for the analysis. Section 5 discusses our empirical framework and identification strategy. Section 6 discusses the main findings. Section 7 concludes.

2. Literature Review

There are two unconnected strands of literature relevant to this research: a growing and established literature on the economic legacies of war, and a vast and rapidly evolving literature on the effects of infrastructure on well-being. The former is predominantly split into a macro focused literature -- on the nexus between conflict and economic growth -- and a microeconomic strand focusing on human capital (see Blattman and Miguel (2010) for a succinct discussion of this literature). The literature on infrastructure focuses on indicators of welfare, such as aggregate productivity (usually measured by gross domestic product or per capita income calculated using gross national income), output elasticity and productivity, as well as household income from agricultural and non-agricultural sources.

The literature on conflict includes contributions by Knight *et al.* (1996), Bannon and Collier (2003), Collier (1999), Cerra and Saxena (2008), Justino and Verwimp (2006) and Hoeffler, and Reynal-Querol (2003), all of whom find evidence of a negative effect of conflict on GDP, output, and by implication adverse effects on poverty. Brück (2004a), Deininger (2003), and McKay and Loveridge (2005) find that households tend to lapse into subsistence farming in times of violence. Deininger (2003) and Brück (2004a, 2004b) report that violence reduces the potential for investment in the non-farm economy.

Complementing this work is a large literature that traces the consequences of violent conflicts on human capital and its determinants. Alderman, Hoddinott and Kinsey (2006) find that young children who suffered from war related malnutrition in Zimbabwe are significantly shorter as adults, which may affect their lifetime labor productivity. In a related paper, Bundervoet, Verwimp, and Akresh (2009) conclude that children who lived in a war-affected region have lower height-for-age ratios. Not surprisingly, Hoeffler and Reynal-Querol (2003) find positive effects of conflict on infant mortality and Bundervoet and Verwimp (2005) find significant negative impacts on the nutritional status, measured as height-for-age. Deininger (2003) and Olga Shemyakina (2011) both find a negative effect of civil wars on educational attainment, and Alderman *et al.* (2006) find a negative impact on height and schooling. These studies suggest that human capital effects may be a powerful mechanism whereby violent conflicts may force individuals and households into long-lasting poverty.

Yet claims of a direct causal line from conflict to poverty should be treated with caution, as causality may be reversed or indirect. It is noteworthy that the outbreak of civil wars is commonly attributed to poverty and the correlation between low per capita incomes and higher propensities for internal war is one of the most robust empirical relationships in the literature (Justino, 2006, Bannon and Collier 2003). A number of recent papers employ within-country data to explore the factors that predict violence and rebellion, and most find strong associations with local economic conditions and inequality (Barron et al 2004; Krueger and Maleckova 2003; Murshed and Gates 2005, Do and Iyer 2007, Macours 2011, Dube and Vargas 2013, Verwimp 2003, 2005, Humphreys and Weinstein 2008; Maystadt et. al 2013).

This paper also contributes to the literature on the effects of transportation infrastructure on welfare. Recent papers provide suggestive evidence on how better transport infrastructure, by enabling greater access to markets, decrease trade costs and interregional price gaps (Donaldson 2013; Casaburi et al, 2013), and affect input and output prices of crops (Khandker et al., 2006, Minten and Kyle 1999). These in turn affect agricultural returns and hence land values (Jacoby 2000, Shrestha 2012, Donaldson 2013). Econometric analysis of household data on the effects of road connectivity on input use, crop output, and household incomes in Madagascar and Ethiopia (Chamberlin et al 2007, Stifel and Minten 2008) suggest that remoteness negatively affects agricultural productivity and incomes at the household level. Not surprisingly the literature also finds that access to good quality roads facilitates economic diversification (Gachassin et al 2010, Fan et. al 2000, and Mu and van de Walle 2007). Several other researchers use microeconomic data to examine transportation infrastructure's impact on welfare variables, such as on income (Donaldson 2013, Jacoby and Minten 2009), consumption per capita (Khandker et al. 2006), and poverty reduction (Fan et al. 2000, Gibson and Rozelle 2003; Warr 2008).

As noted earlier, there is remarkably little empirical evidence on the direct causal impact of access to markets on well-being in countries with conflict, and even less evidence on the combined impact of transportation costs and conflict. For instance, Ulimwengu et al. (2009) analyze the effect of market access on agricultural production and household wealth in DRC, but do not explicitly explore the role of conflict beyond the inclusion of province level fixed effects, which does not capture variations in conflict within the large provinces of DRC. It is imperative to assess the direct impacts of conflict and transport infrastructure and the combined effect of these two in order to evaluate the effectiveness of providing transport infrastructure in areas with conflict. In contrast Martin et al (2008) investigate the links between trade and conflict. They conclude that trade openness may deter the most severe civil wars (those that destroy the largest amount of trade) but may increase the risk of lower-scale conflicts. The implication is that transportation costs can affect the level of conflict indirectly through its impact on the amount of trade that can occur.

3. Theoretical Framework

To guide the empirical analysis we outline a simple theoretical model which describes the manner in which variations in access to markets might influence incentives for rebellion in conflict affected areas. We outline a model where mobilization decisions are endogenous and focus on the effects of transportation costs on welfare, production and conflict incentives of individuals. While there is a significant theoretical literature on the economic determinants of conflict, to our knowledge none of the models explore the role of transport infrastructure on rebel incentives.[3] These issues are arguably of importance for understanding how policies influence individual decisions to engage in productive activities or to join rebel forces, and to guide the design of development strategies.

The model analyzes a situation where households at spatially distinct locations have a choice between either joining a rebel group who loot from other households, or else they engage in some productive activity, such as farming. The role of the government is left in the background, with levels of enforcement taken as given. While not every incidence of civil conflict is of this form, the DRC and many others fit into this broad category where government control is limited and violence often takes the form of looting rather than an outright struggle against the authorities whose power and presence is often circumscribed (see Fearon 2007 for a discussion).

In this framework there are three mechanisms that are key to understanding when insurgency is rendered more (less) profitable than some other economic activity. The first is the usual opportunity cost of conflict: when incomes are higher (say due to lower transport costs), the foregone income from joining the insurgents will be greater, so there is less incentive to participate in rebellion. The second mechanism concerns the size of the prize that is available to the rebels. In most theoretical models the prize is exogenously given and is typically defined by the availability of natural resources, or exogenous factors such as rainfall (Miguel 2004). In the current context the amount that can be looted is endogenously determined by the productive activities of households. Since rebels loot from farm households, higher levels of farmer incomes increase the lootable prize and make conflict more attractive. Third, as payoffs to productive economic activities such as farming decline, there is a greater incentive to join the rebels. All else equal, this implies that a smaller lootable prize will need to be shared among a larger number of looters. Our results therefore suggest that there is no simple linear relationship between policies that promote access to markets and development outcomes. We identify circumstances where such strategies can both inflame and moderate conflict. To our knowledge these issues appear not to have been explored and formally modeled or empirically tested in the literature in this context.

We consider the simplest possible economic structure and functional forms in order to generate empirically testable solutions. The economy consists of a continuum of individuals who are uniformly

[3] See for example Mesquita 2013, Laitin and Shapiro (2008) and the references therein.

distributed along the unit interval $n_i \in [0,1]$, where the individual at location $n_i = 0$ is closest to the market and that at $n_i = 1$ is farthest away from the market. Transport costs to the market are given by $t_i = zn_i$, z>0. Individuals at each location can either choose to farm or join a rebel force who loot from those who farm.

Decisions are made sequentially. In the first stage each individual decides whether to join the rebels or farm. Given the set of rebels (or equivalently the set of farmers at known locations), in the second stage the rebels determine effort levels and a looting strategy which consists of the decision on whether to attack at the market, or where the farmers are located. In the third stage farmers determine the type of goods to produce and the production levels of each. There are two types of produced goods - those produced for sale in the markets which must incur transport costs, and goods for domestic consumption (subsistence farming). Both products are potentially vulnerable to theft by rebels. This structure implies that the rebels have a strategic early-mover advantage relative to farmers. The model is solved by backward induction, beginning with the final stage of the game.

Stage 3- The farmers' decisions. In the third stage farmers determine production decisions. Each farmer is endowed with L units of an input that can be used to produce either a marketed good denoted L_m or a domestically consumed subsistence good denoted L_c, where L is the fixed endowment with $L = L_m + L_c$. The production functions for the two products are given by $M = mL_m$ and $C = cL_c$, (where m > 0 and c > 0). Goods produced for the market are sold to purchase a composite commodity denoted B, which is consumed. Each farmer determines production levels to maximize utility:

$$U_F = B^{\beta}(C(1-f))^{\alpha}; \qquad (2.1)$$

where, C is the self-consumed product produced by the farmer and f is the fraction of this product that is seized by rebels at the location of the farmer and α> 0, β> 0 with $\alpha+\beta=1$.[4] Thus (2.1) defines the utility to the farmer net of theft. Utility is maximized subject to the budget constraint:

$$(t_i + P_B)B = (1-g)M(P - t_i)\}; \qquad (2.2)$$

where t_i = zn_i is the cost of transport, P_B= price of purchased good B, P the price of the marketed (sold) good, and g is the proportion of the marketed good stolen by the rebels at the market.[5] Equation (2.2) simply asserts that that the amount of money spent on purchased goods including transport costs $(t_i + P_B)$, must equal the amount received from selling goods at the market net of transport costs $(P - t_i)$. A fraction g of these marketed goods are stolen by rebels.

[4] Note that the results go through when $U_F = B^{\beta}(C)^{\alpha}(1-f)$ implying theft of both goods.
[5] When not required the location subscript (i) is ignored for notational brevity.

Maximizing (2.1) and (2.2) with respect to B, L_m and L_c generates the reaction functions:

$$B = \frac{\beta(1-g)(P-t_i)\beta mL}{(P_B+t_i)}, \tag{2.3}$$

$$L_m = \beta L, \tag{2.4}$$

$$L_c = \alpha L. \tag{2.5}$$

Substituting in (2.1) yields the indirect utility function for this stage:

$$U_F^*(t_i) = \left[\frac{\beta(1-g)(P-t_i)\beta mL}{(P_B+t_i)}\right]^\beta [c\alpha L(1-f)]^\alpha \tag{2.6}$$

Note for future reference that $\frac{dU_F(t_i)}{dt_i} < 0$. Unsurprisingly in this simple set up, higher transport costs unambiguously lower welfare.

Stage 2 - The rebels' problem. Turning next to the rebels' problem, in stage 2, rebels determine their strategy which consists of deciding whether to loot from the market or at the farm. For simplicity we assume that all looted goods are valued equally by rebels, implying that the results are not influenced by arbitrary assumptions about relative prices of the goods. In this stage of the game the set of farmers (rebels) is taken as given. Let *n** be the given set of farmers (determined in stage 1), then aggregate payoffs to the rebel group is given by:

$$U_R = \int_0^{n*}\big(gmL_m + fc(L-L_m)\big)dx - K_g g^2 - K_f f^2, \tag{2.7}$$

where K_g, K_f are costs of looting at the market and farm, respectively.[6] It is assumed that these costs include risks and consequences of resisting the given levels of government defense. We demonstrate in the Appendix that the existence of an interior equilibrium with both farmers and rebels is contingent upon these costs being at intermediate levels. Excessively high looting costs render rebellion unattractive and vice-versa. The rebels maximize (2.7) taking as given the stage 3 decisions of farmers. Thus substituting from (2.4) for L_m and maximizing with respect to *g* and *f* yields the rebels' aggregate distribution of looting between farms and households:

$$\hat{g} = \frac{n*m\beta L}{2K_g} \tag{2.8}$$

[6] Though desirable for theoretical completeness, since the focus of this paper is on the empirical analysis to save space we do not show results of the perverse case where farmers locate over the interval [n*,1]. Results are available upon request.

$$\hat{f} = \frac{n*c(1-\beta)L}{2K_f} \quad (2.9)$$

Substituting from (2.8) and (2.9) yields the rebel group's aggregate indirect payoff function:

$$U_R^* = \frac{n*^2}{4}\left(\left\{\frac{[cL(1-\beta)]^2}{K_f}\right\} + \left\{\frac{[mL\beta)]^2}{K_g}\right\}\right) \quad (2.10)$$

We assume that these benefits are distributed equally between rebels, so that payoffs to the individual rebel *j* is simply:

$$U_{Rj}^* = \frac{U_R^*}{(1-n*)} \quad (2.11)$$

Stage 1- The decision to farm or rebel. In the first stage of the game each agent decides whether to farm, or join the rebels, given knowledge of the downstream responses. To simplify the analysis we abstract from the problems associated with imperfect or costly monitoring and shirking that might occur in the rebel group. We take the simplest possible case and assume that the marginal agent switches from farming to join the rebels if the payoffs from farming are less than those from joining the rebels. Hence the marginal farmer, located at $\widehat{n_\iota}$, is indifferent between farming and joining the rebels if:

$$U_F(\widehat{n_\iota}) = U_R(\widehat{n_\iota} \quad) \quad (2.11)$$

The value $\widehat{n_\iota}$ is the solution to:

$$\Psi \equiv U_F(\widehat{n_\iota}) - U_R(\widehat{n_\iota} \quad) = 0 \quad (2.12)$$

Where using (2.8) and (2.9) in (2.11) yields:

$$\Psi \equiv \left[\frac{\beta(1-g)(P-t_i)\beta mL}{(P_B+t_i)}\right]^\beta [c\alpha L(1-f)]^\alpha - \frac{n^2}{4(1-n)}\left(\left\{\frac{[cL(1-\beta)]^2}{K_f}\right\} + \left\{\frac{[mL\beta)]^2}{K_g}\right\}\right)$$

It is useful to explore how changes in transportation costs influence the decision to join rebel forces. In general higher transport costs are likely to have ambiguous effects. Intuitively higher transport costs lower output levels and the utility from farming which makes rebellion more attractive, ceteris paribus. However, lower output levels also reduce the amount that is available for looting, so looting benefits decline too. In addition if some households switch from farming to looting, a smaller output will need to be shared between a larger number of looters. The overall decision to farm (or participate in rebellion) will then depend on the relative rates of decline in the payoffs from farming and looting. If farming utility falls more rapidly than that from rebellion there will be a switch by the marginal agent from farming to the rebel forces and vice versa. Result 1 summarizes the cases when higher (lower) transport costs induce less (more) conflict.

Result 1- Lower transportation costs will induce a switch from farming to rebels (and vice versa) when the costs of stealing marketed goods is sufficiently low (high). (i.e. $\lim_{K_g \to 0}, \; then \; \frac{d\Psi}{dt} \to -\infty$)

Proof: See Appendix.

Intuitively, lower transportation costs make production for the market more attractive. Recall that K_g defines the cost of looting. As the costs of looting from the market decline, the strategic advantage (i.e. payoffs) accruing to the rebels increases as they are able to loot a greater share of the aggregate output produced by farmers. As a result, payoffs available to the rebels from theft of the higher aggregate output, can rise faster than the increased output of the marginal farmer. A corollary of result 1 is that when more households join the rebels agricultural output would decline in such situations, ceteris paribus, so there can be no presumption that transportation cost reductions induce the desired outcomes.

The next result explores whether an increase in conflict at the market is more damaging than at the household.[7]

Result 2. In general the welfare cost of a marginal increase in looting of subsistence goods relative to marketed goods is ambiguous, and is increasing in the relative welfare weights of these goods in the utility function. (i.e. $\left|\frac{dU_i^F}{df}\right| > \left|\frac{dU_i^F}{dg}\right| \; if \frac{\alpha}{\beta} > \frac{1-f}{1-g}$)

Proof: See Appendix.

Result 2 suggests that the consequences of increased looting at each location depend critically upon the welfare weights in the utility function. When the marketed good is given sufficient weight relative to the subsistence product the marginal welfare costs of theft at the market are higher and vice-versa. An implication of this result is that the combined effects of simultaneous changes in transport costs and conflict area also ambiguous even in this highly stylized framework.

In sum the model shows that transportation costs affect the consumption/utility levels of households through numerous channels, one by directly affecting the revenue of the households and others indirectly, by affecting the amount available to loot and the number of looters. The model also suggests that lower transportation costs could induce more conflict, and ultimately lead to a reduction in welfare in conflict-prone areas (Result 1). We confront these empirically testable hypotheses with data from DRC. In addition, we empirically analyze the effects of conflict at markets and households (Result 2) as well as the combined effect of conflict and transport cost, as the theory is ambiguous in predicting the consequences.

---

[7] As is conventional rebel utility is ignored.

4. Data

This paper uses several novel data sets to analyze the effect of transportation costs to the nearest market (defined as cities with population of 50,000), and conflict, on several different measures of welfare.[8] In order to do so, a thorough road network data set was constructed for DRC, using several sources of data described in section 4.1 below. To correct for any placement bias inherent in estimating the benefits of transportation infrastructure expenditures, an IV approach is used, and an innovative instrument was generated which we refer to as the natural-historical path, described in section 4.2. Section 4.3 describes the methodology used to calculate the measure of conflict around the households and markets. To account for endogeneity of conflicts, this paper constructs social fractionalization indices to use as instruments, which are described in Section 4.4. Finally, a few different welfare indicators are utilized in this paper, and they are described in section 4.5. Summary statistics for all of the data described in this section are given in Table 2.

4.1 Minimum travel costs to the market

Decades of deterioration and warfare have left DRC's transportation network in notoriously bad shape. Currently, only four provincial capitals out of ten can be reached by road from the national capital, Kinshasa. In many places, the jungle has begun to reclaim roads, paved and unpaved, alike. This makes collecting data on the road network a challenging undertaking.

The GIS vector data of the transportation network was obtained from Delorme.[9] Delorme's data was selected because of its thoroughness; it contains both major trunk roads, as well as rural roads throughout the country. This network was overlaid with another road network constructed for the African Infrastructure Country Diagnostic (AICD) that includes quality attributes such as whether the road is paved/unpaved, the road type (primary, 7m wide roads; secondary, 6m wide roads; and tertiary, 5m wide roads), and road quality (good; fair; poor) [10]. Through a process known as conflation, all these attributes necessary for determining the cost of traveling were transferred to Delorme's vector data. Finally, the road network was updated by making adjustments based on information obtained from transport experts familiar with DRC.

[8]This paper analyzes the combined effect of both large transport infrastructure (such as highways) and rural roads. Thus, it differs from Michaels (2008), Donaldson (2013), Datta (2012), Faber (2014) and Banerjee, Duflo and Qian (2012), that analyze the impact of large transport infrastructures, highways and railways. It also differs from Jacoby and Minten (2009), Dorosh et al (2010), Gibson and Rozelle (2003), Ali (2011), Khandker et al (2011), Mu and van de Walle (2007) that analyze the impact of smaller rural roads.

[9] Delorme is a private company which specializes in GPS devises and has compiled a very thorough network of geo-referenced roads across Africa

[10] Roads not appearing in AICD assumed to be tertiary, unpaved, and of poor quality. Given the state of DRC's road network, we believe this to be a very safe assumption.

In order to calculate the costs of traveling along the road network, the Highway Development Management Model (HDM-4), a standard model frequently used by road engineers, was applied. This model takes as inputs the road attributes available in the AICD data set, the roughness of terrain along the road, as well as country level information on various factors which can affect the price of transporting goods (i.e. price of fuel, labor costs, etc.). The output is the cost per kilometer of transporting a ton of goods in a heavy truck, for every possible road classification combination.[11] Using this model, the cost of traveling from every location within DRC to the cheapest market is calculated, where a market is defined as a city of 50,000 people.

4.2 The Natural-Historical Path instrument

Recognizing that roads are often placed where they will have the biggest economic impact, leading to a biased OLS estimate, we generate a new instrument for this paper, called the Natural-Historical path (NHP). As suggested by its name, the NHP takes into consideration historical data on caravan routes from the 19th century, as well as the terrain and historical land cover within DRC to estimate the quickest path, on foot, to and from anywhere within DRC's borders.

The literature on the economic benefits of roads typically relies on one of two types of IVs; straight line, or "Euclidean distance" IVs, and historical path IVs. Of late, thanks in part to greater access to digitized historical maps and books, the use of historical road IVs has been growing. While these two types of IVs are very different in their formulation, they are both attempts at estimating the same thing; namely, the natural way for humans to travel over land, absent the presence of a road network. Historical paths are useful as IVs, in that they represent the easiest path to travel over land. As these historical paths were constructed with little or no technology, they generally follow a smoother terrain and have been used for hundreds of years, and are thus the most cost effective route to construct a road. At the same time, they are usually not correlated with the current economic benefits that lead to the endogeneity bias, given that in many cases, these routes were constructed well over 100 years ago.[12][13]

A combination of both the natural path and historical caravan data, represents an improved estimate of how people traveled over land in prior centuries, and thus is the best possible IV for transportation cost. A major problem with only using historical path data is that people now live in areas

[11] See Appendix II for more information on the inputs and outputs for this model

[12] While these IVs are desirable, they are not always feasible if data on historical paths are unavailable. In these cases, researchers often rely on straight line IVs. These variables are usually correlated with historical paths, given that the quickest path between two points is usually the path with the shortest length—the straight line. However, they cannot account for the fact that the topography of the land may make traveling in a straight line impossible, or extra costly, making these IVs potentially quite weak.

[13] Some recent examples of papers employing historical route IVs include: Duranton et al. 2014, which used routes from major exploration expeditions in the US between the 16th and 19th centuries as instruments for the US interstate highway system; Garcia-Lopez et al. 2013, which used ancient Roman roads, amongst others, as exogenous sources of variation in Spain's current highway system; and Martincus et al. 2013, which used the Incan road network to instrument for Peru's current road infrastructure.

that may have been uninhabited, or not a part of the trade network, many years ago. Historical path data will therefore not be able to identify the likely paths that would have been used to travel to and from those locations. By using natural path data, we are able to fill in gaps in the historical caravan data, to get a complete picture of the optimal historical travel paths.[14]

4.3. Conflict measure

The study uses the Armed Conflict Location Events Dataset (ACLED) (Raleigh, 2010) version 4 which reports information on the location, date, and other characteristics of politically violent events for all countries on the African continent from 1997-2013.

Using the ACLED data set in its raw original point format brings up a host of technical and methodological problems. First, each conflict is pinned to a single geographic point, and does not capture the effects of conflict on the surrounding area. For instance, battles may have been fought over a large area, and its effects will be felt by an area significantly larger than a solitary point. Second, conflict points cannot capture conflict intensity (for instance, one isolated conflict point versus a cluster of conflict points, or one small conflict with 1 fatality versus a major battle with thousands). Finally, ACLED is subject to some geographic imprecision[15] resulting from how the data was obtained (for instance, conflicts occurring in rural areas are sometimes allocated to the nearest village).

To account for these methodological issues, we employ a kernel density function. This technique allows us to transform conflict points into a smooth surface, and generalize conflict locations. To calculate the value at any point, the kernel density function takes a weighted average of all the conflicts around that point, to create the surface. The magnitude of the weight declines with distance from the point, according to the chosen kernel function.[16] Figure 1 shows the original ACLED conflict data, as well as the kernel density map which was estimated using this technique.

Using these data we construct several measures of conflict. The first is the “kernelly” estimated number of fatalities in the 5 years preceding our DHS data set (2003-2007) and local GDP data set (2002-2006). We calculate this variable around each household and also around each market. We also generate a dummy variable which indicates if there are relatively high levels of conflict near households and

[14] For a complete explanation of how this variable was created, see Appendix III
[15] In ACLED geographic uncertainty level is coded with “geoprecision codes” ranging from one to three (higher numbers indicate broader geographic spans and thus greater uncertainty about where the event occurred). A geoprecision code of 1 indicates that the coordinates mark the exact location that the event took place. When a specific location is not provided, ACLED selects the provincial capital. This way ACLED may attribute violent incidents to towns when in fact they took place in rural areas and therefore introducing a systematic bias towards attributes associated with urban areas that can lead to invalid inferences.
[16] For instance, if a conflict occurs exactly on the point which is being calculated, the value of that conflict will receive a weight of 1. A conflict which is 5 kms away from the point will receive a weight of α and a conflict 10 kms away will receive a weight of β, where $1 > \alpha > \beta > 0$. Eventually, at some distance, referred to as the bandwidth, the weight becomes zero. For more information on how the kernel function and bandwidth were chosen, see Appendix IV

markets. It takes a value of 1 if the kernelly estimated fatalities are greater than the median number of fatalities due to violent conflict near both households around the nearest market.

4.4. Fractionalization

Conflict is another variable which, if not treated appropriately in a statistical regression, can lead to biases due to simultaneous causality. Conflict, for obvious reasons, can lead to lower investment levels, lower incomes, and lower welfare in general. At the same time, lack of economic opportunities and poor institutions (e.g. rule of law can) can also lead some to join rebel or insurgent militias.[17] This implies that conflict is likely to occur in poorer areas, and is also likely to depress these areas further, resulting in two-way causality. It is therefore essential to address these biases when estimating conflict's effect on economic variables.

Social fractionalization, which measures religious, ethnic, and linguistic diversity within a country or region, can be used as a valid instrument for conflict. Earlier studies have found ethno-linguistic fractionalization to be a strong determinant of the probability and duration of conflict (Wegenast and Basedau 2014; Esteban and Ray 2012; Esteban and Schneider 2008; Schneider and Wiesehomeier 2008; Montalvo and Reynal-Querol 2005; Reynal-Querol 2002; Gurr 2000; Collier and Hoeffler 1998; Horowitz 1985). The validity of the instrument is based on the observed correlation between higher levels of fractionalization and levels of conflict.[18] Our fractionalization variable is likely to satisfy the exclusion restriction, that is, causality could not run from welfare to fractionalization, as the fractionalization variable is generated using demographic data published in 2001 (details of the methodology used to generate this variable are provided below) while the income measures we use are from 2006 and 2007. Income measures in these years cannot affect predetermined, and hence exogenous, ethnic fractionalization.

Typically, fractionalization indices are calculated using nation-wide census data. However, given that this is an intra-country study, using one index for the entire country would not be useful. Instead, we calculate micro-level ethnic fractionalization; i.e. fractionalization within a 50km radius of households and markets. The distance of 50kms was chosen because that is the same radius used as the bandwidth for the conflict kernels, so we measure conflict and fractionalization within the same area. The 50km radius around each household is used as an instrument for conflict around each household, and the 50km radius around each market is used as an instrument for conflict around markets.

---

[17] The causes of conflict are not always cut and dry, however. While those who create conflict may originate from areas with low economic opportunities, often conflict arises around areas with wealth, as there are more opportunities for theft.

[18] Using a similar argument, Mauro (1995) uses ethnolinguistic fractionalization to instrument for institutional quality, when estimating its effect on GDP growth.

Given that ethnicity census data are not available for DRC, we take a spatial approach to estimating ethnic fractionalization and use the “People's Atlas of Africa”, developed by Felix and Meur (2001), which was later digitized and turned into a shapefile as part of Harvard’s Center for Geographic Analysis’s AfricaMap project.[19]

We use these data to estimate a fractionalization index similar to that of the Herfindahl Index of ethnolinguistic group shares (Alesina et al 2003), equation 4.1 below:

$$FRACT_w = 1 - \sum_K {p_k}^2 , \gamma \in 50km, \qquad (4.1)$$

where p_k is the percentage of land in which ethnicity k is the dominant ethnic group, within the 50km bandwidth of w, where w can be a household or a market. Higher (lower) values of the fractionalization index imply higher (lower) levels of ethnic fractionalization. In the extreme case, observe that if there is only one group that inhabits a region then $p_k = 1$, hence $FRACT_w = 0$. Specifically, this measure gives the probability that any two points of land chosen at random will have different dominant ethnic groups.[20] For example, to estimate the fractionalization index in Kisangani we first create a circle or buffer of 50km radius and we then estimate the area that each ethnicity occupies within that circle, to arrive at the percentage of land area that each tribe or ethnicity dominates in the areas surrounding Kisangani. We obtain a value of 0.83 which means that Kisangani is a very high level of ethnic fractionalization. Figure 2 shows visually how this was done.

4.5. Welfare indicators

Recognizing that no welfare indicator is perfect we estimate the welfare benefits of reducing transportation costs using different indicators from the Demographics and Health Survey (EDS-RDC), as well as data on local economic activity obtained from the Global Distribution of Economic Activity data set for the entire world developed by Ghosh et al. (2010). By using nighttime satellite imagery collected by the National Oceanic and Atmospheric Administration (NOAA) and LandScan’s population density grid, Ghosh et al. (2010) estimate a raster data set of local economic activity, which we refer to as local GDP. This data set spatially disaggregates DRC’s (among other countries) 2006 GDP into square pixels 30 arc seconds wide (approximately 1km^2), using the fact that brighter lights at night are associated with higher levels of economic activity (see Ghosh et al. 2010 for additional details about how these data were

[19] See https://worldmap.harvard.edu/data/geonode:etnicity_felix
[20] This index traditionally gives the probability that any two persons chosen at random will be from different groups (ethnic, religious, linguistic, etc. depending on what the researcher is studying). We modify this index slightly to accommodate for the fact that our ethnicity data does not observe individuals, but only land area. Rather than considering how many people from different ethnicities live in a certain area, we calculate how much of the land area belongs to a plurality of each ethnicity. We then replace the number of people from a given ethnicity variable in the fractionalization index with land area occupied by a given ethnicity.

generated). Given the granularity of our control data, we aggregated this data into square cells with sides measuring 5 arc minutes in length (approximately 10km).

For the EDS-RDC, conducted in DRC in 2007, 8,891 households were interviewed. The data are representative at the national level, for urban and rural residence, and for all eleven provinces. As EDS-RDC does not provide a direct measure of income, consumption or expenditures, we develop several alternative measures of welfare described below. From EDS-RDC we obtain two outcome variables: (i) a wealth index, and (ii) a multi-dimensional poverty indicator. The first indicator captures an economic measure of well-being as indicated by the level of wealth that households accrue over time, and arguably captures the longer term economic effects of infrastructure. It is calculated including all items in the EDS-RDC household questionnaire that are related to household goods, dwelling construction, and access to services and resources like electricity, water, and sanitation. The second indicator, our multi-dimensional poverty measure, summarizes both economic well-being as well as human capital, and captures a more holistic measure of living standards. We follow Alkire and Santos (2010) to calculate the Multi-Dimensional Poverty Index (MPI) for each household. The MPI is a weighted sum of ten indicators of deprivation across three dimensions: education, health, and standard of living. We adhere to convention and use equal weights for each of the three dimensions and for indicators within each dimension. A household is considered to be multi-dimensionally poor if it is deprived in three of the ten weighted indicators. Appendix I gives more specific details on how this index was constructed.

5. Empirical Framework

In estimating the effects of conflict and transportation costs on the wellbeing of households, we consider three alternative specifications. First, we examine whether the impact of conflict differs depending on whether it is located near markets, or near households. Then, we investigate the combined effect of both transportation costs and conflict on household wellbeing by including an interactive term. This section discusses each of these specifications in greater detail and explains our identification strategy.

5.1 Conflict near market versus conflict near household

Does the impact of conflict differ depending on whether it is located near the market or near the household? Presumably, the effects of conflict near markets could be avoided by retreating to subsistence modes of livelihood, whereas conflict near the home is likely more devastating because its negative effects cannot be avoided even by adopting subsistence farming. To investigate this, consider the following specification:

$$W_i = \beta_0 + \beta_1 T_i + \beta_M C_i^M + \beta_H C_i^H + X_i' \gamma + \varepsilon_i. \qquad (5.1)$$

Here, W_i represents the wellbeing of the household, measured by the wealth index, the multi-dimensional poverty index, or local GDP. Transportation cost to the nearest market is given by T_i and is instrumented as described earlier using the Natural-Historical Path variable. Conflict, measured in number of fatalities within a 50 km radius of the market and household, is given by C_i^M and C_i^H, respectively, also instrumented using the Fractionalization Index , and its squared term as well as the Euclidean distance to the eastern border as there is greater conflict along this border and because distance to this border should not directly impact income (for local GDP, the location of the household is merely replaced with the location of the centroid of the grid cell). The coefficients of these variables provide an indication of the differential effects of transport costs.

Finally, X_i denotes a vector of control variables that are likely to affect household wellbeing. Given the difference between household data and spatial data, different control variables were used in the regressions using the wealth index and the multi-dimensional poverty index, than with local GDP. For the first two, this vector includes agricultural variables, including agricultural potential (to account for the fact that areas with greater agricultural potential may naturally be more likely to have greater wealth) and a dummy indicating whether the household is engaged in agricultural activities (to account for the fact that agricultural households may accumulate wealth differently from others). We also control for household demographic characteristics such as the age of the household head, a binary variable indicating whether the household head is female, number of female members aged 15-49, number of male members ages 15-59, and number of children aged 0-5 (with all continuous variables estimated in log form). These variables help account for the fact that households with different demographic characteristics will have different propensities to accumulate wealth, as well as different levels of health and education. Finally, the regression includes a binary variable indicating whether the household is in a rural area and fixed effects indicating the geographic zone (to account for the unobserved characteristics of the area in which the household resides). For local GDP, controls include a quadratic term of population within the grid cell (to control for agglomeration benefits)[21], the agro-ecological potential yield of several important crops within the gridcell[22], the distance to the nearest mining facility[23], and province fixed effects.

---

[21] Population data is from Landscan and is available here: http://web.ornl.gov/sci/landscan/

[22] Agro-ecological potential data is from GAEZ, a product of FAO. It considers climate and soil conditions to estimate the maximum potential yields in each region for a large number of crops. The data used in this model assumes climactic conditions similar to the 1961-1990 baseline level, and is calculated assuming low input systems.

[23] Distance to the nearest mining facility is included because mines tend to be areas of great economic activity. In addition to the economic activity at the mine, mines can often generate economic spillovers for industries which service the mine and the mine's workers and their families. Data on mining facilities throughout DRC was obtained from the National Minerals Information Center of the USGS. The data set includes geo-referenced data on all mining facilities, active or closed, between 2006 and 2010. Because of the wide definition of what a mining facility actually is, we selected only a subset of mining facilities available to include in our data set. Facilities selected were

5.2 Joint effects of conflict and transportation costs

As an alternative, to examine interactions, we create a binary variable for high conflict areas and introduce an interactive term to the model. Specifically, Equation 5.1 is modified as follows:

$$W_i = \beta_0 + \beta_1 T_i + \beta_2 dC_i + \beta_3 T_i * dC_i + X_i'\gamma + \varepsilon_i \qquad (5.2)$$

Here, the binary variable dC_i takes the value of one if conflict near the household and markets are both greater than the median, and zero otherwise. The other variables are defined as in Equation 5.1 above. By including an interaction term, we are able to estimate how transport cost in areas with high level of conflict.

We instrument for cost to market using our Natural-Historical Path variable, and the instrument for conflict is the level of social fractionalization, and distance to east border. We include both a linear and squared term of fractionalization to account for the diminishing marginal effects of fractionalization on a fixed portion of land.

As a robustness check, we calculate a set of Conley Bounds for the coefficient of interest. To illustrate this, we can rewrite equation (5.1) as follows:

$$W_i = \beta_0 + \beta_1 T_i + \beta_M C_i^M + \beta_H C_i^H + X_i'\gamma + \lambda_0 NHP_i + \sum_{K=\{H,M\}} \lambda_{1K} F_i^K + \sum_{K=\{H,M\}} \lambda_{2K} \left(F_i^K\right)^2 + \sum_{K=\{H,M\}} \lambda_K^D D_i^K + u_i \qquad (5.3)$$

The traditional IV strategy assumes that the parameters $\lambda_0, \lambda_{1K}, \lambda_{2K}$, and λ_K^D in Equation 4 are all equal to zero. The Conley bound specification, shown in Conley et al (2012), allows these parameters to be close to zero, but not actually equal to zero. In other words, we allow the IVs to be only plausibly exogenous. By allowing the values of these parameters to vary, we can test how sensitive the estimates are to different degrees of exogeneity, thereby testing the validity of our instruments as well as our estimates.

6. Empirical Results

6.1 Main findings

The first indicator of household welfare we analyze is the "wealth index", which is readily available in the DHS data, and is an estimate of a household's long term standard of living. The second indicator, a multi-dimensional poverty measure, is generated specifically for this study following Alkire

those which involved the extraction of minerals or hydrocarbons from the ground. Mining facilities that were in the USGS data set but not included in this analysis include facilities like cement plants, or steel mills, which are likely concentrated in large cities or manufacturing areas. We also excluded plants that were labeled as being closed.

and Santos (2010). To further analyze the effect of transport costs and conflict, we analyze their effects on the decomposed MPI; namely on the standard of living, health, and education dimensions separately.

Table 3 presents the results from regressing the wealth index (columns 1 and 2) and multi-dimensional poverty indicator dummy (columns 3 and 4) on transport cost and conflict near markets and households (providing estimates of equation 5.1 above. Columns (1) and (2) (OLS and 2SLS, respectively) both indicate that transportation costs have a significantly negative effect on the wealth index. The IV result in column 2 shows that a 10 percent increase in transport cost to market decreases the wealth index by about 1.7 percent, significant at the 1% level. This indicates that households in areas with better access to markets are able to accrue more wealth than those with poor access to markets. Further, we find that conflict near households is highly detrimental to the wealth of households, as a 10 percent increase in conflict fatalities in the past 5 years decreases wealth by about 3.6 percent. Turning to the other control variables, we find that these IV results indicate that households engaged in agriculture, female-headed households and rural households tend to have less wealth. Additionally, holding all else equal, larger households have more wealth. As the number of male and female members in working age group increases, household wealth also increases. Conversely, as the number of dependents (i.e. children in the age range 0-5 years) increase, household wealth decreases.

Columns 3 and 4 report a significantly positive impact of transportation costs on the probability of being multi-dimensionally poor. The IV result in column 4 show that a 10 percent increase in transportation costs increases the probability of being multi-dimensionally poor by about 2.4 percent. Conflict—measured as the number of fatalities from violent conflict around markets—has a statistically significant positive effect on multi-dimensional poverty. Coefficients from the IV regression presented in column 4 indicate that a 10 percent increase in the number of fatalities around markets increases the probability of being multi-dimensionally poor by about 6 percent. Reflecting results from the Wealth Index we find that agriculturally involved households, female headed households and households in rural areas are more likely to be multi-dimensionally poor. Further, households with older heads are less likely to be multi-dimensionally poor. Larger households and households with more children in the age range 0-5 years are more likely to be multi-dimensionally poor.

Recall that the theoretical model (presented in section 3) suggests that the consequences of conflict vary with its geographical severity – between markets and households. To test for these effects we also analyze the combined effect/interactive effect of transportation cost on the wealth index and multi-dimensional poverty indicator. These results are shown in Table 4. Column 2 of Table 4 shows that when the conflict near both the household and the nearest market is low, the effect of transportation cost on the wealth index is negative and statistically significant. In this low conflict scenario, a 10 percent increase in transportation costs decreases the wealth index by 1.14 percent. When conflict near both the household and the nearest market is high the effect of transportation costs is given by the sum of the coefficients on transportation costs and the interaction between the high conflict dummy variable and transportation costs. In this high conflict scenario, a 10 percent increase in transportation costs increases the wealth index by 1.8 percent. This indicates that when there is high conflict near both the nearest market and the household, households that are farther away from markets are probably better off. This is likely because households farther away from markets are less affected by conflicts near markets. This is a new finding in the literature which suggests that in areas of high conflict, being near a road or a market may be detrimental to economic well-being. The coefficients of the control variables are similar to those reported for prior regressions.

To check whether the effect of reducing transportation cost and conflict on household wealth as estimated using EDS-DRC are robust, we provide estimates using local GDP (obtained from Ghosh et al. 2012). The results are in Table 5. Columns (1) and (2) present the OLS and IV results using data for all of DRC, while columns (3) and (4) present the results for only rural areas in DRC. Column (1) which presents the OLS results, shows a smaller elasticity of transport costs than the elasticity estimated using instrumental variables methodology. The IV estimate of transport cost elasticity indicates that when transportation cost decreases by 10 percent local GDP increases by 1 percent. This estimate is very close to the effect on wealth that was estimated for EDS-RDC[24]. Column (2) in Table 5 shows that a 10 percent increase in the number of fatalities will lead to a 1.2 percent decrease in local GDP. The first stage results presented in the second panel of Table 5 indicate that the instruments strongly predict transportation cost and fatalities in the area. All instruments pass the Angrist-Pischke F Test of Weak Identification, with the F statistic far exceeding 10, the rule of thumb. The results obtained for the entire sample are very similar to that obtained for the rural areas.

The regression results in for Table 6, seek to identify whether impacts differ by intensity of conflict as suggested by the DHS data. The regression includes a dummy variable for cases when the number of fatalities is higher than the mean[25] level of fatalities and an interaction term between this dummy and the natural logarithm of transport costs. The results suggest that when conflict (number of fatalities in the past five years) is below the mean, a ten percent decrease in transportation cost would increase local GDP by 1 percent. However, when conflict is above the average, a 10 percent reduction is transportation cost decreases local GDP by 2.2 percent. This indicates that when conflict is high, reducing transportation cost may not be a priority intervention, perhaps because as the model suggests, in these cases better roads enhance the payoffs from rebellion more than those from farming since opportunities for peaceful economic activities are limited in high conflict areas. When transportation costs are at the mean (US$45.1) a 10 percent increase in conflict leads to a 0.8 percent decrease in local GDP. Overall,

[24] The OLS estimated a counter-intuitive positive effect of conflict on local GDP while the IV estimate of the effect of conflict on local GDP is negative, further evidence that it is necessary to instrument for conflict

[25] The median value of conflict corresponding to the lights data set is zero hence the mean value is used instead which is very close to median value of conflicts near markets in EDS-RDC.

results from analysis of EDS-RDC data and local GDP data predict qualitatively and quantitatively consistent results.

6.2 Robustness checks

To check the robustness of the IVs to relaxation of the exclusion restriction, the Conley Bounds are calculated following Conley et al (2012), and reported in Tables 7. Table 7 shows the 95% confidence intervals for three endogenous variables -- transportation cost, number of fatalities near households and number of fatalities near markets – used in the estimation presented in Table 3. The 95% confidence interval suggests that the coefficient on the log of transportation remains consistently negative for the wealth index and consistently positive for the multi-dimensional poverty indicator, even if the assumption of exogeneity of the IV were not entirely accurate.

Similarly, Table 7 also shows that the coefficients on the number of fatalities near households and number of fatalities near markets remain consistently negative for the wealth index and consistently positive for multi-dimensional poverty. Taken together with the Angrist-Pischke F statistic and first stage results, these findings indicate that the instruments significantly predict the endogenous variables. The Conley bounds estimates shown in Table 7 allows us to test how sensitive the estimates are to different degrees of exogeneity and provide us the range of effects of transport cost and conflict when the exclusion restriction assumption is relaxed to some extent. To our relief, the effect of transport cost reduction remains positive on wealth index and negative on multi-dimensional poverty indicator and the effect of conflict reduction remains positive on wealth index and negative on multi-dimensional poverty as the exclusion restriction assumption is relaxed upto 1 percent[26].

7. Conclusion and Policy Implications

This paper presents new results on the effects of transportation infrastructure in areas of high conflict. It is widely assumed in policy circles, as indicated by the large investments, that rebuilding infrastructure that is damaged by prolonged conflict, or perhaps even constructing new infrastructure, could induce economic benefits. The results of this analysis, however, present a more nuanced story. We have shown that in areas with high conflict, the economic and social benefits of roads may be negated, and in some cases, actually reversed. While these results in the context of transport infrastructure are new, they are reminiscent of past research on the effects of foreign aid and conflict. Collier and Hoeffler (2004) show that although countries tend to receive a large influx of foreign aid in the years right after a civil war

[26] The range of possible values for δ follows previous application of the Conley Bound method in the literature (e.g. Emran et al 2009)." (p.8)

has formally ended, the assistance provides little benefit during this period due to limited absorptive capacity and risk of economic bottlenecks. Instead the benefits of foreign aid are greatest after a lapse of 4-7 years after the end of the civil war.

More generally however, in areas of DRC with low or no conflict, investment in decreasing transportation cost emerges as a highly effective way of generating economic growth. While it is unknown how current or future conflicts in DRC will evolve, this study has shown that upon the cessation of conflict, new road construction in one of the most isolated and disconnected countries in the world can be an important tool toward catalyzing growth and ending the perpetual conflict trap. So, it is perhaps important to emphasize that the results in this paper do not call for an unqualified proscription of road investments in all conflict-prone areas. But at a minimum, they do suggest the need to be mindful of unintended consequences and the desirability of combining such investments with a commitment to enhanced security in order to realize the benefits.

References:

Acemoglu, Daron, James A. Robinson, and Dan Woren. Why nations fail: the origins of power, prosperity, and poverty. Vol. 4. New York: Crown Business, 2012.
Ali, R. (2011). “Impact of Rural Road Improvement on High Yield Variety Technology Adoption: Evidence from Bangladesh,” Working paper, University of Maryland, College Park.

Alkire, S. & Santos, M.E. (2010). Acute multidimensional poverty: A new index for developing countries. OPHI Working paper 38.
Alderman, Harold, John Hoddinott, and Bill Kinsey. 2006. “Long Term Consequences of Early Childhood Malnutrition.” Oxford Economic Papers, 58(3): 450–74.

Alesina, Alberto, et al. "Fractionalization." Journal of Economic growth 8.2 (2003): 155-194.

Annan, Jeannie, Christopher Blattman, and Roger Horton. 2006. “The State of Youth and Youth Protection in Northern Uganda: Findings from the Survey for War Affected Youth.” Kampala, Uganda: UNICEF

Banerjee, A., Duo, E., & Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China. Yale Department of Economics mimeo.

Bannon, Ian, and Paul Collier, eds. Natural resources and violent conflict: options and actions. World Bank Publications, 2003.

Barron, Patrick, Kai Kaiser, and Menno Pradhan. 2004. “Local Conflict in Indonesia: Measuring Incidence and Identifying Patterns.” World Bank Policy Research Working Paper 3384.

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

Brück, T. (2004a), “Coping Strategies Post-War Rural Mozambique”, HiCN Working Paper No. 02, Households in Conflict Network, University of Sussex, UK

Brück, T. (2004b), “The Welfare Effects of Farm Household Activity Choices in Post-War Mozambique”, HiCN Working Paper No. 04, Households in Conflict Network, University of Sussex, UK (www.hicn.org).

Bruck, Tilman. 1996. “The Economic Effects of War.” Unpublished.

Bundervoet, Tom, Philip Verwimp, and Richard Akresh. 2009. “Health and Civil War in Rural Burundi.” Journal of Human Resources, 44(2): 536–63.

Bundervoet, T. and Verwimp, P. (2005), “Civil War and Economic Sanctions: An Analysis of Anthropometric Outcomes in Burundi”, HiCN Working Paper no. 11, Households in Conflict Network, University of Sussex, UK (www.hicn.org).

Casaburi, Lorenzo and Glennerster, Rachel and Suri, Tavneet, Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone (August 13, 2013). Available at SSRN:http://ssrn.com/abstract=2161643 or http://dx.doi.org/10.2139/ssrn.2161643

Cerra, Valerie, and Sweta Chaman Saxena. 2008."Growth Dynamics: The Myth of Economic Recovery." American Economic Review, 98(1): 439–57.

Chainey

Collier, P. (1999), "On the Economic Consequences of Civil War", Oxford Economic Papers 51(1): 168-183.

Collier, Paul, and Anke Hoeffler. "Aid, policy and growth in post-conflict societies." European economic review 48.5 (2004): 1125-1145.

Collier, Paul, and Anke Hoeffler, ''On Economic Causes of Civil War,'' Oxford Economic Papers, 50 (1998), 563–573.

Collier, P., Elliot, L., Hegre, H., Hoeffler, A., Reynal-Querol, M. and Sambanis, N. (2003), Breaking the Conflict Trap: Civil War and Development Policy, World Bank, Oxford University Press.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi (2012) "Plausibly Exogenous", The Review of Economics and Statistics, 94(1): 260-272
Congo Free State, 1885-1908 (http://www.yale.edu/gsp/colonial/belgian_congo/)

Datta, Saugato (2012), "The Impact of Improved Highways on Indian Firms", Journal of Development Economics, 99(1): 46-57.

Deininger, K. (2003), "Causes and Consequences of Civil Strife: Micro-Level Evidence from Uganda", Oxford Economic Papers 55: 579-606.

Deininger, Klaus W., and Derek Byerlee. Rising global interest in farmland: can it yield sustainable and equitable benefits?. World Bank Publications, 2011.

Do, Quy-Toan, and Lakshmi Iyer. 2007. "Poverty, Social Divisions and Conflict in Nepal." Harvard Business School Working Paper 07-065.

Dorosh, Paul, et al. "Road connectivity, population, and crop production in Sub-Saharan Africa." Agricultural Economics 43.1 (2012): 89-103.

Dave Donaldson, 2010. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure", NBER Working Papers,16487, National Bureau of Economic Research, Inc.

Deininger, Klaus W., and Derek Byerlee. Rising global interest in farmland: can it yield sustainable and equitable benefits?. World Bank Publications, 2011.

Donaldson, D. (2013). Railroads of the raj: Estimating the impact of transportation infrastructure. American Economic Review, forthcoming.

Dube, O. and Vargas, J. (2013) Commodity price shocks and civil conflict: evidence from Colombia, Review of Economic Studies, 80, 1384–421.

Duranton, Gilles, Peter M. Morrow, and Matthew A. Turner. "Roads and Trade: Evidence from the US." The Review of Economic Studies 81.2 (2014): 681-724.

Eck, Kristine. "In data we trust? A comparison of UCDP GED and ACLED conflict events datasets." Cooperation and Conflict 47.1 (2012): 124-141.

Emran, Shahe, Fenohasina Maret, and Stephen Smith (2009) “Education and Freedom of Choice: Evidence from Arranged Marriages in Vietnam,” Working Paper

Esteban, Joan, Mayoral, Laura, Debraj Ray. 2012. Ethnicity and Conflict: An Empirical Study. American Economic Review 102:4, 1310-1342.

Esteban, Joan, and Debraj Ray. 1999. “Conflict and Distribution.” Journal of Economic Theory, 87(2): 379–415.

Esteban J and Schneider G (2008) Polarization and conflict: Theoretical and empirical issues. Journal of Peace Research 45(2): 131–141.

Faber, Benjamin, (2014), “Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System,” The Review of Economic Studies, 0, 1–25

Fan, S., P. Hazell, and S. Thorat (2000). “Government Spending, Growth and Poverty in Rural India,” American Journal of Agricultural Economics 82 (4), 1038{1051.

Fearon J 2007 Economic Development Insurgency and Civil War. In Helpmann Elhanan (eds) *Institutions and Economic Performance* Harvard University Press

Felix, M. L., & Meur, C. (2001). Peoples of Africa: Enthnolinguistic map. Brussels: Congo Basin Art History Research Center.

Foster, Vivien, and Daniel Benitez. "Congo, Democratic Republic of-The Democratic Republic of Congo's infrastructure: a continental perspective." World Bank Policy Research Working Paper 5602 (2011).

GAO Report, (2008), “Afghanistan reconstruction: Progress Made in Constructing Roads but Assessments for Determining Impact and a Sustainable Maintenance Program are Needed”, United States Government Accountability Office.

Gachassin, Marie and Najman, Boris and Raballand, Gaël, The Impact of Roads on Poverty Reduction: A Case Study of Cameroon (February 1, 2010). World Bank Policy Research Working Paper Series, Vol. , pp. -, 2010. Available at SSRN: http://ssrn.com/abstract=1559726

Garcia-López, Miquel-Àngel, Adelheid Holl, and Elisabet Viladecans-Marsal. Suburbanization and highways: when the romans, the bourbons and the first cars still shape Spanish cities. No. 2013/5. 2013.

Gibson, John and Scott Rozelle, (October 2003), “Poverty and Access to Roads in Papua New Guinea,” Economic Development and Cultural Change, Vol. 52, No. 1, pp. 159-185

Gersony, Robert. 1997. “The Anguish of Northern Uganda: Results of a Field-Based Assessment of the Civil Conflicts in Northern Uganda.” Report submitted to the United States Embassy, Kampala.

Ghosh, Tilottama, et al. (2010) "Shedding light on the global distribution of economic activity." The Open Geography Journal 3.1: 148-161.Hoeffler, A. and Reynal-Querol, M. (2003), “Measuring the Costs of Conflict”, Oxford University.

Gurr TR (2000) Peoples Versus States: Minorities at Risk in the New Century. Washington, DC: United States Institute of Peace Press.

Gwatkin, D.R., S. Rutstein, K. Johnson, R.P. Pande, and A. Wagstaff. 2000. Socio-economic differences in health, nutrition, and population. HNP/Poverty Thematic Group. Washington, D.C.: World Bank.

Horowitz DL (1985) Ethnic Groups in Conflict. Berkeley, CA: University of California Press.

Humphreys, Macartan, and Jeremy M. Weinstein. 2008. “Who Fights? The Determinants of Participation in Civil War.” American Journal of Political, Science, 52(2): 436–55.

Jacoby, H. G. and B. Minten (2009). “On Measuring the Benefits of Lower Transport Costs”. Journal of Development Economics 89 (1).

Jacoby, H.G., (2000). “Access to markets and the benefits of rural roads”. The Economic Journal 110, 713–737.

Justino, Patricia. 2007. “On the Links between Violent Conflict and Household Poverty: How Much Do We Really Know?” Micro Level Analysis of Violent Conflict, Working Paper 1.

Justino, P. (2006), “Collective Action and Economic Development”. World Development, Volume 34, Issue 7, July 2006, Pages 1254–1270.

Justino, P. and Verwimp, P. (2006), “Poverty Dynamics, Conflict and Convergence in Rwanda”, Working Paper No. 16, Households in Conflict Network, Institute of Development Studies at the University of Sussex, Brighton (http://www.hicn.org/papers/wp16.pdf)

Khandker, Shahidur R., Zaid Bakht, and Gayatri Koolwal. 2006. “The Poverty Impact of Rural Roads: Evidence from Bangladesh.” Policy Research Working Paper Series 3875, World Bank, Washington, DC.

Knight, M., N. Loayza and D. Villanueva (1996). “The Peace Dividend: Military Spending Cuts and Economic Growth.” IMF Staff Papers 43(1): 1-37. Washington DC, United States.

Krueger, Alan B., and Jitka Maleckova. 2003. “Education, Poverty and Terrorism: Is There a Causal Connection?” Journal of Economic Perspectives, 17(4): 119–44.

Levine (2013)

Levine (2006)

Macours, K. 2011. “Increasing inequality and civil conflict in Nepal.” Oxford Economic Papers, 63(1): 1–26.

Martin, Philippe, Thoenig, Mathias and Mayer, Thierry, 2008, “Civil wars and International Trade” Journal of the European Economic Association Vol. 6, No. 2/3, pp. 541-550

Martincus, Christian Volpe, Jerónimo Carballo, and Ana Cusolito. "Routes, Exports, and Employment in Developing Countries: Following the Trace of the Inca Roads♦." Inter-American Development Bank, mimeograph (2012).

Miguel E, S Satyanath and E Sergenti 2004. “Economic Shocks and Civil Conflict” Journal of Political Economy, 112, 4, 725-753.

Montalvo JG and Reynal-Querol M (2005) Ethnic polarization, potential conflict, and civil wars. American Economic Review 95(3): 796–813.

Mauro, Paolo. "Corruption and growth." The quarterly journal of economics (1995): 681-712.

Maystadt J.-F., G. De Luca, P. Sekeris, and J. Ulimwengu. 2013. Mineral Resources and Conflicts in the Democratic Republic of Congo: A Case of Ecological Fallacy? Oxford Economic Papers. DOI:10.1093/oep/gpt037.

McKay, A. and Loveridge, S. (2005), “Exploring the Paradox of Rwandan Agricultural Household Income and Nutritional Outcomes in 1990 and 2000”, Staff Paper 2005-06, Department of Agricultural Economics, Michigan State University, Michigan.

Michaels, Guy (2008), “The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System”, Review of Economics and Statistics, 90(4): 683-701.

Minten, Bart, and Steven Kyle, (1999) “The effect of distance and road quality on food collection, marketing margins, and traders’wages: evidence from the former Zaire”, Journal of Development Economics, Vol. 60. 467–495

Montalvo, Jose G., and Marta Reynal-Querol. 2005.“Ethnic Polarization, Potential Conflict, and Civil Wars.” American Economic Review, 95(3): 796–816.

Murshed, S. Mansoob, and Scott Gates. 2005. “Spatial- Horizontal Inequality and the Maoist Insurgency in Nepal.” Review of Development Economics, 9(1): 121–34.

Mu, Ren, and Dominique van de Walle. 2007. “Rural Roads and Poor Area Development in Vietnam.” Policy Research Working Paper Series 4340, World Bank, Washington, DC.

Olga Shemyakina, “The effect of armed conflict on accumulation of schooling: Results from Tajikistan”, Journal of Development Economics, Volume 95, Issue 2, July 2011, Pages 186-200, ISSN 0304-3878, http://dx.doi.org/10.1016/j.jdeveco.2010.05.002.

Pushak, Nataliya and Briceno-Garmendia, Cecilia, The Republic of Congo's Infrastructure: A Continental Perspective (October 1, 2011). World Bank Policy Research Working Paper No. 5838. Available at SSRN:http://ssrn.com/abstract=1940341

Review of the Implementation Status of the Trans African Highways and the Missing Links. Rep. no. Volume 1. Stockholm: SWECO International AB, 2003.

Raleigh, Clionadh, et al. "Introducing ACLED: An armed conflict location and event dataset special data feature." Journal of peace Research 47.5 (2010): 651-660.

Reynal-Querol M (2002) Ethnicity, political systems, and civil wars. Journal of Conflict Resolution 46(1): 29–54.

Rutstein, Shea O. and Kiersten Johnson. 2004. The DHS Wealth Index. DHS Comparative Reports No. 6. Calverton, Maryland: ORC Macro.

Rutstein, Shea O. and Kiersten Johnson. 2004. The DHS Wealth Index. DHS Comparative Reports No. 6. Calverton, Maryland: ORC Macro.
Schneider G and Wiesehomeier N (2008) Rules that matter: Political institutions and the diversity–
conflict nexus. Journal of Peace Research 45(2): 183–203.

Silverman, B. W. Density Estimation for Statistics and Data Analysis. New York: Chapman and Hall, 1986. p. 76, equation 4.5
<http://books.google.com/books?id=exsrjsL7WkC&lpg=PP1&pg=PA76#v=onepage&q&f=false
>.

Shreshtha, Slesh A. (2012). Access to the North-South Roads and Farm Profits in Rural Nepal. Working Paper, National University of Singapore.

Stifel, David, and Bart Minten. 2008. “Isolation and Agricultural Productivity.” Agricultural Economics, 39 (1): 1–15.

Ulimwengu, John, Funes, Jose, Headey, Derek and Liandzhu, You, 2009, “Paving the Way for Development? The Impact of Transport Infrastructure on Agricultural Production and Poverty

Reduction in the Democratic Republic of Congo”, IFPRI Discussion Paper 00944.
Verwimp, Philip. 2003. “Testing the Double-Genocide Thesis for Central and Southern Rwanda.” Journal of Conflict Resolution, 47(4): 423–42.

Stewart F (2002) Horizontal inequalities: A neglected dimension of development. QEH Working Paper Number 81. Queen Elizabeth House, University of Oxford.

UNEP (2011) “UNEP Study Confirms DR Congo’s Potential as Environmental Powerhouse but Warns of Critical Threats”
<http://www.unep.org/newscentre/Default.aspx?DocumentID=2656&ArticleID=8890>

Verwimp, Philip. 2005. “An Economic Profile of Peasant Perpetrators of Genocide: Micro-level Evidence from Rwanda.” Journal of Development Economics, 77(2): 297–323.

Warr, Peter, (2008), “How Road Improvement Reduces Poverty: The Case of Laos,” Agricultural Economics, 39, pp. 269-279.

Wegenast, Tim C. and Matthias Basedau. (2014). "Ethnic fractionalization, natural resources and armed conflict" Conflict Management and Peace Science, Vol. 31(4) 432–457

Table 2: Summary Statistics of variables used in analysis of EDS-RDC and local GDP data

Variable	Mean	Std. Dev.	Min	Max
Outcomes:				
Wealth Index	4621.605	103646.1	-107521	345191
Dummy: Multi-dimensional poverty indicator	0.816736	0.386912	0	1
Wealth deprivation score	0.7273733	0.280303	0	1.002
Health and educational deprivation score	0.3539926	0.223059	0	1
Variables of interest:				
Cost to market	19.83385	21.69498	0.7342197	124.2525
No. of fatalities around the nearest market in the past 5 years	0.0438948	0.144152	0	0.9066222
No. of fatalities around household in the past 5 years	0.0171938	0.072611	0	0.8148879
Dummy=1 if conflict near household and market is high	0.328478	0.469695	0	1
Instruments:				
Time taken to reach market using natural and historical path	23.55944	26.48469	0.8028989	152.413
Fractionalization index within 50 km radius of market	0.6114013	0.207289	0	0.8391254
Fractionalization index within 50 km radius of household	0.5741688	0.20693	0	0.8456575
Distance to east border from household (km)	880.9615	486.3099	0.6008699	1715.954
Distance to east border from market (km)	882.179	486.5322	2.126719	1682.461
Controls:				
Agricultural potential (factor of ln agri. potential for cassava, maize and ground nut)	-0.0096933	1.007303	-7.474744	1.851499
Dummy: HH agriculturally involved	0.5426576	0.498214	0	1
Age of HH head	41.4731	13.03215	15	93
Dummy: Female HH head	0.188912	0.391468	0	1
No. of HH members	5.846463	2.960631	1	28
No. of female HH members aged 15-49 yrs	1.344828	0.855208	0	9
No. of female HH members aged 15-59 yrs	0.6578478	0.940663	0	13
No. of children in HH aged 0-5 yrs	0.5815993	0.936135	0	10
Dummy=1 if type of residence is rural	0.5551427	0.496987	0	1
Dummy=1 if region is bas-congo	0.0954221	0.293819	0	1
Dummy=1 if region is bandundu	0.1068668	0.308967	0	1
Dummy=1 if region is equateur	0.1037455	0.304953	0	1
Dummy=1 if region is orientale	0.0771403	0.266834	0	1
Dummy=1 if region is nord-kivu	0.0324019	0.177078	0	1
Dummy=1 if region is maniema	0.0936385	0.291347	0	1
Dummy=1 if region is sud-kivu	0.0425089	0.201762	0	1
Dummy=1 if region is katanga	0.1086504	0.311223	0	1
Dummy=1 if region is kasaï oriental	0.1030024	0.303985	0	1
Dummy=1 if region is kasaï occident	0.0921522	0.289262	0	1
No. of observations: 6728				

Table 2 (continued)

	Mean	Std. Dev.	Min	Max	Label
Local GDP	0.8355642	23.57658	0	3058.596	Total income per cell, millions USD (2006 PPP)
Transportation cost to market	45.10281	28.44443	0.3138928	165.4458	Transportation cost in US dollars
No. of fatalities within cell	0.0063334	0.0268568	0	0.3062945	
Population	2517.983	28082.6	0	3,327,414.00	Population per cell (thousands), Landscan 2009
Cassava Potential Yield	6960.317	2120.109	0	11161	Yield Kg/ha, GAEZ FAO
Groundnut Potential Yield	184.5166	124.6515	0	945	Yield Kg/ha, GAEZ FAO
Maize Potential Yield	949.325	362.9007	0	3380	Yield Kg/ha, GAEZ FAO
Distance to nearest mine	337.3761	147.4657	1.270597	500	Distance in km
No. of Observations: 26,330					

Table 3: Effect of transportation cost and conflict on welfare indicators in Democratic Republic of Congo

	Wealth Index		Multi-dimensional poverty indicator	
	(1)	(2)	(3)	(4)
VARIABLES	OLS	IV	OLS	IV
Log of transport cost	-0.0903***	-0.170***	0.0222***	0.0235**
	(-9.679)	(-10.04)	(4.094)	(2.398)
Log of no. of fatality near markets in the past 5 yrs (50 km Kernel)	0.0125	0.0203	0.00348	0.0613***
	(0.826)	(0.482)	(0.376)	(2.646)
Log of no. of fatality near households in the past 5 yrs (50 km Kernel)	0.0532***	-0.359***	0.00785	0.0230
	(2.998)	(-5.163)	(0.935)	(0.693)
Agri. potential	0.00435	-0.00634	-0.00201	-0.00175
	(0.395)	(-0.535)	(-0.346)	(-0.301)
Dummy: HH agriculturally involved	-0.384***	-0.437***	0.0898***	0.0980***
	(-16.69)	(-15.91)	(6.810)	(7.209)
Log of age of HH head	-0.0253	-0.0268	-0.0515***	-0.0508***
	(-0.859)	(-0.791)	(-2.870)	(-2.803)
Dummy: Female HH head	-0.139***	-0.147***	0.0531***	0.0523***
	(-5.880)	(-5.813)	(3.761)	(3.675)
log of no. of hh members	0.103***	0.0806***	0.0603***	0.0645***
	(3.763)	(2.674)	(3.491)	(3.716)
log of no. of female HH members aged 15-49 yrs	0.188***	0.207***	-0.0793***	-0.0825***
	(5.807)	(5.804)	(-4.136)	(-4.283)
log of no. of male HH members aged 15-59 yrs	0.130***	0.141***	0.0308**	0.0313**
	(6.288)	(6.336)	(2.292)	(2.322)
log of no. of children in HH aged 0-5 yrs	-0.109***	-0.105***	0.0312**	0.0290**
	(-5.240)	(-4.683)	(2.546)	(2.358)
Dummy: Rural area	-0.539***	-0.531***	0.0982***	0.110***
	(-19.65)	(-16.25)	(6.246)	(6.123)
Control for region dummies	Yes	Yes	Yes	Yes
Constant	12.55***	11.36***	0.571***	0.785***
	(107.7)	(48.09)	(8.586)	(8.960)
Observations	6,550	6,550	6,550	6,550
R-squared	0.654	0.599	0.251	0.243
Robust t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Table 3 (continued): first stage results

	Dependent variables:		
First stage results:	ln(Transportation cost)	ln(No. of fatalities near market)	ln(No. of fatalities near households)
ln (natural historical path)	0.977***	0.061***	-0.141***
	(300.450)	(9.950)	(-21.15)
Fractionalization near market	-0.047	-0.858***	1.292***
	(-0.51)	(-5.85)	(5.240)
Square of fractionalization near market	-0.218***	1.352***	-1.384***
	(-2.25)	(7.930)	(-5.25)
Fractionalization near households	0.264***	1.532***	-0.708***
	(3.690)	(13.130)	(-3.7)
Square of fractionalization near households	-0.058	-1.973***	1.235***
	(-0.8)	(-15.4)	(6.090)
ln (Distance to east border from household in km)	-0.016***	-0.279***	-0.282***
	(-2.09)	(-16.62)	(-15.86)
ln (Distance to east border from market in km)	-0.043***	-0.477***	-0.166***
	(-3.02)	(-17.18)	(-3.52)
Angrist Pischke test of weak identification			
F(5, 6523)	9559.55	179.92	121
Prob > F	P-value=.0000	P-value=.0000	P-value=.0000

Table 4: Combined effect of transport cost and conflict on wealth index and multi-dimensional poverty

	(1)	(2)	(3)	(4)
VARIABLES	OLS	IV NP	OLS	OLS NP
Log of transport cost	-0.0865***	-0.114***	0.0199***	0.0224***
	(-9.241)	(-8.508)	(3.550)	(3.119)
Dummy: Conflict near household and markets high	0.170***	-0.00177	-0.0167	-0.00709
	(4.772)	(-0.0173)	(-0.636)	(-0.147)
High conflict dummy interacted with log of transport cost	-0.00812	0.132***	-0.000246	-0.00571
	(-0.583)	(4.028)	(-0.0239)	(-0.308)
Agricultural potential	0.0123	0.0202	-0.00289	-0.00316
	(1.069)	(1.639)	(-0.486)	(-0.487)
Dummy: Household agriculturally involved	-0.371***	-0.358***	0.0851***	0.0841***
	(-16.38)	(-14.87)	(6.404)	(6.266)
Log of age of household head	-0.0230	-0.0199	-0.0517***	-0.0518***
	(-0.798)	(-0.681)	(-2.886)	(-2.908)
Dummy: Female household head	-0.139***	-0.142***	0.0552***	0.0555***
	(-6.010)	(-6.109)	(3.985)	(4.007)
Log of no. of household members	0.0960***	0.0979***	0.0601***	0.0602***
	(3.591)	(3.630)	(3.472)	(3.490)
Log of no. of female household members aged 15-49 yrs	0.184***	0.187***	-0.0776***	-0.0776***
	(5.763)	(5.764)	(-4.055)	(-4.047)
Log of no. of male household members aged 15-59 yrs	0.132***	0.130***	0.0307**	0.0309**
	(6.443)	(6.315)	(2.293)	(2.308)
log of no. of children0_5 in HH aged 0-5 years	-0.106***	-0.105***	0.0331***	0.0330***
	(-5.225)	(-5.133)	(2.732)	(2.727)
Dummy: Rural area	-0.535***	-0.556***	0.0948***	0.0948***
	(-20.21)	(-20.72)	(6.164)	(6.128)
Control for region dummies	Yes	Yes	Yes	Yes
Constant	12.19***	12.24***	0.551***	0.545***
	(113.0)	(85.93)	(8.238)	(7.063)
Observations	6,728	6,728	6,728	6,728
R-squared	0.657	0.651	0.248	0.248
Robust t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Table 4 continued: first stage results

	Dependent variable:		
First stage results:	ln(Transportation cost)	Dummy: Conflict near households and markets	ln(Transportation cost)*Dummy: Conflict near households and markets high
ln (natural historical path)	0.952***	0.171***	0.929***
	(24.820)	(4.170)	(6.720)
Fractionalization near market	-2.768***	-0.958***	2.987***
	(-9.65)	(-1.9)	(2.440)
Square of fractionalization near market	2.319***	-0.863**	-5.067***
	(8.220)	(-1.75)	(-3.82)
Fractionalization near households	3.223***	2.611***	-0.364
	(13.330)	(6.030)	(-0.37)
Square of fractionalization near households	-2.667***	-0.353	1.32
	(-11.07)	(-0.83)	(1.240)
ln (Distance to east border from household in km)	0.365***	0.027	0.791***
	(8.600)	(0.770)	(5.850)
ln (Distance to east border from market in km)	-0.387***	-0.091***	-0.695***
	(-7.79)	(-2.29)	(-4.43)
ln (natural historical path)*Fractionalization near market	0.676***	0.273**	-0.638*
	(8.110)	(1.820)	(-1.73)
ln (natural historical path)*Square of fractionalization near market	-0.619***	0.212	1.230***
	(-7.84)	(1.460)	(3.090)
ln (natural historical path)*Fractionalization near households	-0.842***	-0.826***	-0.316
	(-11.4)	(-6.68)	(-1.14)
ln (natural historical path)*Square of fractionalization near households	0.722***	0.216**	0.129
	(9.950)	(1.820)	(0.450)
ln (natural historical path)*ln(Distance to east border from household)	-0.129***	-0.007	-0.250***
	(-8.92)	(-0.52)	(-5.22)
ln (natural historical path)*ln(Distance to east border from market)	0.104***	-0.007	0.092***
	(7.360)	(-0.6)	(1.920)
Angrist Pischke test of weak identification	5575.07	70.77	148.56
	P= 0.0000	P= 0.0000	P= 0.0000

Table 5: Effect of transportation cost and conflict on local GDP

	(1)	(2)	(3)	(4)
VARIABLES	OLS (Full sample)	IV (Full sample)	OLS (rural only)	IV (rural only)
ln(Cost to Market)	-0.0115**	-0.105***	-0.0118**	-0.0822***
	(-2.429)	(-5.425)	(-2.465)	(-4.523)
Ln(No. of fatalities)	0.00789***	-0.121***	0.00713***	-0.0902***
	(4.761)	(-4.464)	(4.311)	(-3.557)
ln(Population)	0.530***	0.530***	0.527***	0.528***
	(170.1)	(153.3)	(169.7)	(159.5)
ln(Population)^2	0.0359***	0.0357***	0.0365***	0.0363***
	(128.8)	(114.5)	(127.9)	(117.7)
ln(Cassava potential yield)	0.488***	0.548***	0.547***	0.595***
	(5.666)	(5.648)	(6.375)	(6.418)
ln(Cassava potential yield)^2	-0.0366***	-0.0395***	-0.0420***	-0.0444***
	(-5.774)	(-5.568)	(-6.649)	(-6.543)
ln(Maize potential yield)	0.451***	0.348**	0.367**	0.286*
	(3.003)	(2.065)	(2.459)	(1.775)
ln(Maize potential yield)^2	-0.0420***	-0.0388***	-0.0340***	-0.0311**
	(-3.686)	(-3.058)	(-2.995)	(-2.566)
ln(Groundnut potential yield)	-0.322***	-0.384***	-0.308***	-0.354***
	(-6.244)	(-6.562)	(-6.017)	(-6.360)
ln(Groundnut potential yield)^2	0.0422***	0.0509***	0.0406***	0.0468***
	(8.274)	(8.590)	(8.010)	(8.345)
ln(Distance to mine)	-0.0764***	-0.0298***	-0.0709***	-0.0346***
	(-13.55)	(-2.601)	(-12.52)	(-3.115)
Province fixed effects	Yes	Yes	Yes	Yes
Constant	-8.953***	-9.561***	-8.962***	-9.426***
	(-42.29)	(-35.67)	(-42.65)	(-37.00)
Observations	26,330	26,330	26,179	26,179
R-squared	0.920	0.901	0.916	0.905
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				
Data: Nighttime Lights local GDP, Ghosh (2010)				

First stage results:

	Full Sample		Rural only	
	(1)	(2)	(3)	(4)
Depedentent variables:	ln(Cost to Market)	Ln(No. of fatalities)	ln(Cost to Market)	Ln(No. of fatalities)
First stage results				
ln(Natural historical Path)	0.875***	-0.583***	0.873***	-0.584***
	(356.710)	(-34.95)	(352.070)	(34.700)
Fractionalization	-0.0003881	-0.006***	-0.000365	-0.006***
	(-1.44)	(-3.3)	(-1.36)	(3.120)
(Fractionalization index)^2	0.0000136***	0.00001	0.0000134***	0.000008
	(4.490)	(0.790)	(4.460)	(0.390)
Distance to east border	0.0044	-0.187***	0.0038	-0.186***
	(1.280)	(-7.97)	(1.050)	(7.660)
Angrist-Pischke Test of weak indentification				
F(4, 26306)	3672.93	354.56	31705.7	352.55
Prob>F	P-value=.0000	P-value=.0000	P-value=.0000	P-value=.0000
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				
Data: Nighttime Lights local GDP, Ghosh (2010)				

Table 6: Combined effect of transport cost and conflict on local GDP

	(1)	(2)	(3)	(4)
VARIABLES	OLS (Full sample)	IV (Full sample)	OLS (rural only)	IV (rural only)
ln(Cost to Market)	-0.00253	-0.104***	-0.00328	-0.0974***
	(-0.495)	(-8.984)	(-0.639)	(-8.062)
Dummy: Conflict is high=1	0.212***	-1.251***	0.204***	-1.248***
	-6.188	(-7.353)	(5.700)	(-6.540)
ln(Cost to Market)*Dummy: Conflict is high	-0.0540***	0.324***	-0.0524***	0.334***
	(-5.326)	-4.392	(-4.965)	-4.297
ln(Population)	0.531***	0.527***	0.528***	0.525***
	(170.3)	(159.2)	(169.9)	(162.0)
ln(Population)^2	0.0359***	0.0358***	0.0366***	0.0362***
	(128.8)	(120.0)	(128.1)	(120.4)
ln(Cassava potential yield)	0.500***	0.436***	0.558***	0.502***
	(5.813)	(4.830)	(6.506)	(5.608)
ln(Cassava potential yield)^2	-0.0374***	-0.0329***	-0.0427***	-0.0390***
	(-5.901)	(-4.983)	(-6.761)	(-5.938)
ln(Maize potential yield)	0.438***	0.489***	0.356**	0.398**
	(2.919)	(3.102)	(2.385)	(2.552)
ln(Maize potential yield)^2	-0.0411***	-0.0462***	-0.0332***	-0.0372***
	(-3.611)	(-3.879)	(-2.931)	(-3.152)
ln(Groundnut potential yield)	-0.330***	-0.300***	-0.316***	-0.279***
	(-6.398)	(-5.293)	(-6.168)	(-4.975)
ln(Groundnut potential yield)^2	0.0430***	0.0413***	0.0414***	0.0385***
	(8.423)	(7.347)	(8.162)	(6.936)
ln(Distance to mine)	-0.0771***	-0.0529***	-0.0718***	-0.0486***
	(-13.69)	(-7.205)	(-12.68)	(-6.612)
Province fixed effects	Yes	Yes	Yes	Yes
Observations	26,330	26,330	26,179	26,179
R-squared	0.920	0.914	0.916	0.911
Sargan statistic (overidentification test of all instruments):		108.047		104.147
		0.0000		0.0000
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				
Data: Nighttime Lights local GDP, Ghosh (2010)				

First stage results:

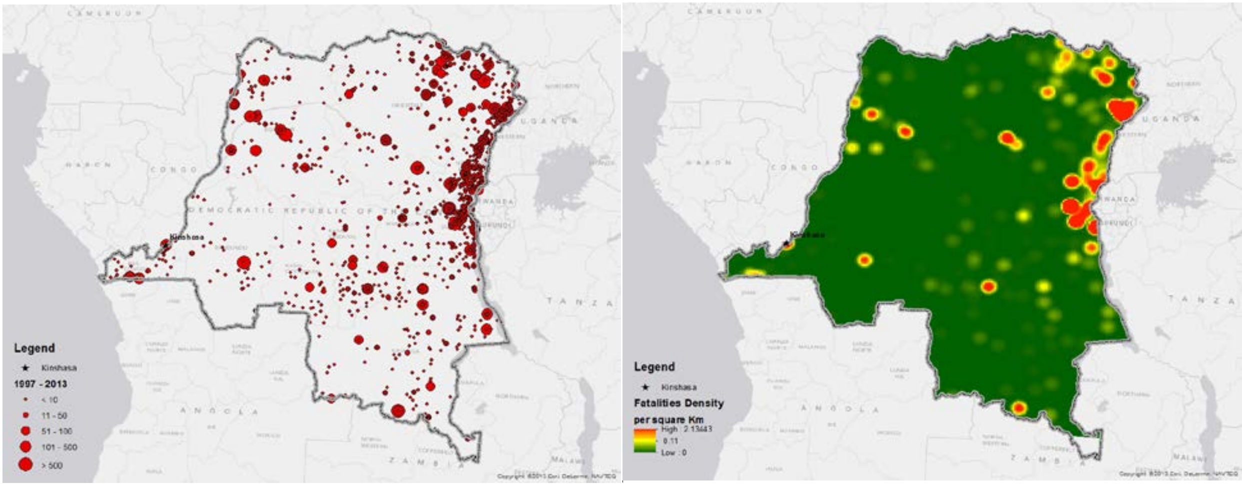
	Full Sample			Rural only		
	(1)	(2)	(3)	(4)	(5)	(6)
Depedentent variables:	ln(Cost to Market)	Dummy: Conflict is high	ln(Cost to Market)*Dummy: Conflict is high	ln(Cost to Market)	Dummy: Conflict is high	ln(Cost to Market)*Dummy: Conflict is high
First stage results						
ln(Natural historical Path)	0.985***	-0.823***	-1.07***	0.9091***	-0.850***	-1.23***
	(23.540)	(-17.04)	(-6.54)	(20.560)	(-16.7)	(-7.12)
Fractionalization	0.010***	-0.010***	-0.022***	0.01***	-0.010***	-0.023***
	(7.640)	(-6.68)	(-4.25)	(7.760)	(-6.52)	(-4.37)
(Fractionalization index)^2	-0.0001***	0.0001***	0.0002***	-0.00009***	0.0001***	0.0003***
	(-6.67)	(7.210)	(4.320)	(-6.67)	(6.600)	(4.400)
Distance to east border	0.0185*	-0.192***	-0.258***	-0.0016	-0.198***	-0.298***
	(1.700)	(-15.29)	(-6.06)	(-0.14)	(-14.84)	(-6.58)
ln(Natural historical Path)*Fractionalization	-0.003***	0.003***	0.007***	-0.0029759***	0.003***	0.007***
	(-8.1)	(7.130)	(4.910)	(-8.19)	(6.970)	(5.020)
ln(Natural historical Path)*(Fractionalization index)^2	0.00003***	-0.00004***	-0.0001***	0.000031***	-0.00004***	-0.0001***
	(7.760)	(-8.18)	(-5.56)	(-7.72)	(-7.58)	(-5.64)
ln(Natural historical Path)*Distance to east border	-0.004	0.052***	0.062***	0.0016	0.053***	0.073***
	(-1.36)	(14.630)	(5.190)	(-0.49)	(14.270)	(5.770)
Angrist-Pischke Test of weak indentification						
F(4, 26306)	18599.38	218.77	58.99	18171.91	213.06	63.55
Prob>F	P-value=.0000	P-value=.0000	P-value=.0000	P-value=.0000	P-value=.0000	P-value=.0000
t-statistics in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						
Data: Nighttime Lights local GDP, Ghosh (2010)						

Table 7: Conley Bounds

	Endogenous variable: travel cost		
Dependent variable:	IV: Natural path		
	δ: [-0.0001, 0.0001]	-0.14831738	-0.11838858
	δ: [-0.001, 0.001]	-0.14927312	-0.11743313
wealth index	δ: [-0.01, 0.01]	-0.15883091	-0.10787822
	Endogenous variable: conflict near markets		
	IV: Fractionalization near market, squared term of fractionalization near market, distance to east border from market		
	δ: [-0.0001, 0.0001]	-0.12900793	-0.09872961
	δ: [-0.001, 0.001]	-0.12919092	-0.0985378
wealth index	δ: [-0.01, 0.01]	-0.13102262	-0.09661798
	Endogenous variable: conflict near households		
	IV: Fractionalization near households, squared term of fractionalization near households, distance to east border from households		
	δ: [-0.0001, 0.0001]	-0.26683725	-0.19892408
	δ: [-0.001, 0.001]	-0.26779058	-0.19806024
wealth index	δ: [-0.01, 0.01]	-0.27734444	-0.18940086

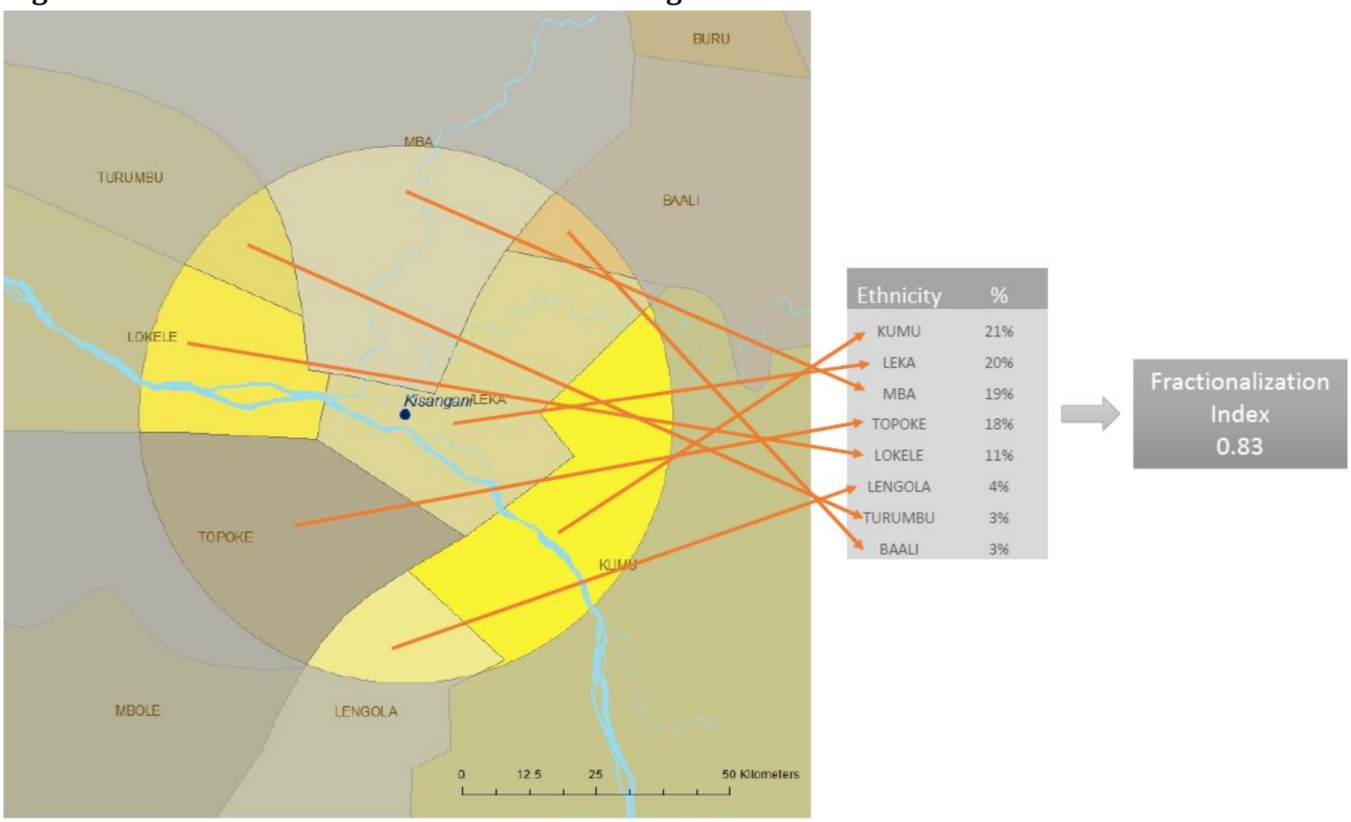
		Lower Bound	Upper Bound
	Endogenous variable: travel cost		
Dependent variable:	IV: Natural path		
	δ: [-0.0001, 0.0001]	0.0256589	0.04427896
	δ: [-0.001, 0.001]	0.02470355	0.04523484
Multi-dimensional poverty indicator	δ: [-0.01, 0.01]	0.01514969	0.0547939
	Endogenous variable: conflict near markets		
	IV: Fractionalization near market, squared term of fractionalization near market, distance to east border from market		
	δ: [-0.0001, 0.0001]	0.01648337	0.03508131
	δ: [-0.001, 0.001]	0.01629361	0.03526633
Multi-dimensional poverty indicator	δ: [-0.01, 0.01]	0.01439252	0.03711992
	Endogenous variable: conflict near households		
	IV: Fractionalization near households, squared term of fractionalization near households, distance to east border from households		
	δ: [-0.0001, 0.0001]	0.01922971	0.05855597
	δ: [-0.001, 0.001]	0.01832708	0.05947112
Multi-dimensional poverty indicator	δ: [-0.01, 0.01]	0.00926055	0.0686626

Figure 1: Estimated Conflict surface



Data source: Raleigh, Clionadh, et al (2010) and author calculations
Left image: Fatalities per violent conflict, 2003-2007. Right image: Kernel density estimate of conflict intensity

Figure 2: Ethnic Fractionalization around Kisangani



Data source: Felix, M. L., & Meur, C. 2001
The Felix and Meur ethnicity data around the city of Kisingani is shown, with a ring of 50km in radius drawn around the center of the city. There are 8 major tribes present in that area, and the percentage of land area which each tribe dominates within that circle is calculated. Those percentages are plugged into equation 4.1 to arrive at the fractionalization index around Kisangani. This calculation is repeated around each household in the DHS survey, and around each market.

Appendix I: Multi-dimensional poverty index construction

Table A1 describes the inputs into the multi-dimensional poverty index (MPI) used in this study. Each of the indicators given in the second column are binary variables, taking the value of 0 if the household is deprived in that vector, according to the third column, and 1 otherwise. The final MPI is generated by calculating the weighted sum of the ten indicators, with weights given in the fourth column. Weights were chosen so that each of the three dimensions—education, health, and standard of living—have equal weight, and each of the indicators have equal weight within each dimension. A household is considered to be multi-dimensionally poor if its weighted sum of indicators is greater than 3. Note that the weights add up to approximately 10, the number of indicators (difference due to rounding).

Table A1: Multi-dimensional Poverty Index

Dimension	Indicator	Deprived if…	Relative Weight
Education	Highest degree earned	No household member has completed five years of education.	1.67
	Child School Attendance	Household has a school-aged child not attending school	1.67
Health	Child Mortality	Household has had at least one child aged 0-5 years die in the past 5 years.	1.67
	Nutrition	Household has a malnourished woman aged (15-49) or child aged (0-5).	1.67
Standard of Living	Electricity	The household has no electricity.	0.56
	Improved Sanitation	Household does not have improved sanitation.	0.56
	Safe Drinking Water	Household does not have access to improved water source.	0.56
	Flooring	The household has a dirt floor.	0.56
	Cooking fuel	The household uses dirty cooking fuel.	0.56
	Asset Ownership	The household does not own more than one bicycle, motorcycle, radio, fridge, TV, or phone and does not own a car.	0.56

World Heath Organization (WHO) standards were followed in determining what to consider unimproved water sources, inadequate sanitation, and dirty cooking fuel.

Appendix II: Highway development and management model (HDM-4)

The Highway Development Management Model (HDM-4) considers several different variables in order to estimate the cost of traveling along each segment of the road network. The data used for the estimates used in this paper was collected specifically for DRC, to best characterize the transportation conditions one would find there.

In order to estimate the unit cost (in ton per km), the cost of transporting a vehicle with an average weight of 25 tonnes, one kilometer, was first estimated. The unit cost per ton-km was derived from the costs per vehicle using a factor of 15 ton per vehicle (average net weight). This factor was obtained based on the assumption of a 30 ton gross vehicle weight, with a 10 ton tare weight and a 75% loading factor.

Characterization of network type and terrain

The road network of DRC includes three classes of roads: primary, secondary, and tertiary. Average vehicle speed and width of the main carriage road were used to characterize the differences among network types as follows:

Paved Road Speed (km/hr) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Flat	100	80	70
Rolling	80	70	60
Mountainous	60	50	40

Unpaved Road Speed (km/hr) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Flat	80	70	60
Rolling	60	50	40
Mountainous	40	30	20

where terrain type is defined using the following concepts and road geometry parameters:

- Flat. Mostly straight and gently undulating
- Rolling. Bendy and gently undulating
- Mountainous. Winding and gently undulating

TERRAIN TYPE	**Number**			**Super_ elevation**
	Rise & Fall	**Rise & Fall**	**Horizontal Curvature**	
	(m/km)	**(#)**	**(deg/km)**	**(%)**
FLAT	10	2	15	2.5
ROLLING	15	2	75	3.0
MOUNTAINOUS	20	3	300	5.0

Characterization of network type and condition

The International Roughness Index IRI (m/km) was used to define the differences in road condition by network as follows:

Paved Road IRI (m/km) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Good	2	3	4
Fair	5	6	7
Poor	8	9	10

Unpaved Road IRI (m/km) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Good	6	8	10
Fair	12	13	14
Poor	16	18	20

Finally, using these parameters above, a final cost per ton-km for each road type is estimated ($/ton/km):

Paved FLAT			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.1174	0.1192	0.1237
Fair	0.1226	0.1264	0.1293
Poor	0.1286	0.1299	0.1349

Paved ROLLING			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.1190	0.1191	0.1231
Fair	0.1241	0.1268	0.1302
Poor	0.1305	0.1315	0.1367

Paved MOUNTAINOUS			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.1283	0.1292	0.1312
Fair	0.1333	0.1318	0.1382
Poor	0.1410	0.1391	0.1449

Unpaved FLAT			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.1401	0.1463	0.1559
Fair	0.1622	0.1755	0.1901
Poor	0.1976	0.2133	0.2290

Unpaved ROLLING			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.1348	0.1453	0.1588
Fair	0.1638	0.1771	0.1921
Poor	0.1991	0.2147	0.2305

Unpaved MOUNTAINOUS			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.1390	0.1570	0.1806
Fair	0.1681	0.1857	0.2091
Poor	0.2014	0.2186	0.2379

Appendix III: Natural-historical path

Since the early arrival of the Portuguese Mariner Diego Cão in 1483, the Congolese (Kingdom of Kongo at that time) has had cultural, social and economic connections with Europe. Western religions, literacy, the wheel, the plow, the gun and many other technologies were quickly adopted by the Congolese (Acemoglu and Robinson 2012).[27] All these came at very large expense: one of the principally traded goods in exchange were slaves[28].

As their contact deepened other types of goods were introduced such as ivory, rubber, copper, diamonds, raffia cloth, and pottery among other natural resources. The European trade was based in the coastal cities of Sonyo and Pinda so it required an extensive trade network toward the eastern part of the country (primarily near present day Kivu and Katanga provinces) where much of the mineral deposits and other natural resources were mined. Fueled by the industrial revolution and new inventions such as the inflatable rubber tubes, the demand for goods increased dramatically.[29] By the end of the 19th century the Congo was a personal possession of King Leopold II (not an official Belgian Colony). The king was engaged in a vigorous publicity campaign aimed at convincing the other European powers to recognize the legitimacy of his rule, a difficult task in view of the notorious brutality of his administration in Africa. One of the products of King Leopold's "Office of Publicity" is a very detailed "Carte du Congo Belge"[30] (Map of the Belgian Congo), which includes caravan routes and existing and projected railways (see figure A4). This map, which shows the transport network constructed to move slaves, ivory and mineral resources between the interior of Congo and the coastal harbors, is the main input used to construct the natural-historical instrumental variable.

The natural-historical IV is constructed by merging two sources: the historical caravan route map described above, and a natural walking path map calculated for this study. The natural walking path, or natural path, is created by estimating the time minimizing route a pedestrian would travel over land, absent the benefits of a road network. The subsequent section details how the natural path network was calculated.

Natural Path walking time calculation

The first step is creating a GIS cost surface model that accounts for all the traffic off-road or outside the caravan network for the year 1900. For that purpose, this report followed a similar approach to that which was used to construct a global map of accessibility in the World Bank's World Development

[27] Why Nations Fail. Acemoglu and Robinson 2012
[28] "The wealth of Africa, The kingdom of Kongo". The British Museum 2011.
[29] Congo Free State, 1885-1908 (http://www.yale.edu/gsp/colonial/belgian_congo/)
[30] Carte du Congo Belge / éditée par l'Office de Publicité, anciens établissements J. Lebègue & Cie. - Éditeurs, Bruxelles (1896). Stored at the Library of Congress and downloaded from http://www.wdl.org/en/item/59/

Report 2009 Reshaping Economic Geography (Uchida and Nelson 2009).[31] The surface model, or off-path friction-surface raster,[32] is a grid where each pixel contains the estimated time required to cross that pixel by walking. To create this raster two basic layers are combined: terrain slope and land cover.

The slope raster was calculated from NASA's Shuttle Radar Topography Mission (SRTM)[33] Digital Elevation Models (DEMs) with a resolution of 90 meters. Even though the original topography data was obtained in February 2000, it was assumed that there has not been any drastic change in DRC's terrain in the 20th century. Therefore, it is safe to assume that the SRTM 90-meters data set provides a fairly good representation of the elevation terrain circa 1900.

Land cover data is far more challenging. In the last few years, with the surge of remote sensing technology, several land cover and land use data sets have been created. These high-resolution data sets are a very accurate representation of the current state of physical material at the surface of the earth. However, these cannot be used in the analysis because they are not a good representation of the land cover for year 1900. Land has changed rapidly in the last 100 years in DRC: deforestation, open pit mining and urbanization in some cities have drastically transformed the surface. Therefore, other data sets were used. The Oak Ridge National Laboratory (ONRL)[34] has developed a Historical Land Cover and Land Use data estimate (Goldewijk 2010). This data set describes historical land use changes over a 300-year historical period (1700-1990) and was modeled based on a deep understanding of the global environment, historical statistical inventories on agricultural land (census data, tax records, land surveys, etc), and different spatial analysis techniques. A shortcoming of this data set is that the resolution is approximately 55km per cell, making a clear tradeoff between space (resolution) and time (representative for 1900). Given the importance of obtaining an accurate picture of the historical land cover, this low resolution, but better account of the land surface types circa year 1900 was chosen (i.e. ORNL Historical Land Cover) over the better resolution but newer data. Figure A1 displays the terrain roughness (left) and land cover (right) for DRC, circa 1900.

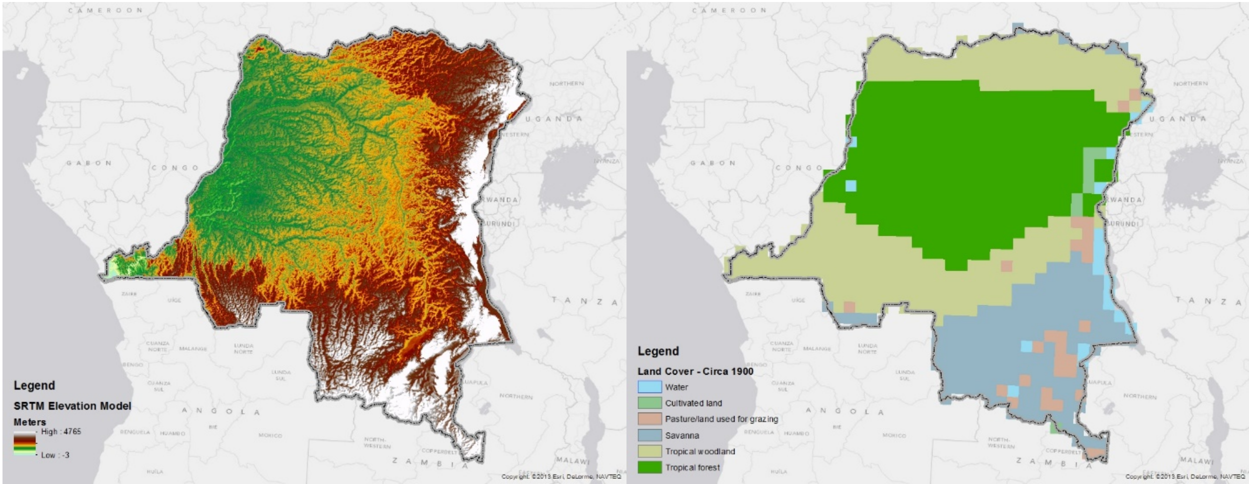
[31] Uchida, H. and Nelson, A. Agglomeration Index: Towards a New Measure of Urban Concentration. Background paper for the World Bank's World Development Report 2009

[32] A raster is a geographic map with information (e.g. elevation, landcover) in a matrix of pixels.

[33] http://www2.jpl.nasa.gov/srtm/

[34] ORNL is a multiprogram science and technology laboratory managed for the U.S Department of Energy by UT-Battelle,LLC

Figure A1: Terrain roughness and land cover circa 1900



Source: Authors' calculations using NASA and ORNL

The off-path friction-surface raster is created by combining the land topography raster and the land cover map. This is based on the guiding assumption that all travel in 1900 was on foot and walking speed is therefore determined by the land cover class and slope. The typical velocity of a hiker when walking on uneven or unstable terrain is 1 hour for every 4 kilometers (4 km/hr) and diminishes on steeper terrain. A hiking velocity equation35 (Tobler 1993) was used to reflect changes in travel speed as a function of trail slope:

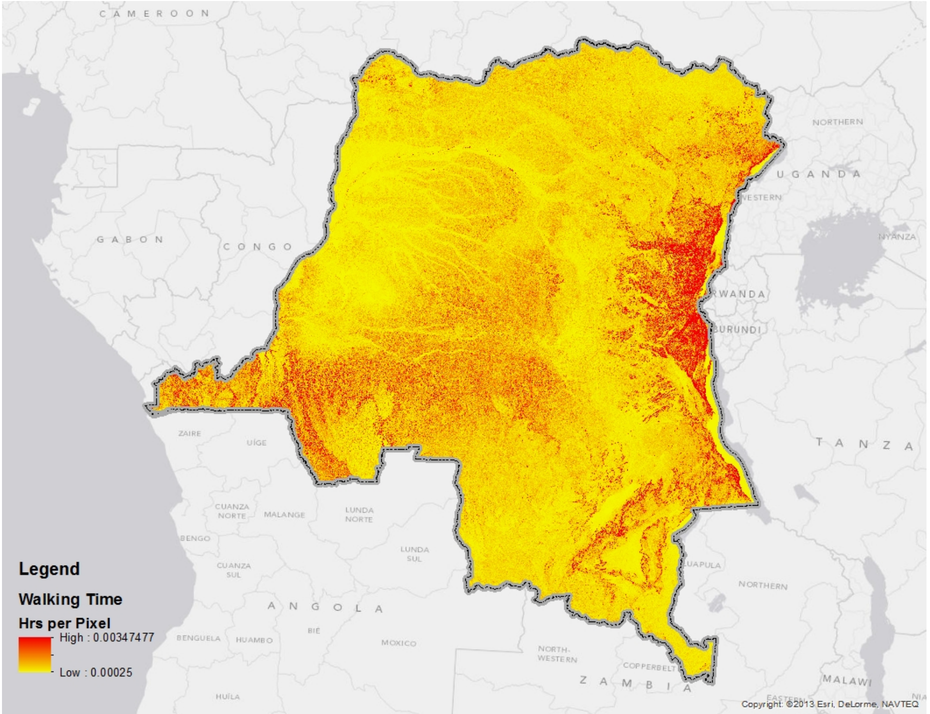
$$W = 6*\exp(-3.5* |S + 0.05|)$$

where W is the hiking velocity in km/hr and S is the slope or gradient.

By applying the speed formula, he time it required to cross 1 pixel (92.5 meters) was computed. In this way, the time (hours) that it takes to walk through any pixel— only taking into account the topography—was calculated, as shown below in Figure A2. Note that the more mountainous regions of DRC near the Kivu provinces in the east have significantly higher walking times.

[35] "Three presentations on geographical analysis and modeling non-isotropic geographic modeling speculations on the geometry of geography global spatial analysis". National Center for Geographic Information and Analysis. Technical Report 93-1. February 1993.

Figure A2: Walking time in DRC given land topography (hours per pixel)



Source: Authors' calculations

Next, a delay factor to account for effect of walking through different land classes was estimated. The historical land cover raster resolution was changed from half degree to 90 meters and each class was assigned a speed reducing factors according to the following table:

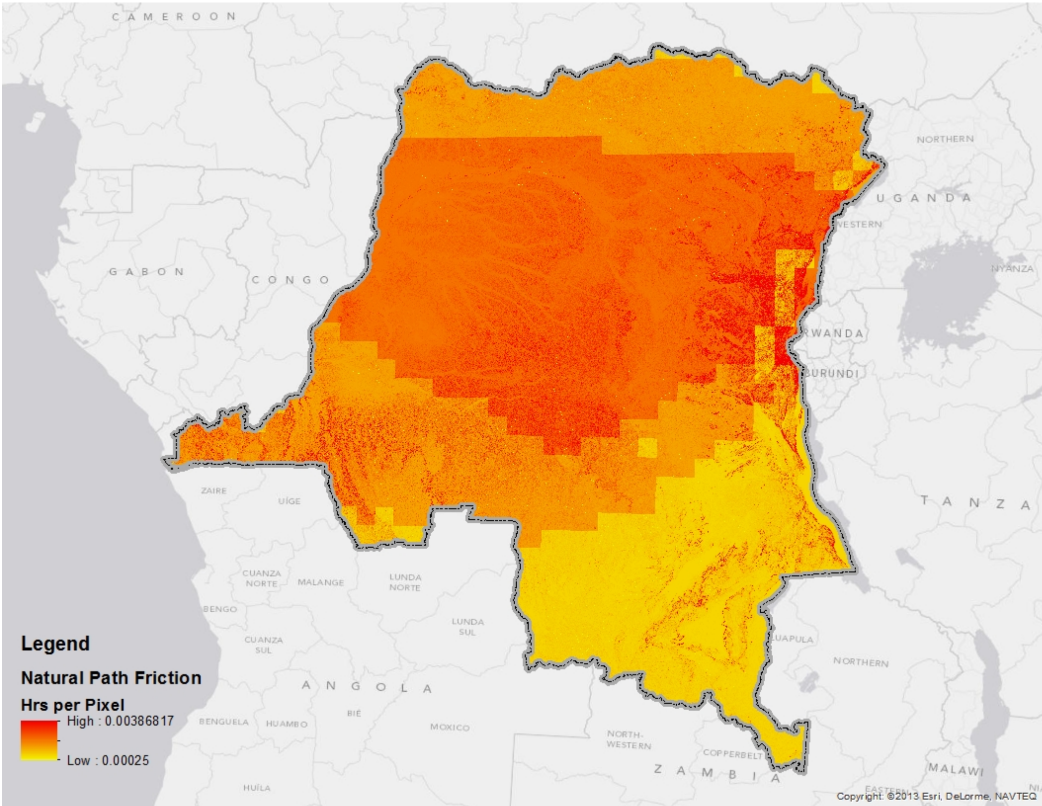
Class #	Biome Type	Delay Factor
0	Oceans/Water	N/A
1	Cultivated land	1.00
2	Pasture/land used for grazing	1.00
5	Ice	1.33
6	Tundra	1.00
7	Wooded tundra	1.00
8	Boreal forest	1.17
9	Cool conifer forest	1.00
10	Temperate mixed forest	1.17
11	Temperate deciduous forest	1.33
12	Warm mixed forest	1.17
13	Grassland/Steppe	1.00
14	Hot desert	1.00
15	Scrubland	1.00
16	Savanna	1.00
17	Tropical woodland	1.33
18	Tropical forest	1.67

19	No data over land (e.g. Antarctica)	N/A

Source: Own calculations based on Uchida, H. and Nelson, A. (2009)

Lastly, walking travel speed (slope variable) was multiplied by the delay factor (land cover variable) to obtain the off-path friction-surface raster that models the time that it takes to walk 92.5 meters anywhere in DRC circa 1900.

Figure A3: Final natural-historical path raster (hours per pixel)

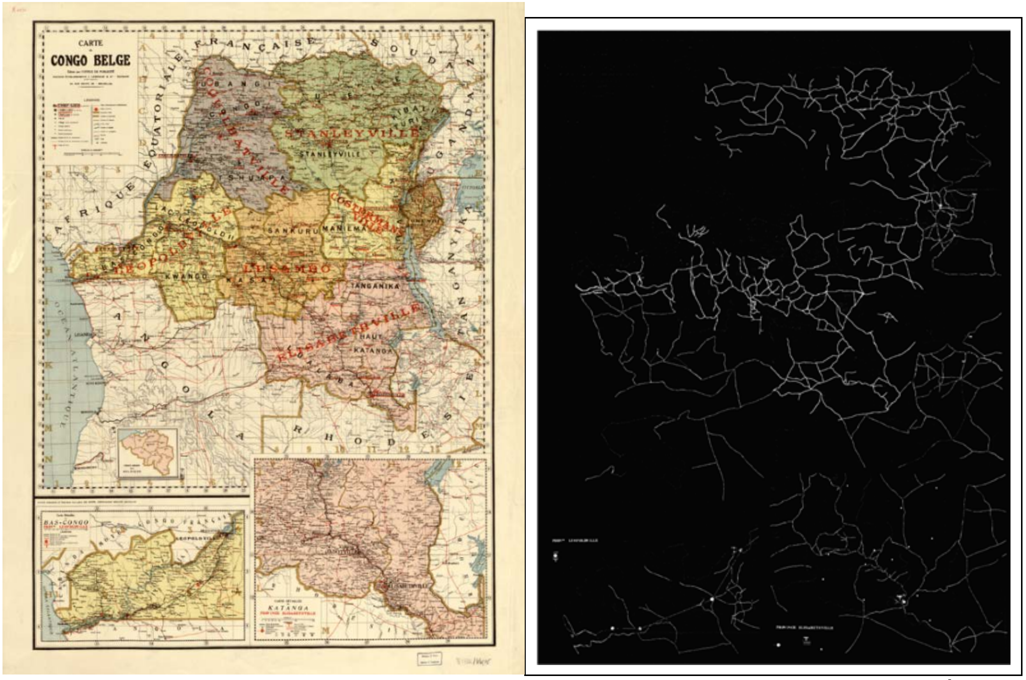


Source: Authors' calculations

Historical Path creation

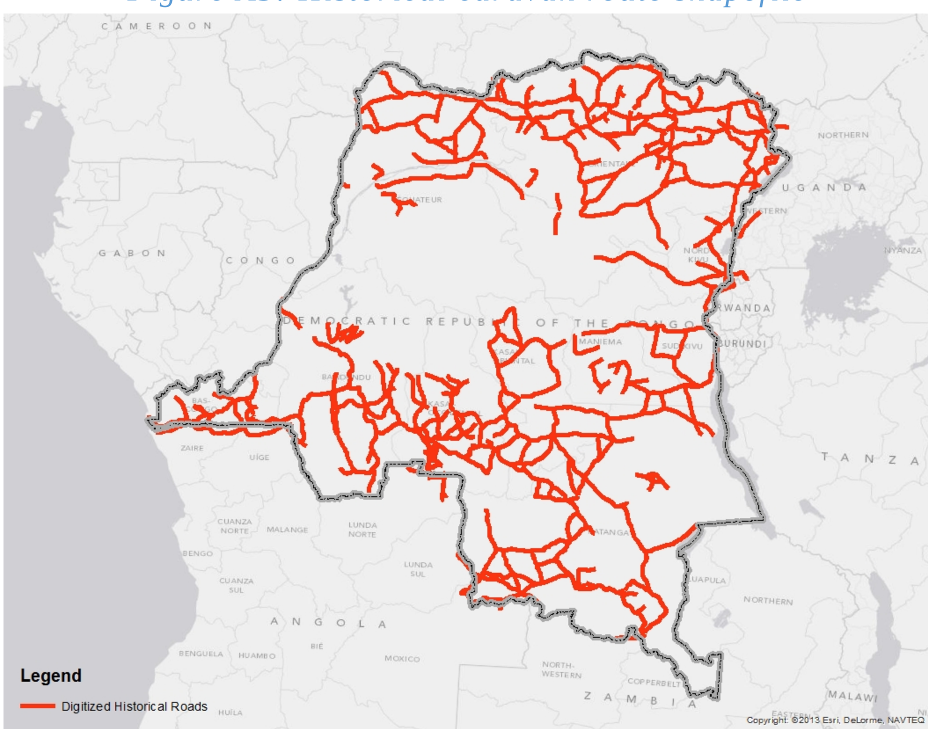
The second step was to digitalize a historical map of DRC to create shapefile that could be added as a layer for spatial analysis. A mix of image manipulator (open source GIMP - http://www.gimp.org/) and GIS software (ESRI's ArcGis Desktop) were used to separate the routes from other map features and then digitize the map. Figure A4 shows the historical map which was then converted into the shapefile shown in Figure A5.

Figure A4: DRC historic trade routes circa 1896 and digitally manipulated map with routes



Source: Carte du Congo Belge / éditée par l'Office de Publicité, anciens établissements J. Lebègue & Cie. - Éditeurs, Bruxelles (1896)

Figure A5: Historical caravan route shapefile



Source: Authors' calculations

Then the newly created shapefile was converted into a raster with a resolution of 92.5 meters to match that of the natural path friction-surface raster. The pixel value assigned to every cell where there is a caravan route passing through is approximately 0.02 hours or 1.2 minutes. This was arrived at by assuming the caravan route travel speed at 5km/hr[36]; equivalent to the human average walking speed on a stable terrain.

Cost-Distance function: Calculating travel time

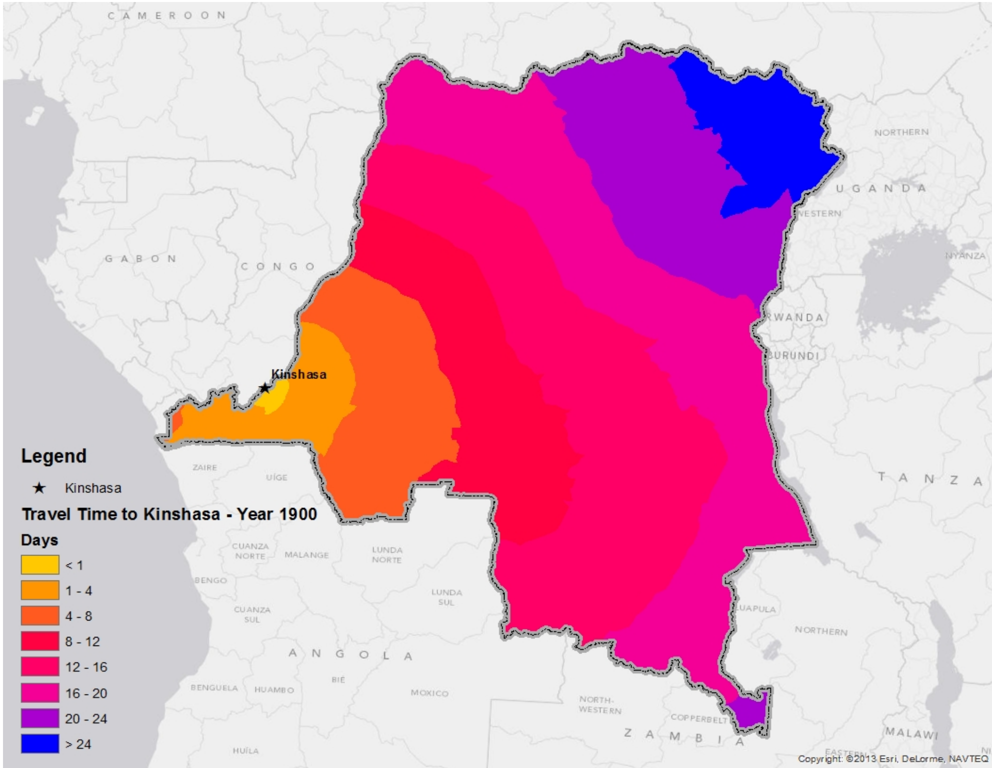
The third step was to merge the off-path and the caravan route friction surfaces. ArcGIS Desktop's tool MERGE was used to combine the two rasters into a single one where the order of the input defines the order of precedence, in this case the caravan routes overlay the off-path walking. Then the friction surface was obtained to model the time that it takes to move around the entirety of DRC around the year 1900, taking into account terrain, land cover type, and transport infrastructure.

To create the final variable, which was used as an instrumental variable in this study, the time that it takes to travel on foot from each pixel in the study area to different selected cities or target destinations, was estimated. ArcGIS Cost distance tools were used to calculate, for each pixel, the least cumulative amount of time it takes to walk to a specified locations (market). The algorithm utilizes the node/link cell representation, whereby the center of each cell is considered a node and each node is connected to its adjacent nodes by multiple links. Every link has an impedance derived from the costs (measured in units of time) associated with the cells from the natural path friction cost surface and from the direction of movement through the cells. See Figure A6 for an example of a raster measuring travel time from each point to Kinshasa.

The creation of a least cumulative cost raster was replicated for each of DRC's 57 selected cities and then then cell values from the travel time raster at SPAM locations was retrieved.. As a result an origin/destination travel time matrix of 27,500 rows (number of pixels in DRC) and 57 columns (selected cities) is obtained. Finally, this data set was compared with the current travel cost data set, and the appropriate city for each pixel was selected for the econometric model.

[36] http://www.princeton.edu/~achaney/tmve/wiki100k/docs/Walking.html

Figure A6: Travel time to Kinhsasa in 1900



Source: Authors' calculations

Appendix IV: Conflict Kernel

In recent years, researchers studying conflict have increasingly embraced a "micro-level" approach as an alternative to macro, country–year level analyses. The micro-level approach suggests that researchers should focus on subnational or individual conflict units of analysis, which are better suited to studying the internal dynamics across time and space and to be able to draw better inferences about the local conditions which affect, and are affected by, conflict. Within the drive for empirical disaggregation, there are two trends. The usual approach of using household or individual surveys and the newer one of geographical and temporal disaggregation of conflict events. For this report the latter approach was selected. Specifically the Armed Conflict Location Events Dataset (ACLED) was used.

The ACLED project codes reported information on the location, date, and other characteristics of politically violent events in unstable and warring states. ACLED focuses specifically on: tracking rebel, militia and government activity over time and space; locating rebel group bases, headquarters, strongholds and presence; distinguishing between territorial transfers of military control from governments to rebel groups and vice versa; recording violent acts between militias; collecting information on rioting and protesting; and non-violent events that are crucial to the dynamics of political violence (e.g. rallies, recruitment drives, peace talks, high-level arrests). ACLED Version 4 data cover all countries on the African continent from 1997 to 2013.

However, due to the nature of conflict, the ACLED data set could not be used in its raw format since it brings up some technical and methodological problems. First, using points pins the conflict to a specific location and does not capture the effects of conflict to the surrounding area (for instance battles that may have been fought in a large area). Second, conflict points cannot capture conflict intensity very well (for instance one isolated conflict point versus a cluster of conflict points). Finally, ACLED is subject to some geographic imprecision due to how the data was obtained (for instance rural conflicts are often allocated to nearby villages). [37]

To account for these methodological issues, a 'hot spot' strategy was followed. Hot spots are concentrations of incidents within a limited geographical area that appear over time (Braga and Weisburd 2010)[38]. Particularly a kernel density interpolation technique was chosen since it allows to transform conflict points into a smooth surface, and generalize conflict locations. To calculate the value at any

[37] In ACLED geographic uncertainty level is coded with "geoprecision codes" ranging from one to three (higher numbers indicate broader geographic spans and thus greater uncertainty about where the event occurred). A geoprecision code of 1 indicates that the coordinates mark the exact location that the event took place. When a specific location is not provided, ACLED selects the provincial capital. This way ACLED may attribute violent incidents to towns when in fact they took place in rural areas and therefore introducing a systematic bias towards attributes associated with urban areas that can lead to invalid inferences.

[38] Braga, A. & Weisburd, D. (2010). Policing Problem Places: Crime Hot Spots and Effective Prevention. Oxford: Oxford University Press.

point, the kernel density function takes a weighted average of all the conflicts around that point, to create the surface. The magnitude of the weight declines with distance from the point, according to the chosen kernel function.[39]

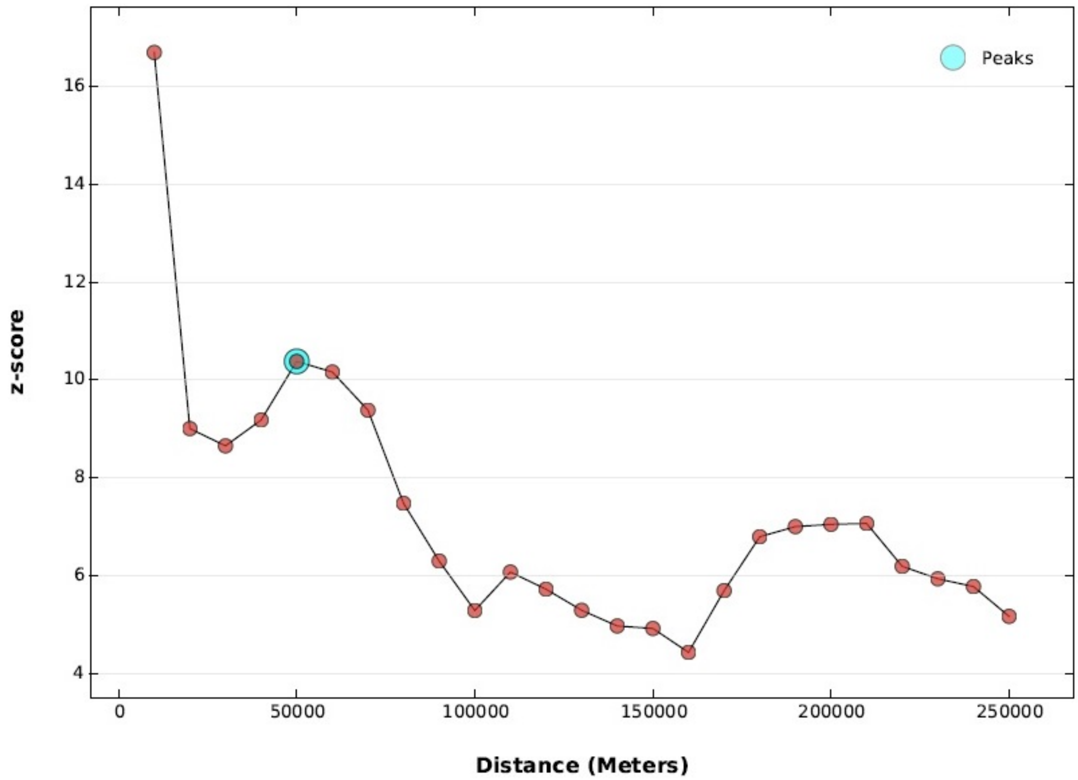
When using kernel density estimation, two decisions must be made: what kernel function to use, and what bandwidth to search over. The literature using this technique in the context of conflict is not very large, and therefore does not offer much guidance on these decisions. While, other researchers have used similar techniques to estimate crime densities (Levine 2006, Chainey et al 2008, Eck et al 2005, among others), there does not appear to be agreement on what kernel function is most appropriate. Without an obvious candidate, a quadratic kernel function, as described in Silverman (1986)[40], was employed. This is a fairly common kernel function. Nevertheless, when other kernel functions are employed in its stead (i.e. Gaussian, Epanechnikov, Quartic), there is not a large effect on the kernel surface.

The other key parameter in the kernel function is the bandwidth. Without any robust theory telling us how far the effects of conflict can permeate, this study relies on spatial autocorrelation. Using the Moran correlogram (Levine 2013) the degree of spatial autocorrelation as a function of distance was calculated. By measuring spatial autocorrelation for a series of distances and their corresponding z-scores a sense of the intensity of spatial clustering can be obtained. Statistically significant peak z-scores will show distances where spatial processes promoting clustering are most pronounced. Figure A7 below shows that the distance where autocorrelation is most pronounced is approximately 50kms, and this is the bandwidth employed in the kernel function.

[39] For instance, if a conflict occurs exactly on the point which is being calculated, the value of that conflict will receive a weight of 1. A conflict which is 5 kms away from the point will receive a weight of α and a conflict 10 kms away will receive a weight of β, where $1 > \alpha > \beta > 0$. Eventually, at some distance, referred to as the bandwidth, the weight becomes zero. For more information on how the kernel function and bandwidth were chosen, see Appendix IV

[40] (Silverman, B. W. Density Estimation for Statistics and Data Analysis. New York: Chapman and Hall, 1986. p. 76, equation 4.5) http://books.google.com/books?id=e-xsrjsL7WkC&lpg=PP1&pg=PA76#v=onepage&q&f=false.

Figure A7: Moran's I, Spatial Autocorrelation by Distance



Source: Authors' calculations

The last ingredient to run the analysis is the conflict data itself. As explained above, the data not only includes where and when the events happened, but also its type (violent, battles, riots, etc.) and outcome (number of fatalities), among other variables. Therefore, several measures of conflict are estimated. The first is the "kernelly" estimated number of fatalities in the 5 years preceding our DHS data set (2003-2007) and local GDP data set (2002-2006). This variable was calculated around each household and also around each market. Also, a dummy variable which indicates if there are relatively high levels of conflict near households and markets was generated; it takes a value of 1 if the kernelly estimated fatalities are greater than the median number of fatalities due to violent conflict near both households around the nearest market, and zero otherwise.

To illustrate how the conflict data was modified to facilitate analysis, Figure 1 above shows the original ACLED conflict data (left), as well as the kernel density map (right).

Appendix V: Proof of results from theoretical model

Heuristic proof for existence of the equilibrium

Without loss of generality assume that $K_f = K_g = K$. Observe that as $K \to 0$ then $U_R(n) = \frac{n^2}{4(1-n)}\left(\left\{\frac{[cL(1-\beta)]^2}{K}\right\} + \left\{\frac{[mL\beta)]^2}{K}\right\}\right) \to \infty$ and $U_F(n_i) = \left[\frac{\beta(1-g)(P-t_i)\beta mL}{(P_B+t_i)}\right]^{\beta}[c\alpha L(1-f)]^{\alpha} \to 0$ (since $g \to 1$ and $f \to 1$). Hence as $K \to 0$ then $\Psi \equiv U_F(n_i) - U_R(n_i) \to -\infty$.

Similarly, as $K \to \infty$ then $U_R(n) \to 0$ and $U_F(n_i) \to \left[\frac{\beta(P-t_i)\beta mL}{(P_B+t_i)}\right]^{\beta}[c\alpha L]^{\alpha} > 0$ hence $\Psi > 0$. Since Ψ is continuous in K it follows that there exists some point such that $\Psi \equiv U_F(n_i) - U_R(n_i) = 0$. ■

Result 1

Consider a change in transport costs z for the marginal household. Note that $\frac{d\Psi}{dz} = \frac{\partial\Psi}{\partial z} + \frac{\partial\Psi}{\partial n}\frac{\partial n}{\partial z}$ and

$\frac{\partial U^F}{\partial t} = \frac{\partial t}{\partial z}\left(\frac{nA^{\alpha}(1-g)mL\beta B^{\beta-1}(-\beta-(p-t)}{p_B+t}\right) < 0$; where A = $c\alpha L(1-f)$ and B = $\frac{\beta(1-g)(P-t_i)\beta mL}{(P_B+t_i)}$. Similarly

$\frac{\partial U^{\ R}}{\partial t} = \frac{2-n}{4(1-n)^2}\left(\frac{\left(cL(1-\beta)\right)^2}{K_f} + \frac{(mL\beta)^2}{K_g}\right)\frac{\partial n}{\partial t}\frac{\partial t}{\partial z} < 0$ (since $\frac{\partial n}{\partial z}\frac{\partial t}{\partial z} < 0$ and $1-n > 0$ and $1-\beta > 0$ by assumption.

Further note from equation (2.8) as $K_g \to 0$ then $g \to 1$, hence $\frac{\partial U^F}{\partial t} \to 0$; $\frac{\partial U_R}{\partial t} \to -\infty$ and $\frac{\partial\Psi}{\partial t} \to \infty$.[41] ■

Result 2

$\frac{\partial U^F}{\partial f} = -c(1-\beta)B^{\beta}L\alpha A^{\alpha-1}$ and $\frac{\partial U^F}{\partial g} = \frac{-\beta B^{\beta-1}LA^{\alpha}(p-t)}{(p_B+t)}$. Thus $\left|\frac{\partial U^F}{\partial f}\right| - \left|\frac{\partial U^F}{\partial g}\right| > 0$ whenever

$\frac{LA^{\alpha-1}B^{\beta-1}}{p_B+t}(p-t)(1-\beta)cmL\beta\big(\alpha(1-g)-\beta(1-f)\big) > 0$. Which holds whenever $\frac{\alpha}{\beta} > \frac{1-f}{1-g}$.

---

[41] Alternatively this result may be derived as follows: $\frac{d\Psi}{dz} = \frac{\partial(U^F-U^R)}{\partial z} + \frac{\partial(U^F-U^R)}{\partial n}\frac{dn}{dz}$. From the stage 1 equilibrium condition we have that in equilibrium by total differentiation of Ψ we get $\frac{dn}{dz} = \frac{\partial U^R/\partial n}{\partial U^F/\partial t}$. Upon substitution in the preceding equation and rearranging yields $\frac{d\Psi}{dz} = \left(\frac{\partial U^F}{\partial z}\right)^2 - \left(\frac{\partial U^R}{\partial n}\right)^2$. Lim $K_g \to 0$ then $g \to 1$, hence $\frac{\partial U^F}{\partial z} \to 0$; $\frac{\partial U^R}{\partial n} \to -\infty$ and $\frac{d\Psi}{dz} \to \infty$.