Who Benefits from Free Trade?

Mathias Bühler*

23rd December 2023

Abstract

How is wealth distributed when the economy grows? I study this question in the context of African countries and ethnic groups. If wealth is distributed proportional to population, larger ethnic groups should benefit more when economic activity increases. Using nighttime light and individual level data to geographically locate wealth, I find the exact opposite: Smaller ethnic groups, particularly those in political power, benefit more from increased economic activity than larger ones. The results indicate that political elites in power redistribute wealth from larger ethnic groups. As a result, people's satisfaction with democracy and trust in institutions reduces, casting a shadow on the implementation of trade liberalization policies in developing countries. Instrumental variables estimating exploiting exogenous variation in trading activity confirm initial results.

Keywords: Elite capture, development policy, wealth distribution

JEL codes: O24, N77, F14

*Ludwig-Maximilians-University, Ludwigstrasse 33, 80539 Munich, Germany. email: Mathias.Buehler@econ.lmu.de. I benefited greatly from discussions with Alessandra Allocca, Dorothee Bühler, Davide Cantoni, Andrew Dickens, Andreas Madestam, and Martina Magli. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

1 Introduction

People engage in economic activity to exchange goods and increase their wealth. It is then a foundation of economics that more exchanged goods in the aggregate also leads to more wealth. But who reaps this wealth? Those who produce and exchange goods, or their ruling elites? In this paper I ask who benefits from the 'gains from trade' and how it affects social cohesion and democratic development.

Africa provides the ideal setting to study these questions, because the arbitrary placement of country borders during colonization split some ethnic groups into multiple parts, but not others. This 'scramble for Africa' arguably contributed to the relative economic underperformance of, and ethnic favoritism in, African countries today (Alesina et al., 2016; Michalopoulos and Papaioannou, 2016; Clochard and Hollard, 2018; Dickens, 2018). In addition, trade liberalization policies have been touted as a panacea to boost economic development and are thus part of virtually all major multilateral agreements signed by African countries today (Smeets, 2021; Lejarraga, 2022).

In this paper, I analyze how wealth is distributed by assigning increased economic activity to ethnic groups by their population shares: an equal distribution would imply that larger ethnic groups benefit more than smaller groups when the economy grows. I use data on bilateral trade between African countries and the distribution of ethnic groups prior to colonialization (Murdock, 1959) to assess wealth gains measured by nighttime light data and seven georeferenced surveys from the Afrobarometer project. Controlling for an extensive range of fixed effects, I isolate the impact of increased trade exposure by interacting trade activity with the population shares of each group.

The analysis reveals three insights. First, gains in wealth are not equally shared: Nighttime light data and individual survey data reveal a negative relationship between trade exposure and wealth across the entire African continent. Second, wealth gains accumulate in ethnic groups in political power, providing evidence for elite capture. Third, this undermines the democratic process: Elite capture of wealth has negative consequences for individuals' satisfaction with democracy and trust in institutions. This paper thus yields new insights on the distribution of wealth gains and casts a shadow on trade policies' impact on development.

The analysis unfolds in two parts. First, how are wealth gains distributed, and second, how does this affect social stability. There are two issues related to reversed causality and omitted variable bias that have to be addressed throughout the analysis. First, trade increases wealth, but richer countries also trade more. Second, government policies or infrastructure might influence trade and are often correlated with the ethnic group in power (Burgess et al., 2015). I thus utilize insights from the trade literature's study on the effects of china's accession to the WTO to obtain quasi-exogenous variation. Similar to Autor et al. (2013) I capture each groups' exposure to trade flows by aggregating realized bilateral exports to the country-of-origin level (*shift*) and interact this trade activity with each group's pre-colonial population (*share*). Country-by-year and country-by-ethnicity fixed effects then address these concerns by capturing the average effect of increased trade activity and ethnic status in each country, respectively, isolating how increased wealth is distributed among ethnic groups: If wealth gains are distributed proportional to size, this interaction term will equal to zero.

Both nighttime light data (2012-2020) at the ethnic-group level and georeferenced data at the individual level from the Afrobarometer (1999-2018) reveal a significant negative relationship between trade exposure and wealth gains. A 10% increase in economic activity reduces nighttime luminosity by 37% and individual wealth by 10%. Groups in power, however, benefit from increased economic activity. Using data on the political status of ethnic groups, I provide evidence that elite capture distorts and redirects wealth, hurting economic development in Africa .

The second part of the analysis concerns the impact on social stability. When wealth gains are not equally shared, it might lead to lower trust, social stability, and a deterioration of democratic institutions. Using the Afrobarometer's questions on satisfaction with democracy and trust in institutions, I highlight a significant negative impact on social stability. Elite capture of the gains from trade thus shape the the distribution of wealth in Africa today. Relocating factories and economic activity into their own ethnic homelands likely explains the negative impacts on trust institutions and satisfaction with democracy: Being left behind by the elites that govern the country, people lose trust in the democratic process. This paper is thus among the first to causally show that including free trade policies in multilateral agreements might thus add to the growing dissatisfaction with democracy in many developing countries.

I argue that the interaction between trade activity and population share, conditional on country-by-year and country-by-ethnicity fixed effects, identifies how wealth is distributed among ethnic groups in Africa. However, endogeneity concerns regarding the political status of ethnic groups as well as its interaction with trade activity may remain. If the political status of ethnic groups is determined by an endogenous component capturing each group's propensity to rule and a random component determining whether the group actually rules, country-by-ethnicity fixed effects perfectly capture the endogenous component, isolating the variation of the random shock to its political status.

However, if ethnic groups in political power actively pursue policies to increase trade activity because they believe to benefit from it, the interaction of trade activity and ethnic group may still be biased. Then, in order to obtain a causal estimate, either shifts, i.e. trade activity, or shares, i.e. ethnic groups, need to be exogenous (Borusyak et al., 2022). Thus, I propose two entirely different instruments to obtain exogenous variation that shifts trade activity.

The first strategy exploits cross-sectional variation in cross-border networks. I utilize the pre-colonial distribution of ethnic groups in all continental African countries and exploit that colonial powers set country borders irrespective of the underlying ethnic homelands. This creates ethnic networks across country borders that are essentially random, as country borders were drawn in 1884 without taking into consideration that countries could become independent more than 60 years later (Michalopoulos and Papaioannou, 2016). Leveraging the strength of this initial ethnic network across neighboring countries, I obtain a valid instrument that exogenously shifts bilateral exports.

The second strategy is akin to a leave-one-out estimator and is identified from panel variation in economic activity. Each country c's realized trade flows are replaced by the average trade flows from all non-bordering countries to all other non-bordering African countries. This ensures that no characteristic of country c is directly used to predict its trade activity and that all variation comes from the average increase in economic activity of Africa.

Similar to Frankel and Romer (1999), I then aggregate predicted exports between neighboring countries ($\widehat{shift}$) and interact this predicted trade activity with population shares (*share*) to instrument realized trade exposure in the *shift-share* estimation on the distribution of wealth gains. With F-statistics on the first stage of 79 for the cross-border instrument and 13 for the leave-one-out instrument, the resulting point estimates confirm initial results.

These findings contribute to our understanding of Africa's long-run development and the important role its colonial history plays. In related work, Michalopoulos and Papaioannou (2016) show that ethnic groups split across country borders are poorer and lag behind non-split ethnic groups. Split ethnic groups were also less politically centralized in the pre-colonial period, which further emphasizes the fact that they exhibit lower levels of economic development today (Michalopoulos and Papaioannou, 2013). My findings suggest that split ethnic groups with large ethnic networks across borders benefit from increased trade activity, yet these gains disproportionately accumulate with the ethnic groups that hold political power. This is suggestive of a mechanism that aligns with the insights of Dickens (2018), who documents evidence of ethnic favoritism within split groups throughout sub-Saharan Africa. More broadly, evidence of ethnic favoritism in

African politics is well documented in the literature (Frank and Rainer, 2012; Burgess et al., 2015; Kramon and Posner, 2016). Overall, my results highlight a novel channel through which patterns of development have persisted throughout the African continent.

I also contribute to the emerging discussion on the distributional effects of trade activity. While it is clear that liberalizing trade generates winners and losers, identifying them empirically was near impossible. In this paper, I show how to identify winners and losers from aggregate data in developing countries, where firm-level data is non-existent or unreliable, yet identifying them is of paramount importance for social stability. Engel et al. (2021) provides an overview of the distributional effects of trade across regions and demographic groups over time. At the firm level, Baccini et al. (2017) highlight how preferential trade agreements increase trade disproportionately for large firms. This evidence is corroborated in the developing countries setting, where Dhingra and Tenreyro (2020) evaluate agribusinesses providing access to farmers and show that while businesses gained, farmers in villages that produced policy-affected crops saw reductions in consumption. Using the staggered implementation of the Africa Growth and Opportunity Act, Desmet and Gomes (2023) show that trade access increases income in general, but decreases it for remote ethnic groups. In contrast to existing studies focusing on tariff reductions, I provide evidence how trade flows differential affect groups based on their power status within government. Thus, my findings add to academic and policy debates on the distributional impacts of trade policies.

The findings in this paper also relate to the discussion on whether trade causes growth (Frankel and Romer, 1999). This literature has used gravity equations to study this relationship, exploiting airplanes (Feyrer, 2019), the Suez canal closure (Feyrer, 2021), or the evolution of the steam ship in the 19th century (Pascali, 2017). Similar to papers that broaden the scope of this question to intra-national trade costs (Donaldson and Hornbeck, 2016) or information frictions (Steinwender, 2018), I add a political economy dimension to this question that hitherto has not been studied in the literature.

This paper is structured as follows. Section 2 presents the data and variable definitions used throughout the paper. Section 3 presents the empirical strategy. Section 4 estimates the effect of trade exposure and elite capture on economic and societal development in Africa. Section 5 presents robustness using two instrumental variables strategies. Section 7 concludes.

2 Data

Economic activity Data on bilateral trade are obtained from UN Comtrade World Bank Integrated Trade Systems from 1990–2020. I use import and export data to maximize coverage of reported trade, acknowledging that the point estimates are likely lower bounds on the true effect of exports between countries.[1] Exports for every country c to every destination d on the African continent are aggregated to the country-by-year level $Export_{c,t} = \sum_{d \in D} Export_{c,d,t}$ and used as *shifters* to the economic activity.

Ethnic population The ethnic group of each individual or region is derived from the spatial intersection of the map in Murdock (1959) with modern country borders.[2] In total Africa contains 833 ethnic groups in 48 African countries. The population share of ethnicity e in country c is then calculated by aggregating detailed grid-cell population data from the United Nations Environment Program in 1960 to the ethnicity-by-country level. These population shares then act as *shares* to assign economic activities to individuals and groups.

Nighttime lights Data on economic development is derived from the most recent satellite data on nighttime lights (Visible Infrared Imaging Radiometer Suite, VIIRS) at a resolution of 500m at the equator (Elvidge et al., 2021). This data is an improvement over the older DMSP-OLS Nighttime Light Series Elvidge et al. (1997) and the new standard in the literature.[3] Two variables are constructed to measure economic development: $Lid_{e,c,t}$ calculates the fraction of pixels with a luminosity greater than zero for each country-ethnic group observation. $\log(NTL+1)_{e,c,t}$ calculates the logarithm of average luminosity for each country-ethnic group observation.

Household wealth Information on household wealth is derived from the georeferenced version of the Afrobarometer survey rounds 1–7 (BenYishay et al., 2017). The ethnic group of each individual is determined by the spatial intersection of Murdock (1959) with the individual's location. I create three standardized indexes from a list of questions capturing household wealth, satisfaction democracy, and trust in institutions. Appendix C explains

[1] If the data is split up into reported or unreported trade, the true estimate will be $\beta = \left(\beta^{reported} X_{cd}^{reported} + \beta^{unreported} X_{cd}^{unreported}\right) / (X_{cd}^{reported} + X_{cd}^{unreported})$. As long as $\beta^{reported} \leq \beta^{unreported}$, I estimate a lower bound effect.

[2] The results are robust to using modern day ethnic distributions and ethnolinguistic distribution of ethnic groups in Weidmann et al. (2010).

[3] Results using the older series confirm the main result: Political Elites capture 20-100% more of the wealth than non-connected groups.

the methodology, shows p-values adjusted for multiple hypothesis testing, and lists all used questions from the latest round.

Ethnic power relations The political status of every country-ethnic group observation is derived from Wimmer et al. (2009). The georeferenced data is interesected with the country-ethnic group from Murdock (1959) and the spatial location of the individual. In case an exact match cannot be found, I follow a two step procedure: First, I use the closest ethnic group within 250km in the same country before linking the remaining ethnic groups based on their names and country.[4] $InPower_{e,c,t}$ denotes whether the individual or group belongs to an ethnicity e enjoying a monopoly or dominant status in country c and year t.

Conflict I obtain georeferenced conflict data from https://ucdp.uu.se/. $Conflict_{e,c,t}$ is defined as any conflict occurring in country c, ethnicity e, and year t. Results are robust to using number of conflicts, various definitions of deaths, or conflict intensity.

3 Empirical Strategy

I study how wealth is distributed geographically, using nighttime light and individual level data in Africa. The unit of observation is a country-ethnic group that is derived from the intersection of 833 precolonial homelands of ethnic groups (Murdock, 1959) with 53 modern-day country borders. In total, there are 1,383 country-ethnic group observations in each year. Nearly half of all ethnic groups in Africa are split between two or more countries.

The estimation equation is derived from a simple principle: Increased economic activity, proxied by trade activity, should increase wealth of ethnic groups.

$$Y_{e,c,t} = \beta_1 \log\left(\sum_{d \in D} Export_{c,d,t}\right) + \beta_2 Population\ Share_{e,c} + \alpha_c + \alpha_t + \varepsilon_{e,c,t} \tag{1}$$

$Y_{e,c,t}$ captures wealth as either satellite data capturing luminosity or individual wealth from the Afrobarometer surveys. I expect $\beta_1 > 0$ as trade should increase wealth. This coefficient can be interpreted as how much increased trade activity affects wealth *on average*. The second variable $Population\ Share_{e,c}$ captures the share of an ethnic group e in country c. The sign of β_2 is, however, ambiguous. If larger ethnic groups are more developed and capture *on average* a larger share of the economic activity, we would expect

[4] I use record linking and compare the string differences between ethnic group's names. I only use perfect matches. Results are robust to only using spatial matches.

$\beta_2 > 0$. Alternatively, if smaller ethnic groups are located close to the capital and occupy influential positions, we would expect $\beta_2 < 0$.

Equation (1) does, however, also capture several factors that compound and bias the treatment effect. First, the sum of exports is correlated with GDP and population, likely biasing the estimate on β_1, motivating the inclusion of economic fixed effects $\alpha_{c,t}$ to hold GDP, population, political system, and aggregate trade flows of country c in each time period t constant. Then, however, β_1 is not identifiable using equation (1). Second, ethnic groups size is likely correlated with economic development, but also to their political status, the fertility of their ethnic homelands, or historical political development (Michalopoulos and Papaioannou, 2013). I thus include country-by-ethnicity fixed effects ($\alpha_{c,e}$) to hold observable and unobservable characteristics for ethnicity e in country c, including its population share, homeland size, average economic and political status, as well as conflict prevalence, constant. Then, again, β_2 is not identifiable using equation (1).

Variation in trade exposure Thus, to estimate ethnic-group level exposure to aggregate trade flows and how wealth is distributed, I estimate the following equation interacting aggregate bilateral exports with population shares for each group:

$$Y_{e,c,t} = \gamma \underbrace{\log\left(\sum_{d\in D} Export_{c,d,t}\right) \times PopulationShare_{e,c}}_{TradeExposure_{e,c,t}} + \alpha_{c,t} + \alpha_{c,e} + \varepsilon_{e,c,t} \tag{2}$$

Conditional on a large set of fixed effects, *Trade Exposure*$_{e,c,t}$ is identified from the interaction of aggregate bilateral exports from country c to all destinations $d \in D$ with the population share of ethnicity e in country c. The comparison is thus strictly within each country-year observation, comparing ethnic-groups to their long-term average. In this setup, aggregate bilateral exports act as a *shifter* that is assigned to each ethnicity by its *population share*. Standard errors are clustered at the country-ethnic group level.

Contrary to equation (1), where more trade implies larger economic development, the sign of γ is unclear as it captures how trade activity differentially affects ethnic groups. Consider a stylized example of country c having two ethnic groups $e_1 = 30\%$ and $e_2 = 70\%$. If the gains from trading were proportionally shared among all ethnic groups, we would expect that γ is zero as the average (level-) effect is captured by β_1 inside the country-by-year fixed effects. A ten unit increase in wealth (ΔY_e) is then proportionately shared among all ethnic groups $\Delta Y_1 = 3;\ \Delta Y_2 = 7$.

A positive coefficient would suggest that large ethnic groups capture a disproportionate share of the benefits, redistributing from small ethnic groups to larger ones. Instead of their proportional share from the ten unit increase in wealth, a share γ is redistributed from e_1 to e_2: $\Delta Y_1 = 3 - \gamma$; $\Delta Y_2 = 7 + \gamma$. A negative coefficient suggests the opposite and is indicative of elite capture: The smaller ethnic group e_1 captures a share γ from e_2 and receives a disproportionate amount of the wealth gains $\Delta Y_1 = 3 + \gamma$; .[5]

Appendix A.1 explains in great detail how equation (2) can be derived and discusses why the inclusion of $\alpha_{c,e}$ and $\alpha_{c,t}$ capture unobserved selection biases in the original equation (1). As *Trade Exposure*$_{e,c,t}$ is now an unbiased estimate of how ethnic groups benefit from economic output, I continue and test the effect of political capture.

Variation in political status To verify the presence of elite capture, I utilize exogenous variation in each group's political status. Each group's political status depends on both endogenous and random factors: As the endogenous factors can be approximated by each group's historical judicial development, propensity to rule, or economic development today (Michalopoulos and Papaioannou, 2013, 2014), country-by-ethnicity fixed effects isolate random variation in power status. Then, additionally controlling for country-by-year fixed effects isolates the random variation that determines each group's access to power.

I thus interact equation (2) with data on political relations and estimate:

$$
\begin{aligned}
Y_{e,c,t} = {} & \gamma\, TradeExposure_{e,c,t} + \delta\, TradeExposure_{e,c,t} \times InPower_{e,c,t} \\
& + \alpha_{c,e} + \alpha_{c,t} \times InPower_{e,c,t} + \varepsilon_{e,c,t}
\end{aligned}
$$

In this setup, γ captures the wealth gains of ethnic groups from additional exposure to trade if they are not in political control. Groups that are in political power then gain an additional δ from additional trade exposure. $\alpha_{c,t} \times InPower_{e,c,t}$ controls for the average political status of country c in time period t, such that δ measures the extent of elite capture.

4 Results

How are the wealth gains from increased economic activity distributed in heterogeneous societies? Who captures the gains from trade and how does this affect societal development? I answer these questions using nighttime light satellite imagery (Table 1) and

[5]This interaction does not capture a simple urban/rural divide in which rural areas are larger with lower population density and thus fewer nighttime lights. Size, location, population density is held constant by $\alpha_{c,e}$.

individual survey data in African countries (Table 2).

I begin by using nighttime light satellite imagery as a proxy for wealth in Table 1. Columns (1) and (2) present results on the fraction of pixels lid for each country-ethnic group. The results suggest that a 10% increase in trade results in a 37% decrease in the fraction of pixels lid.[6] If the group is in power, however, the interaction term in column (2) suggest a 15% increase in nighttime luminosity.[7] These results carry over when considering average luminosity in columns (3) and (4).

Table 1 suggests significant elite capture of the gains from trade. The negative point estimate suggests that most ethnic groups in African countries do not benefit from increased trade activity; gains are squarely located with the group in power, redistributing wealth towards their own group.

These group-level estimates from nighttime lights carry over to individual estimates using the Afrobarometer Surveys. In Table 2 columns (1) and (2), I construct a standardized measure of relative household wealth and use it to assess how wealth is distributed among ethnic groups. The same picture emerges: A 10% increase in exports decreases household wealth by 10% of a standard deviation for groups not in power, and increases wealth by 4.6% of a standard deviation for groups in power.

As the Afrobarometer is mainly a survey about political values, I construct two indices capturing 'satisfaction with democracy' and 'trust in institutions' from questions listed in Appendix C.[8] I show the average effect of trade exposure on these indices in Table 2, columns (3)-(6). Increasing trade activity reduces ethnic groups' satisfaction with democracy and their trust in institutions effectively undermining state building efforts.

Table 1 and 2 thus provide evidence that wealth gains are not shared proportional to each ethnic groups' population share. This result cannot be explained by time-varying country factors or time-invariant characteristics of ethnic groups such as population density, pre-colonial distributions, or the ethnic group being split; country $\times$ year and ethnic $\times$ country fixed effects absorb these confounders completely. These results are also not driven by outliers as dropping countries individually does not alter the estimate significantly (Figure A.1).[9]

[6] A potential concern is differential population density across ethnic groups. Country-by-ethnicity fixed effect capture all constant characteristics, including population density.

[7] Calculated from 10% of the average log exports (13.87) times the point estimate relative to fraction of lid pixels (0.033) on average and for groups in power (0.088).

[8] These questions set the Afrobarometer apart from the DHS that mainly captures health-related questions. Results on wealth using the DHS are replicated in Section D.

[9] Following (Borusyak et al., 2022) I also cluster the standard error at the level that provides exogenous variation; in this case the ethnic group. Standard errors are smaller and thus not reported.

4.1 Mechanism

There are two explanations how powerful elites can benefit from increased economic activity; Bribery and redirection of economic activity. While bribery is certainly widespread among African countries, with many ruling parties benefiting their own group, the focus of this paper lies on the redirection of economic activity. In the spirit of the road building exercise in Burgess et al. (2015), the redirection of economic activity could be the construction or upgrading of roads that attract new businesses, or in the outright relocation of businesses to the ethnic homelands of powerful elites.

Both mechanisms predict that exposure to manufacturing exports has worse impacts on economic development than agriculture or resource exports. Whereas resources and agricultural fields are immutably fixed in space, factories can be relocated. Table A.1 provides suggestive evidence in favor of such relocation as exposure to manufacturing exports significantly reduces nighttime luminosity, and thus wealth.

5 Robustness

Even conditional on country-by-year fixed effects capturing economic activity and country-by-ethnicity effects capturing ethnic relations, these findings could be biased. Larger ethnic groups are more likely to be split into multiple countries, are less likely to gain power, and are less likely to be economically integrated (Michalopoulos and Papaioannou, 2013, 2016). Conversely, smaller ethnic groups are more likely to be closer to the centers of economic and political power. Then, their ability to relocate economic activity towards their homelands might create a reversed causality bias in the interaction term in trade exposure. Do groups benefit from increased trade exposure, or is trade exposure increased because they benefit from it?

In order to assess the severity of this bias and validate my findings I propose two instrumental variables strategies: The first instruments shifts the size of cross-border ethnic networks to predict increased trade volumes between neighbors. This *Cross-Border* instrument exploits cross-sectional variation in the precolonial distribution of ethnic groups between neighboring countries. The second instrument uses non-neighboring countries' exports to other African countries to predict a country's exports. This *Leave-One-Out* instrument exploits panel variation in the average trade activity of non-neighboring countries. Thus, both instruments exploit different sources of variations to predict shifts in trade activity.

5.1 Cross-border instrument

I begin by developing a gravity-type equation that incorporates heterogeneous ethnic groups across multiple country pairs. Then, I exploit the quasi-exogenous placement of borders to obtain exogenous variation in pre-colonial population shares in each exporting country. I then use each ethnic groups' connections to the importing country to exogenously shift trade activity.

A stylized model of trade In the trade literature, the value of bilateral exports is modeled in gravity-type equations (Anderson, 1979). Here, the value of trade is correlated with the size of the exporter and importer economy and the geographic distance between them, as larger and more geographically close economies trade more. In this framework, the addition of a population share of people from country of origin c in destination country d ($PS(c)_{d,t}$) identifies the strength of cross-country networks:

$$\log(X_{cd,t}) = \beta \log(PS(c)_{d,t}) + \Gamma_{cd,t} + \alpha_{c,t} + \alpha_{d,t} + \varepsilon_{c,d,t} \tag{3}$$

Controlling for country ($\alpha_{c,t}$) and destination ($\alpha_{d,t}$) fixed effects interacted with time period fixed effects and bilateral characteristics ($\Gamma_{cd,t}$), β identifies the effect of the population share $\log(PS(c)_{d,t})$ on the log of exports $\log(X_{cd,t})$. The elasticity $\beta > 0$ indicates that trade activity increase if the trading partners share a larger network.

Implicitly, equation (3) assumes that migrants to destination d identify with the nationality of their country of origin c.[10] African countries however, combine a multitude of ethnic groups, each with their own identity. Allowing for multiple ethnic groups (e) from the set of ethnic groups in each country ($e \in E_c \cap E_d$), the general form of equation (3) is given by:

$$\log(X_{cd,t}) = \underbrace{\beta \log\left(\sum_{e \in E_c \cap E_d} PS_{c,t,e} \times PS_{d,t,e}\right)}_{Ethnic\ Connections_{cd}} + \Gamma_{cd,t} + \alpha_{c,t} + \alpha_{d,t} + \varepsilon_{c,d,t} \tag{4}$$

where $PS_{c,t,e} \in (0,1]$ is the population share of an ethnicity e that is common to country c and d, relative to the population of country c at time t. This formulation nests equation (3) if country c has only one ethnic group with $PS_{c,t,e} = 1$. Equation (5) correlates bilateral

[10]The underlying equation is of the form $PS(c)^{\beta}_{d,t} = (Pop(c)_{d,t}/Pop_{d,t})^{\beta}$. The population of migrants from country c in destination d at time t ($Pop(c)_{d,t}$) is denominated by the population size of destination d at time t ($Pop_{d,t}$). The implicit assumption is that all migrants form c identify with country c, and not with a subgroup e. That is, $(Pop(c)_{c,t}/Pop_{c,t})^{\beta} \approx 1$. Combining these yields $PS^{\beta}_{d,t} = (Pop(c)_{d,t}/Pop_{d,t} \times Pop(c)_{c,t}/Pop_{c,t})^{\beta}$.

exports to the probability of a co-ethnic relationship (match) when randomly drawing two individuals from each country. It captures the idea that it is easier to trade with someone from your own ethnicity, but does not exclude the possibility of trading with other ethnic groups.

The formulation of equation (5) is supported by three observations. First, it is the empirical equivalent of an otherwise standard model of international trade (Melitz, 2003; Chaney, 2008) that adds an ethnicity-specific fixed cost capturing lower entry costs into an export market for ethnically connected firms.[11]

Second, the interpretation is equivalent to the search and matching literature where a match is defined when two individuals of the same characteristics are drawn. Since these characteristics are stochastic, the likelihood of a match is given in probabilities. Here, characteristics are distributed along ethnic lines and thus the fraction of the population representing an ethnicity in the importing country is equivalent to the likelihood that an exporting firm from the exporting country finds a match in the importing country. Then, the estimated β can be interpreted as an elasticity that captures the change in match probability of each ethnicity when its population changes on either side of the border.[12] This interpretation is similar to the standard in equation (3); both can be interpreted as a probability of drawing two connected people in each country. Equation (5), however, incorporates the heterogeneous population structures in African countries and allows for a large amount of subgroups within two countries that are connected.

Third, an alternative interpretation of the coefficient β is akin to iceberg trade costs: Ethnic connections capture the 'ethnic distance' between two countries. The ethnic composition of a country can be reflected by a vector $\overrightarrow{e_c}$ that contains the population shares of all possible ethnic groups $e \in E$. The product $\overrightarrow{e_c} \times \overrightarrow{e_d}$ then results in a linear distance measure between countries c and d in terms of ethnicity. Then, similar to the interpretation of larger geographic distances between countries reducing trade, larger 'ethnic distances' also reduce trade by capturing increasing dissimilarity between countries.

[11] These costs can be lower information costs, more reliable information about market structures or bribes, and fewer cases of fraud between business partners. In Appendix E, I show that equation (5) follows if firms face a fixed cost of exporting $PS_{c,e}^{-\eta} f_{cd}$ with $\eta \in [0, 1)$ providing concavity for the impact of fixed costs f_{cd} on the exporting firms' profits. These fixed costs represent costs of setting up a distribution network, informing about markets, administration and paying for permits. A similar model has been suggested by Krautheim (2012) and it nests the established Chaney (2008) model with $\eta = 0$.

[12] The probability that two randomly drawn individuals are not from the same ethnicity is non-zero, but is captured by the country and destination fixed effects in equation (5). This model can be amended to allow for inter-ethnic trade, assuming an increasing cost of trade for ethnic groups that are further away from each other (Appendix E).

Identification assumption I obtain exogenous variation in the population shares determining the ethnic connections $\sum_{e \in E_c \cap E_d} PS_{c,t,e} \times PS_{d,t,e}$ across two countries from the exogenous placement of country borders at the 1884 Berlin conference regulating European colonization in Africa. By the stroke of a pen in Berlin, members of the same ethnic group were placed in different countries. As every different stroke would have resulted in a different composition of ethnic groups in countries c and d, their population shares are essentially random; and so is the cross-country network that I use to shift trade activity.

The construction of the instrument for Zambia is shown as an example in Figure 1. To validate this instrument, I argue that (i) the local dispersion of ethnic groups and (ii) the borders between African countries are placed without the intention to increase trade, migration, or economic activity in modern times.

First, to address endogenous sorting, I obtain exogenous variation in ethnic connections from the precolonial distribution of 833 ethnic groups (Murdock, 1959). I combine the geographic location of each group with grid-cell population data in 1960 to obtain population estimates of ethnic enclaves and their home population at the time of independence. In contrast to modern population figures, my measure of ethnic connectedness is unaffected by migration, catastrophes, hunger, or civil conflict dispersing people across Africa since independence.[13] Similar to the existing literature (Munshi, 2003; McKenzie and Rapoport, 2007), this strategy solves the reverse causality problem if populations were randomly placed in countries.

This assumption is fulfilled as African borders were drawn in 1884 at the Berlin conference. These borders do not reflect the interest of ethnic groups or African countries, but the interest of their colonizers. Most country borders feature parts that follow either latitudinal or longitudinal lines since the exact geography of Africa was largely unknown at the Berlin conference. The exogeneity of these borders has been extensively used in the literature on culture and development, price dispersion across borders as well as ethnic fractionalization (Alesina et al., 2011; Aker et al., 2014; Michalopoulos and Papaioannou, 2014).

I argue that these borders were arbitrarily drawn and do not reflect the interests of ethnic groups; to the contrary, they divide them into more than one country. The only determinant of an ethnic group being divided across two counties is its geographical size (Michalopoulos and Papaioannou, 2013, 2016).[14] To address remaining endogeneity con-

[13] Naturally, this measure includes migration until 1960. However, results are robust to using precolonial- or modern-day population figures (Table B.5).

[14] Using data on historical characteristics of tribes, neither nomadic status, the size of local communities, nor historical institutions predict a future divide into more countries. Estimating all characteristics jointly to account for correlations between variables, the size of ethnic groups is the only determinant that predicts

cerns, I only use borders where ethnic groups have been split when estimating the impact of ethnic networks on trade flows. I thus abstract from comparing influential with negligible ethnic groups and use a balanced sample across similar ethnic groups.

5.2 Leave-One-Out instrument

The second instrument exploits time variation in Africa's trading activity instead of cross-border networks to predict trade flows. Yet, these cross-border networks also motivate a modification to the standard *Leave-One-Out* methodology to completely isolate time variation from cross-sectional variation.

The standard *Leave-One-Out* estimator uses neighboring observations to predict the value of the endogenous variable in a cross-sectional setting. The idea being that geographically neighboring observations are subject to the same cross-sectional shocks, without an endogenous component that determines the strength of this shock. As the *Cross-Border* instrument already exploits this cross-sectional variation, I adjust the *Leave-One-Out* instrument to completely abstract from cross-sectional variation and isolate time-varying trends.

Instead of using the exports of neighboring country d' to predict exports of country c, I treat all neighboring countries of c as a unit ($D'(c) \subset D|d' \in D'(c)$: is neighbor of c). I then use the exports of all non-neighboring countries' $d \notin D'(c)$ to other countries $d \notin D'(c)$ to predict each country's trade activity. Figure 2 highlights the construction of this instrument using the example of Zambia, its neighboring countries, and all other African countries.

This procedure has several advantages. First, it prevents a SUTVA-type violation of the exclusion restriction: If trade activity is redirected towards more connected countries, it is likely that neighboring countries' exports are at least partially redirected from non-neighboring country d to country c. Thus, while using neighboring countries' exports generate a stronger instrument, it is likely that part of the strength arises from a correlation to unobserved networks between the two countries. For a violation of the exclusion restriction, a possible trade-diversion effect would imply that the exports of country d'' are affected by the exports of country d' which is a neighbor of c. Such a second-order violation is unlikely.

Second, this modified *Leave-One-Out* instrument exploits time variation in trading activity, rather than cross-sectional shifts. This is confirmed by the low correlation between the instruments (F-test: 2.55). Thus, finding similar point estimates when using either

the division into multiple countries (Table B.1).

instrument supports the overall finding of the paper, as it is unlikely that violations of the exclusion restriction or the exogeneity assumption affect both instruments to the same extent.

Third, the variation exploited for the *Leave-One-Out* instrument lends itself to an easier interpretation. As I exploit time variation in trading activity of the African continent, the interpretation is the same as in the OLS: how does increased trade activity affect wealth and who benefits?

5.3 First-Stage Results

I now use both the *Cross-Border* and *Leave-One-Out* instrument to predict aggregate trade activity of country c in period t.

$$\log Export_{c,t} = \delta \log \sum_{d' \in D'} \widehat{Export}_{c,d',t} + \alpha_c + \alpha_t + \varepsilon_c \qquad \text{(Cross-Border)}$$

$$\log Export_{c,t} = \delta \log \sum_{d \in D} \overline{Export}_{c,d,t} + \alpha_c + \alpha_t + \varepsilon_c \qquad \text{(Leave-One-Out)}$$

To obtain a valid first-stage F-statistic that is not inflated by multiple observations in each country and year, I begin by estimating the first stage at the country by year level predicting realized trade activity $Export_{c,t} = \sum_{d \in D} Export_{c,d,t}$. For the *Cross-Border* instrument, I obtain predicted values of bilateral trade flows $\widehat{Export}_{c,d',t}$ from the dyadic regression in equation (5) using ethnic connections with neighboring countries d' to shift trade activity.[15] For *Leave-One-Out* instrument I obtain average values of bilateral trade flows $\overline{Export}_{c,d,t}$ from the average export flows of all non-neighboring countries to all other non-neighboring countries excluding c. Both instruments are then aggregated to the country-by-year level and regressed against the realized trade activity $Export_{c,t}$, controlling for country and year fixed effects.

Figure 3 plots the strength of the *Cross-Border* instrument (left) and *Leave-One-Out* instrument (right). A one percent increase in predicted trade activity increases actual trade activity by 0.388 percent in the left panel and 0.856 in the right panel. The difference suggests that time variation, and thus economic growth in Africa, is an important explanatory factor of trading activity for each country. Yet, the F-statistics also show that cross-border connections strongly predict export activity. The 1-percentile bins are closely centered around the predicted values with an F-statistic of 79.85 in the left figure, but more widely

[15]The regressions and the procedure are outlined in Appendix B.

dispersed in the right (F-statistic 13.86).

Thus, Figure 3 reveal two instruments with strong F-statistics above 10 that are uncorrelated with each other (F-statistic 2.55) and exploit different variations. While cross-border networks, and thus cross-sectional shifts, are a strong predictor of the level of trading activity (left Figure), the leave-one-out estimation reveals that trends in economic activity unrelated to ethnic connections predict trends in trading activity.

5.4 Second-stage Results

In the second stage, I predict realized trade activity with predicted trade activity, controlling for country-by-year and country-by-ethnicity fixed effects:

$$Y_{e,c,t} = \beta \widehat{Export}_{c,t} \times Population\ Share_{e,c} + \alpha_{c,t} + \alpha_{c,e} + \varepsilon_{c,e,t}$$

I remain agnostic and cluster standard errors at the same level at the country-by-ethnic group level as bootstrapped standard errors are almost identical (Appendix Table B.8).[16]

In Table 3, I present the results on nighttime luminosity. Columns (1) and (2) replicate earlier findings from Table 1 and serve as a benchmark for IV estimates in the remaining columns. Using the cross-border instrument, the estimated size is within one standard error of the original OLS estimate and thus not statistically different (column 3). Using the leave on out instrument, I obtain slightly larger point estimates in absolute terms. Both estimates, however, confirm the initial result: People do not benefit equally from increased trade activity.

The reduced form estimates in columns (4) and (6) then interact predicted trade exposure with the political power status of the ethnic group. The results are indistinguishable from the OLS suggesting that wealth gains are redistributed from larger ethnic groups to smaller ethnic groups that are in political power.

In Table 4, I present the results on household wealth (Panel A), satisfaction with democracy (Panel B), and trust in institutions (Panel C). Again, the results on reported household wealth mirror the results on nighttime light luminosity: wealth gains are redistributed towards politically powerful groups. People exposed to more trading activity also report less satisfaction with democracy and less trust in institution, regardless of specification or instrument.

[16] As I use predicted values in the interaction term, the standard errors should be corrected for loosing a degree of freedom. However, to ensure comparability with the OLS results, I report standard errors clustered by country and ethnicity.

6 Conclusion

How is wealth distributed? Who benefits from the increased economic activity? The results in this paper provide evidence that trading increases wealth, but only for members of ruling coalitions. Ethnic groups belonging to cross-border ethnic networks are, by construction, at the border of countries and are less likely to be in power of an entire country. However, even though these ethnic groups help bridge the gap between two countries and increase trade, the gains from trade are concentrated among the group that is in power. Relocating factories and economic activity into their own ethnic homelands likely explains the negative impacts on trust institutions and satisfaction with democracy: Being left behind by the elites that govern the country, they lose trust and faith in democratic progress.

References

Aker, J. C., M. W. Klein, S. A. O'Connell, and M. Yang (2014). Borders, ethnicity and trade. *Journal of Development Economics 107*(1), 1–16.

Alesina, A., W. Easterly, and J. Matuszeski (2011). Artificial states. *Journal of the European Economic Association 25*, 246–277.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016). Ethnic inequality. *Journal of Political Economy*.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review 69*(1), 106–116.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–68.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science 293*(5536), 1818–1820.

Baccini, L., P. M. Pinto, and S. Weymouth (2017). The distributional consequences of preferential trade liberalization: Firm-level evidence. *International Organization 71*(2), 373–395.

BenYishay, A., R. Rotberg, J. Wells, Z. Lv, S. Goodman, L. Kovacevic, and D. Runfola (2017). Geocoding afrobarometer rounds 1 - 7: Methodology & data quality. Technical report, AidData. Available online at http://geo.aiddata.org.

Bühler, M. (2018). The effects of migration and ethnicitiy on african economic development. *Working Paper*.

Bühler, M. and A. Madestam (2023). State repression, exit, and voice: Living in the shadow of Cambodia's Killing Fields. *Working Paper*.

Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies 89*(1), 181–213.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padró i Miguel (2015). The value of democracy: Evidence from road building in Kenya. *American Economic Review 105*(6), 1817–1851.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–1721.

Chaney, T. (2016). Liquidity constrained exporters. *Journal of Economic Dynamics and Control 72*, 141–154.

Clochard, G. and G. Hollard (2018). Africa's growth tragedy, 20 years on. *Working Paper*.

Desmet, K. and J. F. Gomes (2023, January). Ethnic remoteness reduces the peace dividend from trade access. Working Paper 30862, National Bureau of Economic Research.

Dhingra, S. and S. Tenreyro (2020, February). The rise of agribusiness and the distributional consequences of policies on intermediated trade. CEP Discussion Papers dp1677, Centre for Economic Performance, LSE.

Dickens, A. (2018). Ethnolinguistic favoritism in African politics. *American Economic Journal: Applied Economics 10*(3), 370–402.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A "Market Access" approach. *The Quarterly Journal of Economics 131*(2), 799–858.

Egger, P. H. and M. Larch (2008). Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of International Economics* (2), 384–399.

Elvidge, C., M. Zhizhin, G. T., and T. J. Hsu FC (2021). Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing 13*(5), 922.

Elvidge, C. D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, and E. R. Davis (1997). Mapping city lights with nighttime data from the dmsp operational linescan system. *Photogrammetric Engineering and Remote Sensing 63*(6), 727–734.

Engel, J., D. Kokas, G. Lopez-Acevedo, and M. Maliszewska (Eds.) (2021). *The Distributional Impacts of Trade : Empirical Innovations, Analytical Tools, and Policy Responses.* Washington, DC: World Bank.

Feyrer, J. (2019, October). Trade and income—exploiting time series in geography. *American Economic Journal: Applied Economics 11*(4), 1–35.

Feyrer, J. (2021). Distance, trade, and income — the 1967 to 1975 closing of the suez canal as a natural experiment. *Journal of Development Economics 153*, 102708.

Frank, R. and I. Rainer (2012). Does the leader's ethnicity matter? ethnic favoritism, education, and health in sub-saharan africa. *The American Political Science Review 106*(2), 294–325.

Frankel, J. A. and D. H. Romer (1999, June). Does trade cause growth? *American Economic Review 89*(3), 379–399.

Grossman, G. M., E. Helpman, and A. Szeidl (2006). Optimal integration strategies for the multinational firm. *Journal of International Economics 70*(1), 216–238.

Kramon, E. and D. N. Posner (2016). Ethnic favoritism in education in kenya. *Quarterly Journal of Political Science 11*(1), 1–58.

Krautheim, S. (2012). Heterogenous firms, exporter networks and the effect of distance on international trade. *Journal of International Economics 87*(1), 27–35.

Lejarraga, I. (2022). Trading aims: The value of africa's deep integration trade agreement. Policy brief, ECFR.

McKenzie, D. and H. Rapoport (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics 84*, 1–24.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica 81*(1), 113–152.

Michalopoulos, S. and E. Papaioannou (2014). National institutions and subnational development in Africa. *Quarterly Journal of Economics 129*(1), 151–213.

Michalopoulos, S. and E. Papaioannou (2016). The long-run effects of the scramble for Africa. *American Economic Review 106*(7), 1802–1848.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the U.S. labor market. *Quarterly Journal of Economics 118*(2), 549–599.

Murdock, G. (1959). *Africa: Its Peoples and Their Culture History*. New York: McGraw-Hill.

Pascali, L. (2017, September). The wind of change: Maritime technology, trade, and economic development. *American Economic Review 107*(9), 2821–54.

Santos-Silva, J. and S. Tenreyro (2006). The log of gravity. *Review of Economics and Statistics 88*(4), 641–658.

Smeets, M. (2021, September). Africa's integration in the WTO multilateral trade system: Academic support and the role of WTO chairs. Working Paper 2021-9, World Trade Organization Economic Research and Statistics Division.

Spolaore, E. and R. Wacziarg (2015). Ancestry, language and culture. *Working Paper*.

Steinwender, C. (2018, March). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review 108*(3), 657–96.

Weidmann, N. B., J. K. Rød, and L.-E. Cedermann (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research 47*(4).

Wimmer, A., L.-E. Cederman, and B. Min (2009). Ethnic politics and armed conflict. A configurational analysis of a new global dataset. *American Sociological Review 74*(2), 316–337.

Table 1: Trade exposure and wealth gains
Using nighttime light satellite imagery

	Fraction lid		Average luminosity	
	(1)	(2)	(3)	(4)
Trade Exposure	-0.008***	-0.009***	-0.029**	-0.024**
	(0.003)	(0.003)	(0.012)	(0.011)
Trade Exposure × In Power		0.010**		0.028*
		(0.004)		(0.016)
Country × year fixed effects	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes
Observations	10,314	8,541	10,314	8,541
Mean dependent variable	0.033	0.034	0.106	0.116

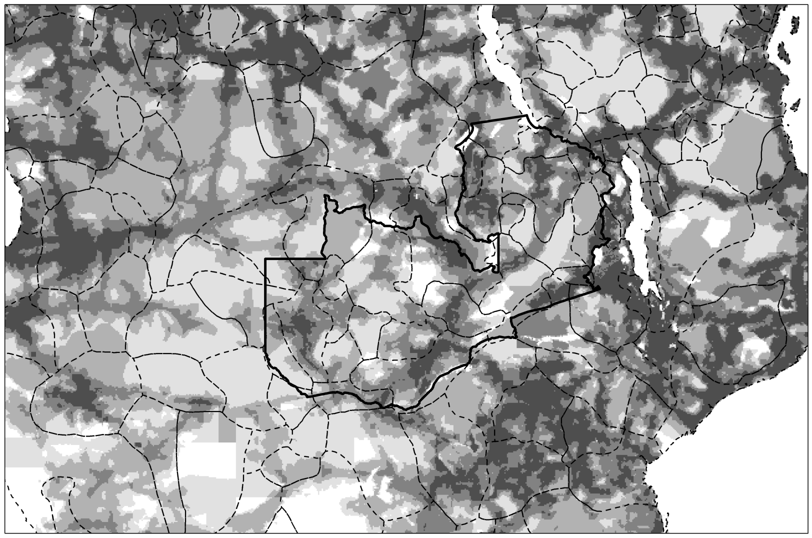
In this table, I show how trade exposure impacts the distribution of wealth as measured by nighttime luminosity. *Trade Exposure* is defined by realized trade flows to all African countries aggregated to the country-by-year level interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Fraction lid* is calculated as the fraction of pixels not zero and *Average luminosity* as the log of average luminosity in each country-ethnic group observation plus one. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Trade exposure, wealth, and society
Household wealth from the Afrobarometer

	Household wealth		Satisfaction with Democracy		Trust in Institutions	
	(1)	(2)	(3)	(4)	(5)	(6)
Trade Exposure	-0.069***	-0.076**	-0.207***	-0.263***	-0.127***	-0.152***
	(0.025)	(0.034)	(0.056)	(0.060)	(0.049)	(0.051)
Trade Exposure × In Power		0.036***		-0.025		-0.012
		(0.013)		(0.020)		(0.011)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	218,950	194,775	218,950	194,775	218,950	194,775

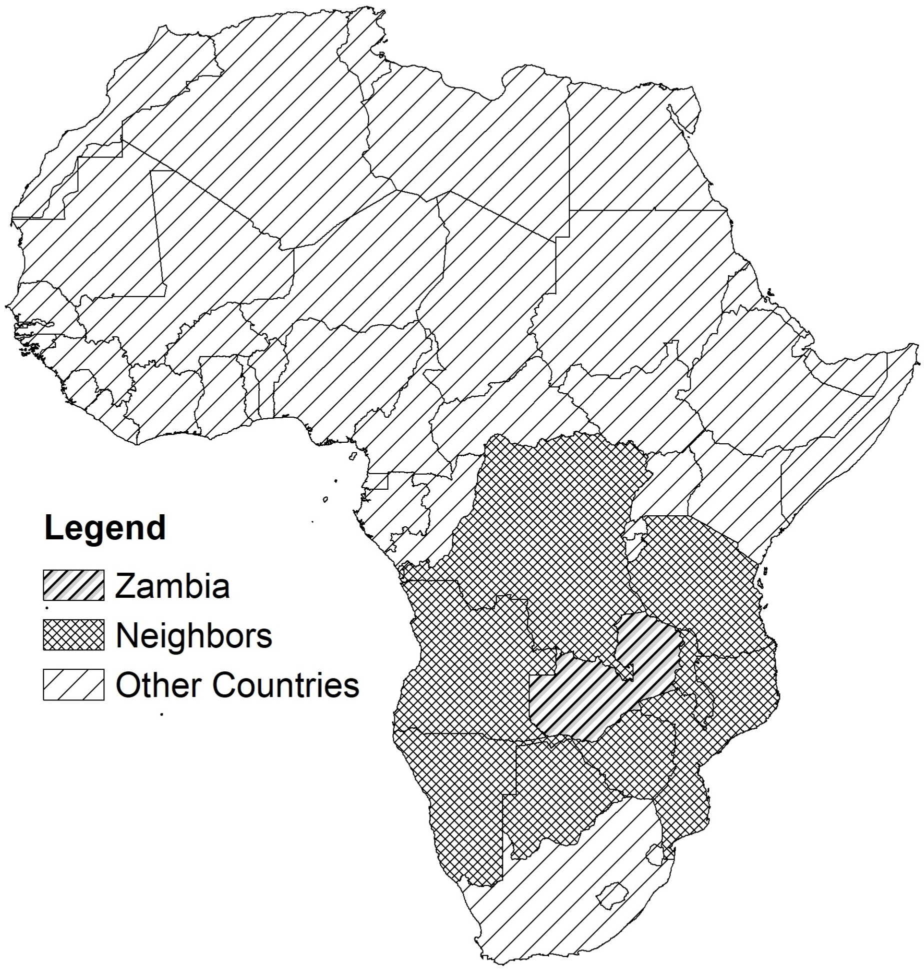
In this table, I show how trade exposure impacts household wealth as measured by the Afrobarometer *Trade Exposure* is defined by realized trade flows to all African countries aggregated to the country-by-year level interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Individual controls* are a full set of age, gender, education, and urban dummies. *Household wealth* represents a standardized index constructed from 9 variables asked in 7 rounds of the Afrobarometer. Details in the Appendix. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 1: Construction of the *Cross-Border* instrument



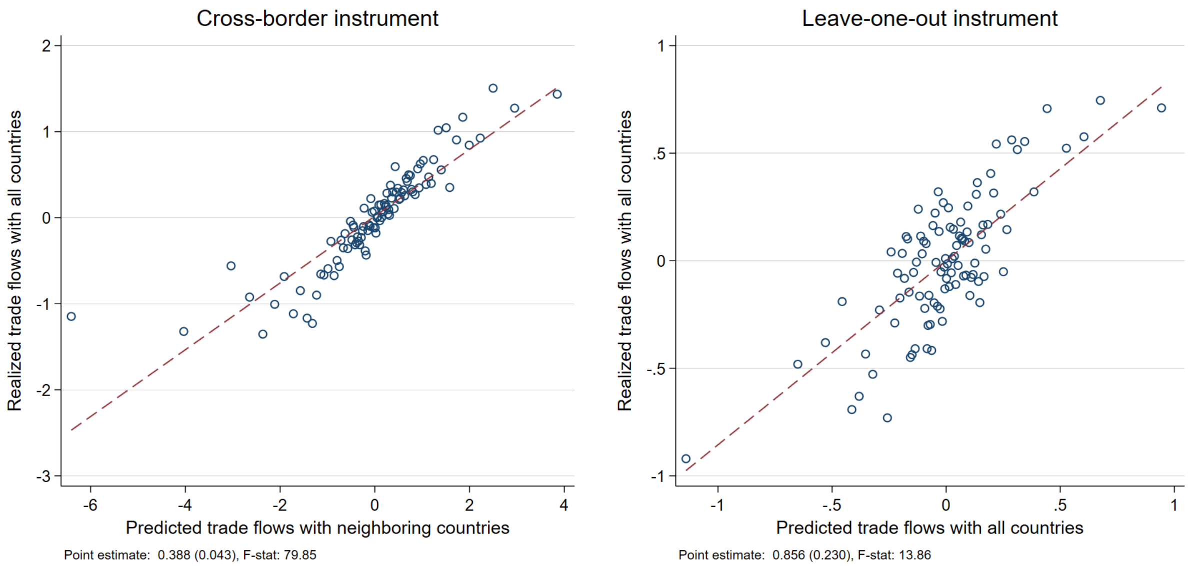
Notes: Using Zambia (solid line), this map exemplifies the construction of the *Cross-Border* instrument. Ethnic groups, as defined by Murdock (1959) span multiple countries (dashed line). Population figures are shown as the shaded background with darker colors representing denser population. The ethnic network is defined as the population share of ethnic group *e* in Zambia multiplied with its population share outside Angola (to the West of Zambia). The *Cross-Border* instrument is then the sum of all ethnic networks between Zambia and Angola. This instrument is valid as no single country border follows an ethnic border and some borders are straight lines. Then, as country borders determine population shares, which in turn define the strength of the instrument, the instrument is exogenous from the individual's perspective.

Figure 2: Construction of the *Leave-One-Out* instrument



Notes: Using Zambia, this map exemplifies the construction of the *Leave-One-Out* instrument. Zambia's exports to all countries are replaced by the average bilateral exports of all "Other countries" to all "Other countries". In the data, this means that instead of the possible 2,256 trade observations (48×47), only 1,854 are used on average. Due to cross-border networks, "Neighbors" are excluded from this to isolate variation that is entirely driven by increasing trade activity in Africa.

Figure 3: First stage relationship of the instruments



Notes: This figure plots the correlation between aggregated predicted and realized trade flows using the cross-border instrument (left) and the Leave-one-out instrument (right). Both plots show residualized values, controlling for country and year fixed effects. The linear fit is shown with the dashed line in each panel and its slope and F-statistic noted below. Both instruments are only weakly correlated (0.026, s.e.: 0.016) indicating that the exploited variation is different: The cross-border instrument (left) exploits cross-sectional variation between neighboring countries; the Leave-on-out instrument (right) exploits time variation in economic activity of non-neighboring countries.

Table 3: Trade exposure and wealth gains
IV results

	OLS		Cross-Border		Leave-One-Out	
	(1) OLS	(2) OLS	(3) IV	(4) RF	(5) IV	(6) RF
Panel A: Fraction of pixel lid						
Trade Exposure	-0.008***	-0.009***	-0.012***		-0.034***	
	(0.003)	(0.003)	(0.004)		(0.013)	
Trade Exposure × In Power		0.010**				
		(0.004)				
Predicted Trade Exposure				-0.008***		-0.042***
				(0.002)		(0.011)
Predicted Trade Exposure × In Power				0.007*		0.011***
				(0.004)		(0.003)
Panel B: Average nighttime luminosity						
Trade Exposure	-0.029**	-0.024**	-0.028**		-0.088*	
	(0.012)	(0.011)	(0.012)		(0.048)	
Trade Exposure × In Power		0.028*				
		(0.016)				
Predicted Trade Exposure				-0.019**		-0.078*
				(0.008)		(0.042)
Predicted Trade Exposure × In Power				0.025		0.035***
				(0.015)		(0.013)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Reduced form				Yes		Yes
Observations	10,314	8,559	10,128	8,398	10,314	8,559
First Stage F-Test			29.452		17.831	

In this table, I show how trade exposure impacts the distribution of wealth as measured by nighttime luminosity. *Trade Exposure* is defined as realized trade flows to all African countries aggregated to the country-by-year level and interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. In columns (3) and (5) it is instrumented by *Predicted Trade Exposure* using either the Border instrument or the Leave-One-Out instrument. *Fraction lid* is calculated as the fraction of pixels not zero, *Average nighttime luminosity* as the log of average luminosity in each country-ethnic group observation plus one. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. The first stage F statistic is given in the last row. Corrected F-Statistics at the country-year level presented in Figure 3. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Trade exposure and society
Democracy, trust, and conflict

	OLS		Cross-Border		Leave-One-Out	
	(1) OLS	(2) OLS	(3) IV	(4) RF	(5) IV	(6) RF
Panel A: Household Wealth						
Trade Exposure	-0.069***	-0.076**	-0.104**		-0.061	
	(0.025)	(0.034)	(0.041)		(0.043)	
Trade Exposure × In Power		0.036***				
		(0.013)				
Predicted Trade Exposure				-0.077**		-0.055
				(0.031)		(0.046)
Predicted Trade Exposure × In Power				0.039***		0.035**
				(0.013)		(0.014)
Panel B: Satisfaction with Democracy						
Trade Exposure	-0.207***	-0.263***	-0.330***		-0.278***	
	(0.056)	(0.060)	(0.071)		(0.066)	
Trade Exposure × In Power		-0.025				
		(0.020)				
Predicted Trade Exposure				-0.247***		-0.303***
				(0.055)		(0.078)
Predicted Trade Exposure × In Power				-0.019		-0.032
				(0.020)		(0.020)
Panel C: Trust in Institutions						
Trade Exposure	-0.127***	-0.152***	-0.187***		-0.137***	
	(0.049)	(0.051)	(0.060)		(0.048)	
Trade Exposure × In Power		-0.012				
		(0.011)				
Predicted Trade Exposure				-0.140***		-0.148***
				(0.045)		(0.054)
Predicted Trade Exposure × In Power				-0.009		-0.013
				(0.011)		(0.012)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Reduced form				Yes		Yes
Observations	218,950	194,775	193,671	193,671	194,775	194,775
First Stage F-Test			1381.449		267.612	

In this table, I show how trade exposure impacts household wealth as measured by the Afrobarometer. *Trade Exposure* is defined by realized trade flows to all African countries aggregated to the country-by-year level interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Individual controls* are a full set of age, gender, education, and urban dummies. In columns (3) and (5) it is instrumented by *Predicted Trade Exposure* using either the Border instrument or the Leave-One-Out instrument. *Household wealth*, *Satisfaction with Democracy*, and *Trust in Institutions* represent standardized indexes constructed from variables asked in 7 rounds of the Afrobarometer. Details in the Appendix. The first stage F statistic is given in the last row. Corrected F-Statistics at the country-year level presented in Figure 3. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For Online Publication: Appendix
23rd December 2023

This appendix provides supporting evidence to the main paper. I cover the following topics:

A **Supporting Evidence**

B **Additional Information for the Instrumental Variables Strategy**

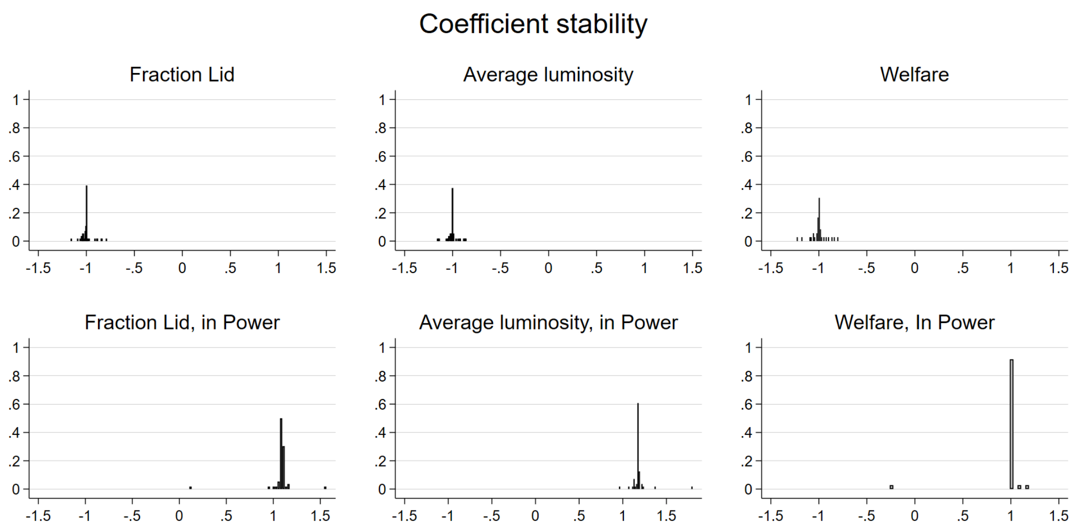
C **Questions of the Afrobarometer Rounds**

D **Replicating Results using the DHS**

D **Technical Appendix**

A Supporting Evidence

Figure A.1: Coefficient stability for the main regressions in Table 1 and 2



Notes: This figure plots the distribution of point estimates when dropping one country at a time relative to the average effect. 1 implies that the point estimate is the same, 0.5 implies its 50% smaller.

Table A.1: Trade exposure and the gains from trade
Sectoral exports and nighttime light satellite imagery

	Fraction lid				Average luminosity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade Exposure	-0.008***				-0.029**			
	(0.003)				(0.012)			
Exposure to manufacturing exports		-0.004**				-0.021**		
		(0.002)				(0.010)		
Exposure to agriculture exports			-0.002				-0.007	
			(0.002)				(0.005)	
Exposure to resource exports				-0.003				-0.008
				(0.003)				(0.006)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,314	10,845	10,845	10,845	10,314	10,845	10,845	10,845

In this table, I show how trade exposure in various sectors impacts economic development as measured by nighttime luminosity. Columns (1) and (3) represent the instrumental variables estimation, columns (2) and (4) the reduced form evidence. *Trade Exposure*, *Exposure to manufacturing exports*, *Exposure to agriculture exports*, and *Exposure to resource exports* are defined as realized trade flows to all African countries aggregated to the country-by-year level in that sector and interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Trade Exposure* is identified only from the interaction of the two. *Fraction lid* is calculated as the fraction of pixels not zero and *Average luminosity* as the log of average luminosity in each country-ethnic group observation plus one. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.1 Derivation of the main estimating equation

In this section, I want to highlight how I derive the fully saturated regression from principles. I report regressions estimates from all iterations between the following two equations in Table A.2:

$$Y_{e,c,t} = \beta_1 \log\left(\sum_{d\in D} Export_{c,d,t}\right) + \beta_2 Population\ Share_{e,c} + \alpha_c + \alpha_t + \varepsilon_{e,c,t} \tag{1}$$

$$Y_{e,c,t} = \gamma \underbrace{\log\left(\sum_{d\in D} Export_{c,d,t}\right) \times PopulationShare_{e,c}}_{TradeExposure_{e,c,t}} + \alpha_{c,t} + \alpha_{c,e} + \varepsilon_{e,c,t} \tag{2}$$

In the first column of Table A.2 I report the simplest regression equation (1). Here, exports are negatively correlated to economic activity at the regional level, whereas GDP per capita and the size of ethnicity *e* are positively correlated as larger ethnic groups have more land area and higher national GDP per capita should be correlated with nighttime light luminosity.

In the second column I include the interaction term capturing trade exposure. While this term enters positively, the population share of ethnicity *e* now has a large negative impact on nighttime light luminosity. This is likely due to unobserved characteristics of the ethnic group. As some ethnic groups have a history of state formation, while others have not, these effect likely confound this estimate. These same effects also confound the estimate on Trade Exposure as historical state formation likely has a positive impact on trade flows; hence this estimate is likely upward biased.

In the third column I thus include ethnicity fixed effects, confirming the previous suspicion on biases as both the estimate on population shares and trade exposure change drastically. In addition, the adjusted R^2 increases from 0.329-0.924, highlighting the importance of these fixed effects.

In the fourth column I now include country-by-year fixed effects to control for total trade flows, GDP, and changing populations in each country. The remaining estimates remain unchanged.

At this point, it becomes clear that the point estimate on the population share of ethnicities (-0.211, column 4) is estimated within ethnic group, but across multiple countries—otherwise it would be dropped from the regression: it captures how the same ethnic group is treated in two different countries. The point estimate suggests that the larger its population share in the respective country, the worse is its economic outcomes, *on average*. While this already provides evidence against an equitable distribution of wealth, I also want to isolate the relative status effect of ethnic groups from the impact of increased economic activity: While trade exposure captures how the ethnic group is benefiting from additional trade, the level effect captures in its population share captures, in addition to the interaction term, how it suffers.

In the final column I thus include country-by-ethnicity fixed effects to capture all observed and unobserved characteristics of each ethnic group in each country, in addition to including country-by-year fixed effects that control for total trade flows, GDP, and changing populations in each country. Now, the point estimate on Trade Exposure only identifies the variation from additional economic output that affects the geographic distribution of economic activity through each ethnic groups' population shares.

Table A.2: Trade exposure and the gains from trade
Deriving the estimating equation

	Fraction lid				
	(1)	(2)	(3)	(4)	(5)
Trade Exposure		0.094***	0.021*	0.021*	-0.008***
		(0.014)	(0.012)	(0.011)	(0.003)
log Realized Exports	-0.001**	-0.005***	-0.002***		
	(0.000)	(0.001)	(0.001)		
Population share Ethnicity	0.214***	-1.013***	-0.219	-0.211*	
	(0.042)	(0.174)	(0.145)	(0.123)	
log(GDP per capita)	0.008***	0.010***	0.008***		
	(0.002)	(0.002)	(0.002)		
Country fixed effect	yes	yes	yes	yes	yes
Year fixed effect	yes	yes	yes	yes	yes
Ethnicity fixed effect			yes	yes	yes
Country × year fixed effects				yes	yes
Country × ethnicity fixed effect					yes
Mean dependent variable	0.033	0.033	0.033	0.033	0.033
Adjusted R2	0.295	0.329	0.924	0.929	0.986
Observations	10,204	10,204	10,204	10,314	10,314

In this table, I show how trade exposure in various sectors impacts economic development as measured by nighttime luminosity, iterating through possible estimation equations. *Country fixed effects* account for unobserved characteristics at the country level, including size and location of country. *Year fixed effect* account for unobserved characteristics at the year level, including the African economic situation. *Ethnicity fixed effects* account for unobserved characteristics at the ethnicity level, including its size and population. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Trade Exposure* is identified only from the interaction of the two. *Fraction lid* is calculated as the fraction of pixels not zero. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Cross-border instrument

In this Appendix, I provide additional evidence for the Cross-Border instrument. I begin by showing that the only determinant of ethnic groups being split between two countries is its geographic size, not its population (Table B.1, Column 8).

I then estimate a simple dyadic equation, regressing the size of the ethnic connections on export flows. Each population share $PS_{c,e}$ is calculated by the intersection of ethnic homelands e with country borders c and sums up all grid-cell population data in 1960 from the United Nations Environment Program in 1960. This number is denominated by country c's population in 1960 to arrive at $PS_{c,e} \in (0, 1]$.

$$\log(X_{cd,t}) = \beta \underbrace{\log\left(\sum_{e \in E_c \cap E_d} PS_{c,e} \times PS_{d,e}\right)}_{\textit{Ethnic Connections}_{cd}} + \Gamma_{cd,t} + \alpha_{c,t} + \alpha_{d,t} + \varepsilon_{c,d,t} \tag{5}$$

Every regression follows the standard in the trade literature and includes time-specific country- and destination fixed effects and country-destination pair characteristics (Γ_{cd}). Country-destination pair characteristics include the log of the distance, number of ethnic connections, sharing a colonial history or language, linguistic and genetic similarity, and geographical features of the border. A positive point estimate, $\beta > 0$, suggests that a larger population on either side of the border for a connected ethnicity yields larger trade flows.

I estimate this equation at the country-destination by year level. The final sample consists of 46 African countries in 91 country pairs with 182 country-destination relationships that share a border. Due to unobserved trade, the sample is further reduced to 169 observations from 1990–2020. Since the exploited variation is at the country-pair level, I cluster the standard errors at this level. I report estimates using ordinary least squares and show robustness to using weighted least squares, a panel estimation with country specific year fixed effects, the Poisson-Pseudo-Maximum-Likelihood estimator as suggested by Santos-Silva and Tenreyro (2006), as well as sectoral exports.[17]

I begin by estimating equation (5) in a yearly panel of country-pairs. In this country-destination panel between neighboring countries, exports increase by 0.17% when networks increase by 1% with an F-statistic of 9.92 (Table B.2, Column 2). This result is robust to controlling for regional trade agreements and conflict (Column 3), using a Poisson-Pseudo-Maximum-Likelihood (PPML) estimator (Column 4), or only exploiting the country-destination level variation in networks in a weighted regression (Table B.3). Finally, I

[17] For a more detailed breakdown of the estimation strategy see Bühler (2018).

replicate the IV-results on wealth gains using the PPML estimator in Tables B.6 and B.7.

Table B.1: Determinants of being divided:
Historical characteristics of ethnic groups in Murdock (1959)

	Ethnic group is divided between two or more countries							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log Population in 1960	0.041***	0.008						0.015
	(0.013)	(0.015)						(0.021)
	[0.011]	[0.011]						[0.017]
log Ethnic Area		0.109***						0.138***
		(0.019)						(0.022)
		[0.013]						[0.016]
log Population Density			-0.031**				-0.050***	
			(0.015)				(0.021)	
			[0.011]				[0.014]	
Cities				-0.087			-0.084	-0.046
				(0.055)			(0.059)	(0.060)
				[0.050]			[0.051]	[0.049]
Mean Size of Local Communities					0.013		0.020*	0.004
					(0.012)		(0.011)	(0.011)
					[0.011]		[0.011]	[0.011]
Political Centralization						0.036	0.038	-0.072
						(0.055)	(0.053)	(0.050)
						[0.051]	[0.051]	[0.051]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	833	833	833	441	441	441	441	441
Adjusted R-squared	0.022	0.086	0.014	0.017	0.014	0.011	0.038	0.134

Every column shows the point estimate from a regression on the probability of an ethnicity being divided between two or more countries. Geographic Controls include latitude, longitude, and their product. log Population in 1960 taken from UNEP SIOUX grid cell data. log Ethnic Area is the total expansion area of an ethnicity as given by the Murdock map. Data in columns (4)–(8) taken from Michalopoulos and Papaioannou (2013) and coded as follows. 'Cities': If at least one ethnicity that crosses the border historically had permanent or complex settlements. 'Political Centralization' If at least one ethnicity that crosses the border historically had a jurisdictional level beyond the local level: centralized tribe$\geq$2, where 'centralized tribe' is the count variable of jurisdictional level beyond the local level (range: 0-3). Standard errors corrected for spatial correlation within 500km shown in parenthesis. Lower cutoffs decrease the standard errors to the robust standard errors level shown in brackets. Symbols reflect the significance level for spatially corrected standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Ethnic connections and trade flows
log(Exports) between neighboring countries

	OLS			PPML
	(1)	(2)	(3)	(4)
Ethnic connections	0.183**	0.173**	0.210***	0.278***
	(0.078)	(0.070)	(0.080)	(0.066)
	[0.044]	[0.055]	[0.058]	[0.107]
Linguistic distance		-5.270**	-4.840**	-8.222***
		(2.167)	(1.908)	(0.964)
		[-0.049]	[-0.050]	[-0.128]
Genetic distance		26.575***	24.480***	31.519***
		(5.242)	(5.279)	(5.445)
		[-0.085]	[-0.051]	[0.346]
Regional trade agreements			-0.279	-0.301
			(0.523)	(0.462)
			[-0.009]	[-0.036]
log(# Conflict)			-0.597	-0.212
			(0.576)	(0.337)
			[-0.007]	[-0.019]
log(# Civilian deaths)			0.500	0.246
			(0.308)	(0.218)
			[0.006]	[0.013]
log(# Deaths)			0.035	0.046
			(0.131)	(0.077)
			[0.001]	[-0.002]
lnUnknown			0.124	0.052
			(0.252)	(0.128)
			[0.000]	[0.011]
Country-pair controls	Yes	Yes	Yes	Yes
Exporter and importer $\times$ year fixed effects	Yes	Yes	Yes	Yes
Conflict controls	4195	4195	4195	4198

In this table, I show that ethnic connections predict bilateral exports per capita between neighboring countries. *Ethnic connections* are defined as the log ethnic match probability as defined in equation (5) and capture the likelihood of drawing two individuals from either country with the same ethnicity. Sample consist of bordering countries in Africa and includes country and destination by year fixed effects in all regressions. The main dependent variable are the logarithm of bilateral exports per current capita in the years 1992–2018. The following country-pair controls are added in all columns: log length of the border, log distance between capitals, a binary variable indicating whether parts of the border is a river, a mountain above 1,000 or 2,000 meter. The number of ethnic groups shared between the countries, whether the countries share a colonial past or judicial language, and log border fractionalization (Alesina et al., 2011). Linguistic and genetic distance $\in [0, 1]$ captures the similarity between the countries (Spolaore and Wacziarg, 2015). Regional trade agreements for the years 1989-2020 are obtained from Mario Larch's Regional Trade Agreements Database from Egger and Larch (2008). Conflict at the country level is obtained from UDCP https://ucdp.uu.se/. Inference is made on the basis of standard errors clustered by each country-pair, as shown in parenthesis. Two-way clustered standard errors allowing for separate home- and foreign-clusters shown as robustness in brackets. OLS and PPML denote the estimation method. Significance denoted by standard errors clustered by the country pair: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Ethnic connections and trade flows

Weighted exports between neighboring countries

	OLS			PPML
	(1)	(2)	(3)	(4)
Ethnic connections	0.125	0.158*	0.142**	0.147**
	(0.091)	(0.087)	(0.061)	(0.066)
	[0.066]	[0.074]	[0.057]	[0.085]
Country-pair controls		Yes	Yes	Yes
Weighted regression			Yes	Yes
Observations	169	169	4,195	4,195

In this table, I show that ethnic connections predict bilateral exports per capita between neighboring countries, when the data is collapsed to the level of variation in this setting: the country-destination pair. *Ethnic connections* are defined as the log ethnic match probability as defined in equation (5) and capture the likelihood of drawing two individuals from either country with the same ethnicity. Sample consist of bordering countries in Africa and includes country and destination by year fixed effects in all regressions. The main dependent variable are the logarithm of bilateral exports per current capita in the years 1992–2018. The following country-pair controls are added in all columns: log length of the border, log distance between capitals, a binary variable indicating whether parts of the border is a river, a mountain above 1,000 or 2,000 meter. The number of ethnic groups shared between the countries, whether the countries share a colonial past or judicial language, and log border fractionalization (Alesina et al., 2011). Linguistic and genetic distance $\in [0,1]$ captures the similarity between the countries (Spolaore and Wacziarg, 2015). Inference is made on the basis of standard errors clustered by each country-pair, as shown in parenthesis. Two-way clustered standard errors allowing for separate home- and foreign-clusters shown as robustness in brackets. Significance denoted by standard errors clustered by the country pair: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: Ethnic connections and trade flows
Sectoral exports between neighboring countries

	log(Export)	log(Manufacturing)	log(Agriculture)	log(Resources)
	(1)	(2)	(3)	(4)
Ethnic connections	0.171**	0.258***	0.144**	0.068
	(0.070)	(0.074)	(0.070)	(0.108)
	[0.054]	[0.055]	[0.075]	[0.097]
Country-pair controls	Yes	Yes	Yes	Yes
Exporter and importer × year fixed effects	Yes	Yes	Yes	Yes
Conflict controls	4199	4199	4199	4199

In this table, I show that ethnic connections predict bilateral exports per capita between neighboring countries. *Ethnic connections* are defined as the log ethnic match probability as defined in equation (5) and capture the likelihood of drawing two individuals from either country with the same ethnicity. Sample consist of bordering countries in Africa and includes country and destination by year fixed effects in all regressions. The main dependent variable are the logarithm of bilateral exports per current capita in the years 1992–2018. The following country-pair controls are added in all columns: log length of the border, log distance between capitals, a binary variable indicating whether parts of the border is a river, a mountain above 1,000 or 2,000 meter. The number of ethnic groups shared between the countries, whether the countries share a colonial past or judicial language, and log border fractionalization (Alesina et al., 2011). Linguistic and genetic distance $\in [0, 1]$ captures the similarity between the countries (Spolaore and Wacziarg, 2015). Inference is made on the basis of standard errors clustered by each country-pair, as shown in parenthesis. Two-way clustered standard errors allowing for separate home- and foreign-clusters shown as robustness in brackets. Significance denoted by standard errors clustered by the country pair: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table B.5: Ethnic connections and trade flows
Alternative population data

	OLS			
	(1)	(2)	(3)	(4)
Ethnic connections	0.173**			
	(0.070)			
	[0.055]			
Ethnic connections, pre-colonial population		0.109**		
		(0.053)		
		[0.037]		
Ethnic network according to model			0.211**	
			(0.086)	
			[0.068]	
Ethnic connections, Ethnologue data today				0.146
				(0.111)
				[0.074]
Country-pair controls	Yes	Yes	Yes	Yes
Exporter and importer $\times$ year fixed effects	Yes	Yes	Yes	Yes
Conflict controls	4195	4195	4195	4195

In this table, I show that ethnic connections predict bilateral exports per capita between neighboring countries. *Ethnic connections* are defined as the log ethnic match probability as defined in equation (5) and capture the likelihood of drawing two individuals from either country with the same ethnicity. *Ethnic connections, pre-colonial population* are defined by the population figures for each ethnic group as recorded in Murdock (1959), multiplied with their population share in 1960 to obtain country-by-ethnic group population figures. *Ethnic network according to model* is defined following the model in Appendix E. *Ethnic connections, Ethnologue data today* are defined by current day population shares by ethnicity in each country. Sample consist of bordering countries in Africa and includes country and destination by year fixed effects in all regressions. The main dependent variable are the logarithm of bilateral exports per current capita in the years 1992–2018. The following country-pair controls are added in all columns: log length of the border, log distance between capitals, a binary variable indicating whether parts of the border is a river, a mountain above 1,000 or 2,000 meter. The number of ethnic groups shared between the countries, whether the countries share a colonial past or judicial language, and log border fractionalization (Alesina et al., 2011). Linguistic and genetic distance $\in [0, 1]$ captures the similarity between the countries (Spolaore and Wacziarg, 2015). Inference is made on the basis of standard errors clustered by each country-pair, as shown in parenthesis. Two-way clustered standard errors allowing for separate home- and foreign-clusters shown as robustness in brackets. Significance denoted by standard errors clustered by the country pair: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: IV Results: Trade exposure and the gains from trade
Using a PPML estimation in the first stage

	OLS		Cross-Border (OLS)		Cross-Border (PPML)	
	(1) OLS	(2) OLS	(3) IV	(4) RF	(5) IV	(6) RF
Panel A: Fraction of pixel lid						
Trade Exposure	-0.008***	-0.009***	-0.012***		-0.010***	
	(0.003)	(0.003)	(0.004)		(0.004)	
Trade Exposure × In Power		0.010**				
		(0.004)				
Predicted Trade Exposure				-0.008***		-0.006***
				(0.002)		(0.002)
Predicted Trade Exposure × In Power				0.007*		0.005
				(0.004)		(0.005)
Panel B: Average nighttime luminosity						
Trade Exposure	-0.029**	-0.024**	-0.028**		-0.027**	
	(0.012)	(0.011)	(0.012)		(0.012)	
Trade Exposure × In Power		0.028*				
		(0.016)				
Predicted Trade Exposure				-0.019**		-0.018**
				(0.008)		(0.008)
Predicted Trade Exposure × In Power				0.025		0.022
				(0.015)		(0.017)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Reduced form				Yes		Yes
Observations	10,314	8,559	10,128	8,398	10,128	8,398
First Stage F-Test			29.452		29.690	

In this table, I show how trade exposure impacts the distribution of wealth as measured by nighttime luminosity. *Trade Exposure* is defined as realized trade flows to all African countries aggregated to the country-by-year level and interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$ instrumented by *Predicted Trade Exposure*. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Trade Exposure* and *Predicted Trade Exposure* are identified only from the interaction of the two. *Fraction lid* is calculated as the fraction of pixels not zero, *Average nighttime luminosity* as the log of average luminosity in each country-ethnic group observation plus one. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: Trade exposure, wealth, and society
Using a PPML estimation in the first stage

	OLS		Cross-Border (OLS)		Cross-Border (PPML)	
	(1) OLS	(2) OLS	(3) IV	(4) RF	(5) IV	(6) RF
Panel A: Household wealth						
Trade Exposure	-0.069***	-0.076**	-0.104**		-0.107***	
	(0.025)	(0.034)	(0.041)		(0.039)	
Trade Exposure × In Power		0.036***				
		(0.013)				
Predicted Trade Exposure				-0.077**		-0.076**
				(0.031)		(0.030)
Predicted Trade Exposure × In Power				0.039***		0.037***
				(0.013)		(0.013)
Panel B: Satisfaction with Democracy						
Trade Exposure	-0.207***	-0.263***	-0.330***		-0.311***	
	(0.056)	(0.060)	(0.071)		(0.067)	
Trade Exposure × In Power		-0.025				
		(0.020)				
Predicted Trade Exposure				-0.247***		-0.230***
				(0.055)		(0.052)
Predicted Trade Exposure × In Power				-0.019		-0.025
				(0.020)		(0.020)
Panel C: Trust in Institutions						
Trade Exposure	-0.127***	-0.152***	-0.187***		-0.191***	
	(0.049)	(0.051)	(0.060)		(0.058)	
Trade Exposure × In Power		-0.012				
		(0.011)				
Predicted Trade Exposure				-0.140***		-0.142***
				(0.045)		(0.043)
Predicted Trade Exposure × In Power				-0.009		-0.013
				(0.011)		(0.011)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes				
Reduced form				Yes		Yes
Observations	218,950	194,775	193,671	193,671	193,671	193,671
First Stage F-Test			1381.449		304.244	

In this table, I show how trade exposure impacts household wealth as measured by the Afrobarometer *Trade Exposure* is defined by realized trade flows to all African countries aggregated to the country-by-year level interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Trade Exposure* is identified only from the interaction of the two. *Individual controls* are a full set of age, gender, education, and urban dummies. *Household wealth*, *Satisfaction with Democracy*, and *Trust in Institutions* represent standardized indexes constructed from variables asked in 7 rounds of the Afrobarometer. Details in the Appendix. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: Trade exposure, wealth, and society
Bootstrapped standard errors

	Fraction lid		Average luminosity		Household wealth		Satis. w. Democracy		Trust in Inst.	
	(1) CB	(2) LOO	(3) CB	(4) LOO	(5) CB	(6) LOO	(7) CB	(8) LOO	(9) CB	(10) LOO
b										
Trade Exposure	-0.012**	-0.034*	-0.028*	-0.088	-0.104*	-0.061	-0.330***	-0.278***	-0.187**	-0.137**
	(0.005)	(0.017)	(0.016)	(0.054)	(0.056)	(0.060)	(0.103)	(0.100)	(0.082)	(0.068)
se										
Trade Exposure	0.004**	0.013	0.012*	0.048**	0.041**	0.043**	0.071***	0.066***	0.060***	0.048***
	(0.002)	(0.009)	(0.007)	(0.024)	(0.019)	(0.017)	(0.021)	(0.024)	(0.020)	(0.016)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls					Yes	Yes	Yes	Yes	Yes	Yes
Reduced form	10128	10314	10128	10314	193671	194775	193671	194775	193671	194775

In this table, I replicate the IV results using bootstrapped standard errors. Odd columns with "CB" replicate the cross-border instrument; even columns with "LOO" the leave-one-out instrument. *Trade Exposure* is defined by realized trade flows to all African countries aggregated to the country-by-year level interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. *Trade Exposure* is identified only from the interaction of the two. *Individual controls* are a full set of age, gender, education, and urban dummies. *Fraction lid* is calculated as the fraction of pixels not zero, *Average nighttime luminosity* as the log of average luminosity in each country-ethnic group observation plus one. *Household wealth*, *Satisfaction with Democracy*, and *Trust in Institutions* represent standardized indexes constructed from variables asked in 7 rounds of the Afrobarometer. Details in the Appendix. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Afrobarometer questions

In this Appendix, I highlight the methodology and questions that go into creating the standardized indexes for household wealth, satisfaction with democracy, and trust in institutions.

Standardized Index Following Anderson (2008), I standardize each question within these three categories and sum the standardized outcomes, weighting each question by the inverse of the covariance matrix of the standardized outcomes. The four indices address concerns of multiple hypothesis testing and aggregate changes that individual questions only measure imperfectly. I also present results for individual questions with estimated p-values and false discovery rate-adjusted p-values, computed using the procedure outlined in Bühler and Madestam (2023).

The individual questions are shown in Tables C.2–C.4 at the example of the Afrobarometer Round 7. All questions are recoded such that higher values in the question would correspond a positive impact of trade exposure. For example, question q8a "*Over the past year, how often, if ever, have you or anyone in your family: Gone without enough food to eat?*" is recoded such that the meaning is reversed: While in q8a higher values imply *more* food insecurity, the recoded variable has higher values for *less* food insecurity.

Table C.1: Trade exposure and the gains from trade
Individual responses and multiple hypothesis testing

	(1)	(2)	(3)	(4)
	beta	s.e.	*p*-value	FDR adj. *p*-value
Household Wealth				
Country Condition	-0.125	0.086	0.145	0.159
Household Condition, Compared	-0.025	0.071	0.728	0.282
Household Condition	-0.133	0.078	0.088	0.124
Household Condition, Past	-0.189	0.136	0.165	0.159
Days with Food	-0.130	0.051	0.011	0.032
Days with Water	-0.212	0.083	0.010	0.032
Days with Cash	-0.046	0.055	0.400	0.282
Days with Fuel	-0.167	0.066	0.011	0.032
Satisfaction with Democracy				
Status of Democracy	-0.213	0.079	0.007	0.007
Satisfied with Democracy	-0.184	0.100	0.068	0.024
Freedom of Speech	-0.377	0.110	0.001	0.002
Trust in Institutions				
Trust in President	-0.320	0.166	0.053	0.092
Trust in Electoral Commission	-0.167	0.082	0.043	0.092
Trust in Ruling Party	-0.234	0.164	0.154	0.092
Trust in Traditional Leaders	-0.122	0.066	0.063	0.092

Notes: Results on all individual questions from the Afrobarometer survey. P-values run individually in column (3), p-values adjusted for Multiple Hypotheses Testing in column (4).

Table C.2: Variables and questions in the Afrobarometer
Household Wealth

Variable	Question Number	Question Text
Country Condition	q4a	In general, how would you describe: The present economic condition of this country?
Household Condition, Compared	q5	How do you rate your living conditions compared to those of other people in your country?
Household Condition	q4b	In general, how would you describe: Your own present living conditions?
Household Condition, Past	q6	Looking back, how do you rate economic conditions in this country compared to twelve months ago?
Days with Food	q8a	Over the past year, how often, if ever, have you or anyone in your family: Gone without enough food to eat?
Days with Water	q8b	Over the past year, how often, if ever, have you or anyone in your family: Gone without enough clean water for home use?
Days with Cash	q8e	Over the past year, how often, if ever, have you or anyone in your family: Gone without a cash income?
Days with Fuel	q8d	Over the past year, how often, if ever, have you or anyone in your family: Gone without enough fuel to cook your food?

Notes: The table lists the individual questions included in Table C.1 under the category *Household wealth*. All variables are recoded such that higher values imply higher wealth. Question number and text refer to the Afrobarometer Round 7 questionnaire.

Table C.3: Variables and questions in the Afrobarometer
Satisfaction with Democracy

Variable	Question Number	Question Text
Status of Democracy	q35	In your opinion, how much of a democracy is your country today?
Satisfied with Democracy	q36	Overall, how satisfied are you with the way democracy works in your country?
Freedom of Speech	q42a	In your opinion, how often, in this country: Do people have to be careful of what they say about politics

Notes: The table lists the individual questions included in Table C.1 under the category *Satisfaction with Democracy*. All variables are recoded such that higher values imply higher satisfaction with democracy. Question number and text refer to the Afrobarometer Round 7 questionnaire.

Table C.4: Variables and questions in the Afrobarometer
Trust in Institutions

Variable	Question Number	Question Text
Trust in President	q43a	How much do you trust each of the following, or haven't you heard enough about them to say: The President?
Trust in Electoral Commission	q43c	How much do you trust each of the following, or haven't you heard enough about them to say: Electoral commission
Trust in Ruling Party	q43e	How much do you trust each of the following, or haven't you heard enough about them to say: The Ruling Party?
Trust in Traditional Leaders	q43j	How much do you trust each of the following, or haven't you heard enough about them to say: Traditional leaders?

Notes: The table lists the individual questions included in Table C.1 under the category *Trust in Institutions*. All variables are recoded such that higher values imply higher trust in institutions. Question number and text refer to the Afrobarometer Round 7 questionnaire.

D Replicating Results using the DHS

This section replicates the main results from the Afrobarometer using DHS household level data 1991-2018 from 28 countries. Main effect goes through as expected, while the intereaction effect is of expected sign but marginally insignificant.

Table D.1: Trade exposure and wealth gains
IV results in the DHS

	OLS		Cross-Border		Leave-One-Out	
	(1) OLS	(2) OLS	(3) IV	(4) RF	(5) IV	(6) RF
Trade Exposure	-0.061*	-0.053	-0.070**		-0.087**	
	(0.034)	(0.033)	(0.035)		(0.040)	
Trade Exposure × In Power		0.027				
		(0.018)				
Predicted Trade Exposure				-0.036*		-0.102**
				(0.020)		(0.051)
Predicted Trade Exposure × In Power				0.030		0.024
				(0.019)		(0.019)
Country × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country × ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Reduced form				Yes		Yes
Observations	250,144	249,186	246,866	246,866	249,186	249,186
First Stage F-Test			115.156		144.785	

In this table, I show how trade exposure impacts the distribution of wealth as measured by the DHS. *Trade Exposure* is defined as realized trade flows to all African countries aggregated to the country-by-year level and interacted with the population share of ethnicity e: $\sum_{d \in D} Export_{c,d,t} \times PopulationShare_{e,c}$. In columns (3) and (5) it is instrumented by *Predicted Trade Exposure* using either the Border instrument or the Leave-One-Out instrument. *Fraction lid* is calculated as the fraction of pixels not zero, *Average nighttime luminosity* as the log of average luminosity in each country-ethnic group observation plus one. *Country × year fixed effects* account for unobserved characteristics varying at the country and year level, including total trade flows, GDP, and population. *Country × ethnicity fixed effects* for the size and impact of ethnicity e in country c. The first stage F statistic is given in the last row. Corrected F-Statistics at the country-year level presented in Figure 3. Wealth gains are measured as a standardized index of the following variables in the DHS: Working, whether your house has electricity, and the number of sleeping rooms in your house. Significance denoted by standard errors clustered by country and ethnicity: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Technical Appendix

In this section, I derive a model of international trade with firm and ethnic heterogeneity to provide a motivation for the main estimation equation (5). My framework draws on Chaney (2008) and nests the standard model while remaining tractable.

The economy consists of N countries which contain a subset $e \in E$ of predefined ethnic groups. Not every ethnicity is present in every country. Furthermore, every economy produces a homogeneous composite good q_0, as well as horizontally differentiated goods $q(\omega)$. Any firm of ethnicity $e \in E$ producing a heterogeneous good $\omega \in \Omega$ from country $i \in N$, uses its ethnic counterpart $e' \in E$ in country $j \in N$ to maximize the expected profits from selling in market $j \in N$ according to:

$$\pi_{ij,ee'}(\omega) = p_{ij}(\omega)q_{ij}(\omega) - c_{ij,ee'}(\omega) \tag{6}$$

Where the price of a good $p_{ij}(\omega)$ is country specific, as is the demand for a good $q_{ij}(\omega)$.[18] $\tau_{ij} > 1$ represent variable trade costs, denoted as "iceberg trade costs". A firm needs to produce τ_{ij} goods in order to sell one unit in country j. The cost of producing a good $c_{ij,ee'}(\omega)$ is assumed to be ethnic dependent in home e and foreign e' and of the form:

$$c_{ij,ee'}(\omega) = \frac{\tau_{ij}}{\varphi} q_{ij}(\omega) + \left(\frac{L_{j,e'}}{L_j}\right)^{-\eta} f_{ij} \tag{7}$$

Here, φ denotes productivity which every firm draws from a Pareto distribution $G(\varphi) = 1 - \varphi^{-\gamma}$.[19] γ represents the degree of firm heterogeneity, with increasing values denoting decreasing firm heterogeneity. Firms learn about their productivity when drawing from $G(\varphi)$ and, subsequently, decide to pay country pair specific fixed costs f_{ij} in order to serve market j.[20] These fixed costs are mitigated by the fraction of the population in country j that is of the same ethnicity $e' = e \in E$ as the owner of the firm.[21] I call the effect of

[18] Although Aker et al. (2014) show that ethnic groups affect the prices between two countries, I assume that this is a result of a supply or demand shock. However, including a demand shock here would create a simple demand shift in the gravitiy equation. Alternatively, one could divide the product space into goods consumed by ethnic groups which would yield a result similar to including different sectors.

[19] Following the literature standard I use the Pareto distribution as it mirrors the empirical distributions well (Axtell, 2001) and is notational convenient.

[20] The cost of producing a good are wages times $c_{ij,ee'}(\omega)$. Due to the production in the freely traded homogeneous good q_0 wages in both sectors are normalized to unity to simplify the expressions. Furthermore, since there are infinitely many possible firms of each ethnicity, I can characterize the costs of producing variety ω simply by the ethnicity and the productivity of the firm φ.

[21] A similar approach has been undertaken by Krautheim (2012) where the fraction is the number of domestic firms active in the destination market. In the following, I assume that every ethnicity has at least one member in every country. I can relax this assumption and assume that there is an additional fixed cost

the fraction $\left(\frac{L_{j,e'}}{L_j}\right)^{-\eta}$ the network effect of ethnic ties. This fraction lies within the unit interval and raised to the power of $\eta \in \left[0, \frac{\sigma-1}{\gamma}\right)$ that gives the importance of ethnic networks in decreasing the fixed costs of exporting. It can be interpreted as a decreased costs of acquiring information about the market structure in the destination country or market demand. Alternatively, its interpretation permits lower payments to government officials because of ethnic ties or it serves a proxy for the general trust-worthiness of a society. Empirical evidence by Grossman et al. (2006) suggests that factors like cultural distance and institutional development are particular relevant for the fixed cost of exporting. Ethnic networks should then be beneficial when firms try to circumvent bureaucratic hurdles. The larger the hurdles, the larger should be the impact of ethnic networks.

In every country, households maximize their utility according to:

$$U = q_0^{1-\mu} \left(\int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}\mu} \tag{8}$$

That is, they consume a freely traded homogeneous good q_0 and consume every available variety of the heterogeneous good ω. The share of income spent on the heterogeneous good is given by μ and the elasticity of substitution is given by $\sigma > 1$. Standard results lead to a pricing of $p_{ij}(\varphi) = \frac{\sigma}{\sigma-1}\frac{\tau_{ij}}{\varphi}$ and a demand:

$$q_{ij}(\varphi) = p_{ij}(\varphi)^{-\sigma} P_j^{\sigma-1} \mu \left(1 + \frac{\Pi}{L}\right) L_j. \tag{9}$$

Here, $\left(1 + \frac{\Pi}{L}\right) L_j$ denotes the fraction of world capital Π and labor L income that belongs to country j.[22] Hereof, a fraction μ is spend on heterogeneous goods. Combining the profit function, pricing and demand yield the ethnicity dependent productivity cutoff above which firms start to export due to non-negative profits $\pi_{ij,ee'} \geq 0$:

$$\varphi^*_{ij,ee'} = \left(\frac{\sigma}{\sigma-1}\right) \frac{\tau_{ij}}{P_j} \left[\frac{\mu}{\sigma}\left(1 + \frac{\Pi}{L}\right) L_j\right]^{\frac{1}{1-\sigma}} \left(\frac{L_{j,e'}}{L_j}\right)^{\frac{\eta}{1-\sigma}} f_{ij}^{\frac{1}{\sigma-1}} \tag{10}$$

The price index P_j can be solved explicitly by summing all prices from all exporting countries together, taking their productivity cutoffs into account.[23] Then, the productivity cutoff can be expressed in terms of primitives:

to pay when dealing with non co-ethnic members. The results are robust.

[22]Due to the sector that produces the homogeneous goods, wages are driven down to unity.

[23]$P_j = \left(\sum_{k=1}^{N} L_k \sum_{e \in E} \int_{\varphi^*_{kj,ee'}}^{\infty} \left(\frac{\sigma}{\sigma-1}\frac{\tau_{kj}}{\varphi}\right)^{1-\sigma} dG(\varphi) \right)^{\frac{1}{1-\sigma}}$.

$$\varphi^*_{ij,ee'} = \left[\frac{\gamma}{\gamma-(\sigma-1)}\right]^{\frac{1}{\gamma}} \left[\frac{\mu}{\sigma}\left(1+\frac{\Pi}{L}\right)\right]^{-\frac{1}{\gamma}} L_j^{\frac{\eta-1}{\gamma}} \frac{\tau_{ij}}{\theta_j} f_{ij}^{\frac{1}{\sigma-1}} (L_{j,e'})^{\frac{\eta}{1-\sigma}} \tag{11}$$

As in Chaney (2008), the total foreign population decreases the cutoff due to market size effects $L_j^{\frac{\eta-1}{\gamma}}$. This effect is dampened by $\frac{\eta}{\gamma}$ because the ethnic population has a stronger effect on the cutoff than the total population.[24] θ denotes the multilateral resistance term that approximates how distant a market is in comparison to all other markets.[25] Equation (11) suggests that much of the ethnic network effect will work through the extensive margin of trade. If the fixed costs of exporting are higher due to corruption, the cutoff for ethnically connected and non-connected firms increases, but to a lesser extent for the former group.[26]

In order to obtain a testable equation, I aggregate individual demand[27] to an network extended gravity equation:

$$X_{ij} = \mu\left(1+\frac{\Pi}{L}\right) L_j f_{ij}^{\frac{\sigma-1-\gamma}{\sigma-1}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{-\gamma} \sum_{e\in E_i\cap E_j} L_{i,e}(L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}} \tag{12}$$

Total exports between any pair of countries increase in market size $\mu\left(1+\frac{\Pi}{L}\right)L_j$ and multilateral resistance θ and decrease in variable trade cost τ_{ij} and fixed costs f_{ij}. The network term is increasing the total trade flows since $\nu \equiv \frac{\eta(\sigma-1-\gamma)}{1-\sigma} \in [0,1)$ in order to obtain interior solutions for the system of equations.[28] If the number of ethnic groups is greater then the number of countries, the system of equations is under-identified and individual parameters in ν cannot be identified. A way around is to assume specific

[24]The original cutoff in Chaney (2008) can be recovered by setting $\eta = 0$. The effect of the foreign ethnic population is greater since $\frac{\eta}{\gamma} < \frac{\eta}{\sigma-1}$ due to the assumption $\gamma > \sigma - 1$ that guarantees interior solutions.

[25]$\theta_j = \left[\sum_{k=1}^{N} f_{kj}^{\frac{\sigma-1-\gamma}{\sigma-1}} \tau_{kj}^{-\gamma} \sum_{e\in E} L_{k,e}(L_{j,e})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}\right]^{-\frac{1}{\gamma}}$. A popular example is the comparison between Portugal and Spain with New Zealand and Australia. Similar in terms of GDP, the latter trade relatively more with each other due to their distance to all other markets in the world.

[26]Putting it differently, in a world where all the fixed cost consist of corruption and trust, the ethnic networks are paramount to exporting. We should observe only ethnically connected firms. A similar exercise can be done by changing the cost function into a part which is ethnic dependent (trust and corruption) and a part that is non ethnic dependent. Then ethnic networks do not matter when there is no ethnic dependent fixed costs, but matter a lot when there is no non ethnic dependent fixed cost.

[27]$X_{ij} = L_i \sum_{e'=e\in E} \frac{L_{i,e}}{L_i} \int_{\varphi^*_{ij,ee'}}^{\infty} dG(\varphi)$, where $\frac{L_{i,e}}{L_i}$ is the ethnic fraction in country i. An alternative summation would be to include the non ethnic population in foreign and their cutoffs: $X_{ij} = L_i\left[\sum_{e\in E_i\cap E_j} \frac{L_{i,e}}{L_i} \int_{\varphi^*_{ij,ee'}}^{\infty} dG(\varphi) + \sum_{e'\neq E_i\cap E_j} \frac{L_{i,e}}{L_i} \int_{\varphi^*_{ij,ee'}}^{\infty} dG(\varphi)\right]$. The second term would be condensed to the part in Chaney (2008).

[28]I further require that $\gamma > (\sigma-1)$ and $\eta < \frac{(\sigma-1)}{\gamma}$ to guarantee an interior solution.

values for ν and conduct sensitivity analyses. Specifically, if ν takes on the value one, the ethnic network variable leads to a search and matching interpretation and gives the likelihood that two randomly selected firms from both countries are of the same ethnicity, when controlling for population size.

The introduction of ethnic heterogeneity in the framework of Melitz (2003) and Chaney (2008) introduced a second source of heterogeneity that creates a particular feature regarding export decisions. Firms owned by an ethnic minority might first export to other markets and only later serve their home market. This feature is similar to capital-constraint firms that cannot export in Chaney (2016) and implies imperfect selection into exporting. Firms that export might have lower productivity than firms that do not and, thus, create wealth losses.

The empirical equivalent of this equation is given by:

$$\log(X_{ij,t}) = \beta \log \left(\sum_{e \in i \cap j}^{E} L_{i,e} (L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}} \right) + \Gamma_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t} \tag{13}$$

Since the importer and exporter fixed effect also capture population in each country and $(L_j \times L_i)^{-1} = -\log L_j - \log L_i$ one can rewrite the equation as:

$$\log(X_{ij,t}) = \beta \log \left(\sum_{e \in i \cap j}^{E} \frac{L_{i,e}}{L_i} \times \frac{(L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}}{L_j} \right) + \Gamma_{ij,t} + \delta_i + \delta_j + \varepsilon_{ij,t} \tag{14}$$

which as $\frac{\eta(\sigma-1-\gamma)}{1-\sigma} \to 1$ approaches equation (5). This equation can be interpreted as a search and matching model, where the population in the importing country has to incur a penalty, thus needs a larger population to have the same effect on trade as the exporting population.

E.1 Inter-ethnic Trade

So far I assumed that connections can only exist within ethnic groups and neglected the possibilities of inter-ethnic connections. Here, I relax this initial assumption and assume that every ethnicity has an implicit (weak) ranking of every other ethnicity. Then, for every ethnicity I can order the other ethnic groups according to the cost they have to incur in order to conduct business with them. This cost is similar to the fixed costs discussed earlier, in the sense that it reflects learning costs between ethnic groups. Therefore, I assume there exists a matrix $F_{E \times E}$ that reflects this ordering between every possible combination of ethnic groups. The cost of producing and exporting are then given by:

$$c_{ij,ee'}(\varphi) = \frac{\tau_{ij}}{\varphi} q_{ij}(\varphi) + \left(\frac{L_{j,e'}}{L_j}\right)^{-\eta} f_{ij} f_{ij,ee'} \tag{15}$$

with $f_{ij,ee'}$ being an element from $F_{E\times E}$. Here bilateral fixed costs are disentangled from ethnic specific cost. Every firm has to incur bilateral fixed costs to set up the firm, but also have to invest in ethnic relations in order to mitigate the additional ethnic specific fixed costs.[29] The gravity equation is then given by:

$$X_{ij} = L_j \mu \left(1 + \frac{\Pi}{L}\right) f_{ij}^{1-\frac{\gamma}{\sigma-1}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{-\gamma} \sum_{e \in E \cap E'} L_{i,e} (L_{j,e'})^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}} f_{ij,ee'}^{1-\frac{\gamma}{\sigma-1}} \tag{16}$$

Now, the effect of ethnic match probabilities is not only measured within ethnic groups, but also between ethnic groups. If the fixed costs of creating ties between ethnic groups are low enough, this specification should fit the data better. Combining the findings on the extensive margin formulation and the ethnic specific fixed costs, ethnic groups have a two fold effect on trade flows. They increase the number of firms exporting in distrustful environments by affecting the extensive margin. However, trade volumes between two countries are negatively affected by the ethnic specific fixed costs. Then if these fixed costs represent trust or corruption issues, the above model puts a strong emphasis on reducing corruption and increase trust among ethnic groups.

E.2 Empirical results from the theoretical Model

Table E.1 presents the empirical results to estimating the theoretical model on trade flows between all African countries using OLS in odd columns and PPML in even. The point estimates are not statistically different from the baseline empirical results in Table B.2, and allowing for inter-ethnic trade increases the importance of ethnic networks, as expected.

[29] The basic model is a special case of this case where the off diagonal elements of $F_{E\times E}$ are assumed to be so high that only within ethnicity connections can occur.

Table E.1: Ethnic connections and trade flows
log(Exports) between all countries

	OLS	PPML	OLS	PPML	OLS	PPML
	(1)	(2)	(3)	(4)	(5)	(6)
log(Size of Network)	0.130^{***}	0.085^{***}				
	(0.024)	(0.024)				
	[0.033]	[0.027]				
log(Size of Network, exponent=0.2)			0.197^{***}	0.131^{***}		
			(0.038)	(0.038)		
			[0.054]	[0.049]		
log(Size of Network, allowing for inter-ethnic networks)					0.859^{***}	1.282^{***}
					(0.156)	(0.178)
					[0.369]	[0.245]
Country-pair controls	Yes	Yes	Yes	Yes	Yes	Yes
Exporter and importer $\times$ year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37,983	38,271	37,983	38,271	37,983	38,271

In this table, I show that ethnic connections predict bilateral exports per capita between all countries using the models empirical equation. *log(Size of Network* is estimating equation (13) with the exponent on foreign ethnic population $L_{j,e'}^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}$ being set to one. *log(Size of Network* is estimating equation (13) with the exponent on foreign ethnic population $L_{j,e'}^{\frac{\eta(\sigma-1-\gamma)}{1-\sigma}}$ being set to 0.2 based on reasonable values for the elasticity of substitution σ, the pareto parameter γ and η based on the assumption that $\gamma > (\sigma - 1)$ and $\eta < {}^{(\sigma-1)}/\gamma$. *log(Size of Network, allowing for inter-ethnic networks* estimates equation (16) Sample consist of all countries in Africa and includes country and destination by year fixed effects in all regressions. The main dependent variable are the logarithm of bilateral exports per current capita in the years 1992–2018. The following country-pair controls are added in all columns: log length of the border, log distance between capitals, the number of ethnic groups shared between the countries, whether the countries share a colonial past or judicial language. Linguistic and genetic distance $\in [0, 1]$ captures the similarity between the countries (Spolaore and Wacziarg, 2015). Inference is made on the basis of standard errors clustered by each country-pair, as shown in parenthesis. Two-way clustered standard errors allowing for separate home- and foreign-clusters shown as robustness in brackets. OLS and PPML denote the estimation method. Significance denoted by standard errors clustered by the country pair: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$