The International Transmission of Democratic Values: Evidence from the African Diaspora *

Marco Manacorda† Jacopo Ponticelli‡ Andrea Tesei§

March 20, 2024

Abstract

This paper investigates the effect of rising anti-immigrant sentiment in European democratic host countries on support for democracy in African migrants' origin communities. We propose a novel methodology to estimate migration stocks from sub-national areas of origin to any possible country in the world based on Google searches for origin communities performed in destination countries. We use these data alongside measures of anti-immigrant sentiment at destination to study origin communities exposure to anti-immigrant sentiment. We show that a one standard deviation increase in exposure to anti-immigrant sentiment at destination leads to a decline of between 0.1-0.2 of a standard deviation in support for democracy at origin. We conclude that rising anti-immigrant sentiment in Western democracies has the potential to harm democratic progress in Africa.

Keywords: Africa; Migration; Social Remittances; Google Search.

* We received valuable comments from Eliana La Ferrara, Hilel Rapoport, David Yang and seminar participants at the 2022 ASSA Meeting, CEPR 2023 Symposium, Queen Mary University and Central European University. Edoardo Bollati, Gabriel Chaves-Bosch, Noah Farougi, Antonio Leon, Erick Molina and Jiaqi Zhang provided excellent research assistance. We gratefully acknowledge funding from Queen Mary University of London and Northwestern Kellogg School of Management.

† Queen Mary University of London, & CEPR; contact: m.manacorda@qmul.ac.uk.

‡ Northwestern University, NBER & CEPR; contact: jacopo.ponticelli@kellogg.northwestern.edu.

§ Queen Mary University of London & CEPR; contact: a.tesei@qmul.ac.uk.

1 Introduction

In countries with underdeveloped democratic institutions, the viability of democracy hinges significantly on the attitudes and beliefs of their citizens. With public support, democracy is legitimate and stable, while without such support it becomes vulnerable (Lipset, 1959; Easton, 1965). Understanding where public support for democracy originates is therefore crucial to assess the democratic prospects of countries with weak institutional foundations. A large body of literature in sociology and economics has argued that international migrants can be important agents of change and play a key role in shaping the democratic support in their countries of origin. Because they are typically exposed to more consolidated democratic systems, migrants from countries with weak institutional capacity may send back to their communities of origin so-called "social remittances", ideas and norms - including democratic values - learned at destination, which can foster democratic progress at home (Levitt, 1998; Spilimbergo, 2009; Barsbai et al., 2017).

A cursory look at the democratic progress made by weakly institutionalized countries over the recent past, however, is not immediately consistent with the idea of social remittances. Despite a twofold increase in South-North emigration since 2000, the democratic progress in the global South has stalled, if not reversed (Boese et al., 2022). In this paper, we thus re-examine the international transmission of democratic values, emphasizing an important aspect that has been overlooked by previous studies. Partly in response to increased migration flows, over the past twenty years Western democracies have grown more anti-immigrant and communitarian, adopting policies that favour insiders - natives and those who share prevailing cultural traits - at the expense of outsiders - foreigners, migrants and minorities (Halla et al., 2017; Dustmann et al., 2019; Guriev and Papaioannou, 2022; Manacorda et al., 2023). This suggests an alternative possibility: rather than embracing the democratic values at destination, migrants' exposure to democracies that are fully consolidated but incapable of promoting their integration, well-being and economic success may lead them to revise their expectations of the functioning and desirability of the democratic system, and report their negative experience to their communities of origin.

We pose two questions: do negative attitudes towards migrants and migration at destination impact support for democracy in origin communities? And is the process accelerated through access to information technologies at origin, which allow for greater information about the migrants' experience at destination? We focus of the African diaspora in Europe, which provides an ideal setting to investigate these questions. Africa is by and large the poorest continent in the world, with a relatively weak institutional capacity, while Europe is predominately composed of fully consolidated democracies, potentially representing an ideal democratic benchmark for migrants coming from weakly institutionalized systems. African migration to Europe has more than doubled since 2000 to account for 27 percent of total migration from the continent, with more than 11 million African citizens, or 1 percent of the total population of origin, residing in Europe.

The main empirical challenge to identify this transmission of values from migrants to their communities of origin is the absence of systematic information on sub-national migration stocks. This data limitation arises from the fact that, once migrants are observed at destination, typically no information is collected on their precise place of origin. Census data in destination countries - by far the most reliable sources of information on immigration - collect at best migrants' country of origin. When available, Census data in origin countries typically fail to measure those who left, let aside their destination. This problem is more acute for developing countries, the major net donors of international migrants. Because the data only allow to identify outmigrants from a country irrespective of their precise sub-national place of origin, this implies that measures of exposure to outmigration will be at the national level, possibly over time. This poses clear challenges for the identification of the effect of outmigration on origin communities' outcomes.

The first contribution of this paper is to propose a novel methodology to recover estimates of the migration stock from each origin sub-national region of Africa to each destination country in the world. Our methodology relies on data on Internet searches for the names of African sub-national regions performed in destination countries, that we recover from Google Trends. Search engine data have proved useful in predicting economic and financial activity (Choi and Varian, 2012), but, to the best of our knowledge, have not been used to identify migration flows. Building on a literature in psychology and human geography on individuals' attachment to their places of origin (Tuan, 1990; Hernández et al., 2007), we propose to capture the presence of migrants from a specific region of Africa using Internet searches for region-specific terms at destination.[1] According to this logic, for instance, a greater number of searches recorded in Italy for the term "Kumasi" – the capital city of the Ashanti region in Ghana – relative to searches for the same term in France, would indicate a comparatively higher presence of Ashanti migrants in Italy than in France.

As Google searches might be capturing the popularity of an origin region, irrespective of the number of migrants, we use data from countries with no stock of migrants to purge our estimates of such effects. Because Google Trends only provide information on *relative* searches

[1] Although others have used Internet generated data, and in particular Facebook network data (Spyratos et al., 2019), to estimate bilateral migration stocks across fine geographies in the USA and Europe, these data are unlikely to provide reliable estimates of migration stocks for Africa, as Facebook usage in the continent is still very limited (with an estimated number of users as of 2020 of 25.4 million relative to a population of over 1.2 billion, i.e., a usage rate of 2%), and highly selected in terms of countries, urban and socioeconomic status. Böhme et al. (2020) use Google searches to predict migration intentions.

across terms or locations, we use between-country-of destination variation in Google searches for origin countries' regions together with aggregate data on bilateral migration stocks across country dyads from the United Nations to derive best linear predictions of migration based on Google searches. We exploit within-country-of-destination Google searches across country of origin regions together with such aggregate data to re-apportion migrants to origin regions. This allows us to provide novel estimates of migration stocks from 709 origin regions within Africa to 133 countries. Of course, our methodology has the potential to be applied to other origin countries and hence its purpose goes beyond the specific estimates and application in this paper.

We validate our methodology using the scarce and highly sparse available data on migration from Africa that records information on the sub-national region of origin of migrants. The majority of this data effectively captures return migration – and not current migration status – using 1% samples of national Censuses for Africa. We also collect information on regions of origin of current migrants from small-scale, specific-purpose surveys. We document a clear positive correlation between our estimated sub-national migration stocks from Africa and the measures constructed with these auxiliary data, which lends support to the validity of our methodology.

Next, we use our estimates of migration stocks at sub-national level to study the role of social remittances from African migrants living in Europe to their communities of origin in Africa. We are particularly interested in studying whether changes in attitudes towards migrants in European destination countries – which are mostly consolidated democracies – affect the view of democracy in the migrants' communities of origin. Migrants might update their view about the desirability of the democratic system depending on how they are treated in European democracies, and then share their experiences with their communities of origin. To test this hypothesis, we exploit variation in exposure to anti-immigrant sentiment at destination across sub-national regions of origin throughout Africa. Exposure is based on differential baseline migration patterns in destination countries across origin regions and the time variation in attitudes towards migration across destination countries. In focusing on within country of origin variation, our approach purges the estimates of potential confounding factors that simultaneously affect migration and citizens' view on democracy in origin countries.

We estimate a set of individual level regressions linking support for democracy recorded in the communities of origin of African migrants with the exposure of such communities to anti-immigrant sentiment in destination countries. We use data from the Afrobarometer (rounds 3 to 8), which cover about 250,000 respondents in 39 African countries in the years between 2005 and 2021. We investigate respondents' support for democracy, along with their preferences for the rule of law, electoral competition, and civic participation.

We use two sets of measures to capture changes in attitudes towards migrants at destination. First, we use data on electoral outcomes in national parliamentary elections, combined with data on party platforms from the Chapel Hill Expert Survey (CHES, Jolly et al., 2022), to identify political support for party policies and ideologies that promote nationalistic policy platforms and intolerance towards immigrants and minorities. In particular, we use the CHES data to build a number of measures that relate to support for a closed society, including supporting a restrictive immigration policy, opposing ethnic and minority rights and advocating for migrants' cultural assimilation. Second, we use data from the European Social Survey (ESS) to build country-year level measure of public support for anti-immigrant positions, as well as experience of migrants' discrimination. Our measure of exposure for each African region of origin is the weighted average of anti-immigrant sentiment across destination countries where the community has sent its migrants, where the weights capture the intensity of the migration link.

One potential concern with our empirical strategy is that changes in anti-immigrant sentiment at destination might be correlated with other shocks at destination, which are also transmitted to the communities of origin via the migrant network. For example, anti-immigrant sentiment might intensify during a recession, and the effect of a recession might be felt also in the communities of origin via lower economic remittances. To disentangle the confounding effect of economic remittances, we include in our estimating equation a measure of exposure to changes in economic conditions at destination via the migration network.

Our main empirical finding is that individuals in sub-national regions of origin exposed to an increase in anti-immigrant sentiment at destination experience a relative decline in their support for democracy. The estimates indicate that individuals at origin with a standard deviation higher exposure to anti-immigrant sentiment at destination experience between 0.1 and 0.2 of a standard deviation decline in their probability of supporting democracy. This decline extends to their endorsement of multi-party competitive elections, engagement in voting, and participation in civic activities. We find consistent effects when using alternative measures of anti-immigrant sentiment at destination, using data from both CHES and ESS.

We also study whether variation in the ability to communicate with destination regions affect the strength of social remittances. To this end, we match Afrobaromenter respondents' location with fine geographical data on the diffusion of 3G and 4G mobile phone coverage. This allows us to compare the impact of social remittances on individuals that have different access to the mobile phone network but that reside within the same community of origin in Africa. We find that, for a given exposure to anti-immigrant sentiment, individuals with access to mobile phone coverage experience a larger decline in their support for democracy relative to individuals in the same region but without access to mobile coverage.

Our paper is related to several strands of research. First, we contribute to the literature on migration and culture. A number of influential studies document how migrants maintain a strong connection with their culture of origin, measured in terms of level of trust (Guiso et al., 2004; Algan and Cahuc, 2010), gender attitudes (Fernández and Fogli, 2009) or propensity to violence (Couttenier et al., 2019). Recent work also shows that the presence of immigrants may influence natives' political preferences and ideology in the destination country (Giuliano and Tabellini, 2020) and that opinion changes in the countries of origin spillover to immigrants abroad, potentially with large social and political implications for the host countries (Yarkin, 2023).

More directly related to our paper is the body of work investigating social remittances (Levitt, 1998), defined as the transfer of ideas, behaviors, and norms from migrants in host countries back to their home communities. A number of studies in this area document how exposure to democratic values at destination might promote democratic progress (Spilimbergo, 2009; Docquier et al., 2016; Mercier, 2016; Barsbai et al., 2017), political participation (Chauvet and Mercier, 2014) and demand for political change (Batista and Vicente, 2011; Karadja and Prawitz, 2019) at home, either via transmission of information to home communities or via return migration. In contrast to these studies, we emphasize how negative migrants' experiences in host countries may lead to revised expectations about the desirability of democracy, which may spillover to their communities of origin. Additionally, we enhance previous research by presenting continent-wide findings with a high degree of geographical detail. Existing studies either focus on cross-country analyses (e.g., Spilimbergo, 2009), which may not address the potentially endogenous emigration from countries with diverse domestic trends in democratic support, or utilize finer variation within a single country (e.g., Barsbai et al., 2017), which offers better identification but has limited external validity. We overcome these challenges by leveraging sub-national variations in democratic support across the entire African continent and by exploiting differences in anti-immigrant sentiment among destination countries, which is arguably exogenous to migration patterns from specific regions of origin.

A related strand of literature argues that support for democracy increases with the length of time spent under the system (Persson and Tabellini, 2009; Fuchs-Schündeln and Schündeln, 2015), particularly when democratic periods coincide with economic growth, peace and stability (Acemoglu et al., 2021). In contrast to these studies that center on the experiences of national citizens, our focus is on international migrants. Their encounter with the democratic system may be less satisfactory, especially in cases where the host country is unable to integrate them and provide for their economic success and well-being.

Our results are also connected to existing research showing that information and communication technologies may accelerate the process of change in slow-moving social norms on various fronts, such as fertility choices (Billari et al., 2020) and attitudes towards gender roles (Jensen and Oster, 2009). Recent evidence, however, also suggests that exposure to cultural values perceived as antagonistic to local traditions may lead to communities' backlash, as exemplified by the resistance to the Westernization of radio programming during the liberalization of the radio market in Pakistan (Blumenstock et al., 2022).

Finally, our measure of sub-national migration stocks is based on Google searches (GT) performed in destination countries. GT has been used in a number of applications (e.g., Choi and Varian, 2012; Nuti et al., 2014), including to predict migration intentions (Böhme et al., 2020). While these studies utilize GT to forecast local events, our approach involves key terms associated with foreign places, namely sub-national regions of origin across Africa. In this regard, our measure is similar in spirit to Burchardi et al. (2019), who construct a measure of information demand about foreign countries across US counties using data from Google Internet searches. They show that this measure reflects the composition of local ancestry and use it to study the impact of ancestral ties on foreign direct investment sent and received by local firms. Given that migrants likely have a more direct connection with their place of origin than one mediated through ancestry, this suggests a potentially strong predictive content of searches for places of origin in our context.

The rest of the paper is organized as follows. In section 2 we describe the data used to measure shocks at destination and show the impact they have on migrants' experience. To map these shocks into consequences at origin, in section 3 we introduce our methodology to estimate migrant stocks at sub-national level using data on Internet searches from Google Trends. In the same section we also present a set of validation tests of our estimates based on existing data on sub-national migration from Africa to Europe, as well as data on return migration. In section 4 we introduce the Afrobarometer data and the definition of the outcome variables used, and present the empirical specifications that we use to study the international transmission of support for democracy via migrant networks. In section 5 we discuss the results from the analysis and in section 6 we present our conclusions.

2 Attitudes towards immigrants at destination and their effects on their view of democracy

Before studying how attitudes towards migrants in European destination countries affect the view of democracy in their communities of origin in Africa, we provide some evidence of how such attitudes may influence migrants' own views of democracy. To this end, we combine data from the European Social Survey (ESS), which provides information on respondents' socioeconomic characteristics, place of birth, and perceptions of discrimination and satisfaction with democratic institutions, with data from the Chapel Hill Expert Survey, which allows us to characterize the political climate across European countries of destination. Specifically, the CHES data classify the policy positions and ideologies of European parties over time. We focus on five measures that generally relate to support for a closed society that favors insiders (the nation, the native-born and those sharing prevalent cultural traits) at the expense of outsiders (supranational institutions, foreigners, migrants and minorities).[2] In particular, we focus on parties' positions in terms of support for (1) restrictive immigration policies; (2) migrants' integration as opposed to multiculturalism; (3) and ethnic minority rights. We also consider (4) a traditional measure of the parties' left-right ideological orientation and (5) the parties' positions on the divide between universalism (support for open borders, individual and minority rights and acceptance of global authorities) and communitarianism (support for traditional values, defense of the national community and support for the sovereignty of states). We rescale all CHES variables so that higher values correspond to greater support for a closed society. We use data on electoral outcomes of national parliamentary elections to calculate a weighted average of each CHES variable at the country and election year level, with weights given by the fraction of votes each party received.

Figure 1 shows the evolution of political support for positions in favor of a closed society over the period 2005-2021 on the European continent and in the 19 major European countries separately.[3] The measures are normalized to 100 at the beginning of the period. Three main observations emerge. First, as previous studies have documented (e.g. Manacorda et al., 2023), European politics has shifted markedly toward more closed positions over the past 20 years, with an increase of between 7 and 19 percent, depending on the measure used. Second, there is considerable heterogeneity across countries and over time: while voters in most countries have been more supportive of closed-society platforms, in other countries, notably Spain, Portugal, and Greece, political support has shifted toward more universalist parties. Finally, even within countries there is considerable variation across ideological dimensions, as exemplified by France and Italy, which have both become more liberal in terms of support for civil rights while becoming more conservative on the dimensions of immigration policy and support for ethnic minorities. We exploit the high degree of political variation among European countries to identify the effect of such shocks on migrants' experience at destination on support for democracy in their communities of origin.

[2] The CHES database is based on experts' assessment of parties' platforms and ideologies and it covers the majority of European parties, providing a consistent source across space and time. We use CHES data from waves 2006, 2010, 2014 and 2017.

[3] These countries account for around 450 million people and ninety-six percent of the EU27 population.

In Table 1 we turn to investigate the impact of support for a closed society on perceived discrimination and satisfaction with democracy among ESS respondents. We estimate the following regression:

$$y_{ir_ct} = \alpha_r + \alpha_{ct} + \beta_1 Non\text{-}EU_i + \beta_2 Anti\text{-}Imm_{ct} * Non\text{-}EU_i + \gamma X_{ir_ct} + \varepsilon_{ir_ct} \qquad (1)$$

where $Anti\text{-}Imm$ refers to one of the different dimensions of a closed society discussed above (one for each column of Table 1) and $Non\text{-}EU$ is a dummy equal to one for respondents born outside the EU. The specification also includes individual characteristics (age, gender, education level, employment status, a dummy capturing whether the respondent lives in a large urban center and a dummy for country of birth), along with fixed effects at the regional NUTS2 and country X year level. The main coefficient of interest is on the interaction term, capturing the differential response to closed society politics among Non-EU immigrants compared to natives.

The outcome variable in Panel A of Table 1 is a dummy equal to 1 if a respondent perceives their group as being discriminated against on the basis of their race, country of origin, ethnicity, religion or language. The β_1 coefficients indicate that migrants are on average more discriminated than natives, while the β_2 coefficients on the interaction term show that this differential increases as politics in destination countries turn more communitarian and political support for a closed society increases. In panel B, the outcome variables captures satisfaction with how democracy works in the destination country on a scale from 0 to 1. The negative coefficient on the interaction term indicates that, compared to natives, migrants express reduced satisfaction with democracy as support for a closed society in their host country rises. Overall, the results in Table 1 suggest that shock in local politics at destination are consequential for migrants' experience. Not surprisingly, migrants feel more likely to be discriminated against as host countries become more closed and, relatedly, their level of support for the way the democratic regime works in practice declines. Interestingly, the effect appears to be specific to non-EU migrants. Appendix Table A.1 reports results from the same analysis for EU migrants, who do not appear do be more discriminated against, nor less satisfied with how democracy works as host countries become more communitarian.

Having shown that non-EU migrants are less satisfied with democracy in the host countries as these become less welcoming towards them, in Appendix Table A.2 we present descriptive evidence suggesting the great potential for this information to travel all the way to migrants' community of origins. We use data from the EU-commissioned Migration from Africa to Europe (MAFE) and from Oxford EUmagine, two small-scale migration surveys that provide information, among other things, on the frequency of contact with relatives abroad among about 6,000 respondents from African countries. Participation in the surveys is conditional on having at least one relative residing in a foreign country (not necessarily in Europe, in the case of the EUmagine survey). The upper panel in the table shows that the vast majority of respondents have been in contact at least once over the previous year, while the lower panel shows that the contact is quite frequent: over 70 percent of respondents are in contact at least once a month, while one third is in contact every week.

In the next sections, we turn to investigate whether migrants' disgruntlement with democratic systems at destination that are unwilling to integrate them, translate into lower support for democracy at origin. To do so, we need to map migrants at destination with their communities of origin. We formally turn to this task in the next section, where we introduce our measure of predicted sub-national migration stocks using GT.

3 Estimating migration stocks from Google Trends data

3.1 Methodology

In this section we provide details on how we derive estimates of migration stocks from subnational areas of origin to each destination country based on Google Trends (henceforth GT) data. These data reflect "trending topics" based on searches people make on Google every day (as well as irregular search activity, such as automated searches or queries). A formal discussion is presented in Appendix A.

Our object is of interest is the distribution of migrants from each subnational origin region in Africa across destination countries. We proxy this distribution using GT data about searches performed in each country (other than the African country of interest) for the name of the most populous city in each African region, which is typically the region administrative capital.[4] Intuitively, say, if more searches are performed for "Accra", the capital city of the homonymous region of Ghana, in the UK relative to Switzerland, then this suggests that migrants from the Accra region are more likely to be located in the UK relative to Switzerland. The approach is predicated on the assumption that migrants from a certain place of origin are more likely to perform searches for that place than non-migrants or migrants from other areas and countries. Despite its intuitive appeal, there are some main challenges associated to this approach.

First, the relationship between Google searches and number of migrants needs not to be one to one. This will depend on the intensity by which migrants perform such searches

[4] This is likely to be a good predictor of migrants' interest for that region, either because migrants from the region disproportionately come from that city or because they are likely to have a specific interest in it (due to politics, sports etc.).

relative to natives. The problem is very similar in spirit to Henderson et al. (2012) who use nightlight density to predict GDP across granular geographical areas in Africa. These authors exploit the aggregate relationship between GDP and nightlights intensity across countries to derive the best linear prediction of the former based on the latter and hence estimates of local GDP based on local nightlight intensity. We follow a similar approach. In particular, we use searches for different origin countries in Africa performed across different destination countries together data on the total number of migrants from each country in Africa to each destination country to derive the best linear prediction coefficient of migration conditional on Google searches. In particular, let M_{r_od} denote migrants stocks from origin region r of country o to destination country d, S_{r_od} denote Google searches for region r_o performed in country d (relative to all searches in country d), $M_{od} = \Sigma_{r\in o} M_{rod}$ denote total migrants stocks from country o to country d, $\tilde{s}_{od} = \frac{\Sigma_{r\in o} s_{r_od}}{N_o}$ denotes average log searches from all regions of country o performed in country d, where N_o is the number of regions in country o and $x = ln(1 + X)$. One can obtain the best linear prediction coefficient of migration conditional on Google searches based on a regression of log bilateral migration stocks at the national level, m_{od}, on average log searches for regions of country o performed in d, $\tilde{s}_{od}$, plus country of origin and country of destination fixed effects:

$$m_{od} = \psi\tilde{s}_{od} + \delta_o + \delta_d + e_{od} \tag{2}$$

One limitation of the GT data is that these do not provide the *absolute* volume of searches for a certain term but only *relative* searches. In particular, GT data allow us to recover either within-country-of-destination relative searches, i.e. relative searches performed in a given country across two regions (e.g., searches in the UK for the term "Kumasi", the regional capital of the Ashanti region in Ghana, relative to searches, say, for "Accra") or between-country-of-destination relative searches, i.e. searches for the same region ("Accra") across two countries (say the UK and Switzerland), where the latter are further expressed as the ratio of total Google searches in the two countries (UK vis a vis Switzerland) in order to avoid that these are affected by scale effects.

This means that we cannot estimate equation (2) directly. One can however still use information on between-country-of-destination relative searches to estimate the BLP coefficient. In particular, one can replace s_{od} in the previous regression with $s_{od} - s_{oD}$, where D denotes a numeraire destination country (that we set arbitrarily to Switzerland). As the latter term is de facto a country of origin fixed effect, the rest of the model remains unchanged and our bilateral regression model is :

$$m_{od} = \psi(\tilde{s}_{od} - \tilde{s}_{oD}) + \delta_o + \delta_d + e_{od} \tag{3}$$

Second, searches for a certain regional capital city in Africa might reflect its popularity (e.g., if this is a tourist destination or an economic hub) rather than migration. As long as a place popularity is the same across destination countries, though, one can purge Google searches of this effect. In practice, we exploit relative searches for a certain region of origin (say Ashanti relative to Accra, in Ghana) performed in countries with no migrants from Ghana (say Uruguay, Indonesia, etc,), and hence a fortiori from the Ashanti and Accra regions, to identify the popularity of each regional capital relative to the country capital.

Third, and even once one has netted out these popularity effects, the fact that GT provides only relative rather than absolute searches implies that one will not be able to derive how many migrants from a certain African region, say Accra, reside in the UK but only how this number compares to the number of migrants from Ashanti. Effectively, we are short of one degree of freedom for identification. One can solve this problem, though, again exploiting national migration stocks together with within-country-of-destination Goole searches. Institutively, say, if searches for the city of Accra performed in the UK are twenty times as large as searches performed in the same country for each of the other nine other regions of Ghana, and if there are 100,000 migrants from Ghana to the UK, this implies that the number of migrants from Accra to the UK will be roughly 69,000 (=20/29 X 100,000). One can obtain similar estimates for each destination country and hence the overall distribution of migrants from Accra across destination countries.

A final consideration is that Google searches might reflect bilateral links between African regions and countries over and above migration. These might be due for example to commercial or colonial ties that relate specific areas in origin countries (say the country political or economic capital) to countries of destination. Of course, these will not be absorbed by popularity effects that are common across destination countries. In practice, as long as these bilateral links affect migration, this will not invalidate our approach. As a further check, we can include a variety of additional country of origin X country of destination specific variables in our aggregate bilateral aggregate regression (3). If the results are not affected by their inclusion, this provides further evidence in favor of our claim that GT data genuinely capture bilateral migration stocks.

3.2 Estimation results

In the following we present empirical estimates based on the procedure described above. For all 50 African countries in our sample, we start by identifying the most populous city in each admin1 sub-national region. In total, we identify 709 cities. Figure 2 reports a map of Africa with the precise location and name of each of these 709 cities, while Appendix Figure A.1 zooms in on Western Africa.

We perform separate GT queries for each of these 709 cities relative to the respective country capital across 133 destination countries. We also perform searches for each of these cities in each destination country relative to Switzerland. We average the resulting monthly searches across all months from January 2004 to December 2020. We complement these data with data on bilateral migration stocks from the United Nations Population division.[5]

We first turn to estimates of equation (3). Before presenting regression results and in order to add transparency to the estimates, for the purpose of illustration, we report a set of correlations between log migrant stocks from each African country to each destination country, m_{od}, and average log searches for the origin country's regions relative to Switzerland, s_{od}. We plot residuals of both variables relative to country of origin fixed effects based on population-at-destination weighted regressions. Appendix Figure A.2 reports the results for Algeria, Ghana, Morocco, and Nigeria - four large migrants' donor countries, two of which francophones and two anglophones. The results use data for 2015, although results using data for other years are very similar. Although regressions include all destination countries, we explicitly label observations corresponding to European countries. One can see a clear positive gradient between migrants stocks and Google searches: higher searches at destination are associated to higher outmigration. For example, searches for regions of Algeria and Morocco are higher in France, where large migrants communities from these countries reside, than in the UK. By converse, we see higher searches for countries such as Ghana and Nigeria in the United Kingdom, a preferred destination for such migrants, relative to France. This evidence is consistent with our hypothesis that GT are good proxies for migration stocks.[6]

A concern is that these estimates might be capturing common country of destination effects. This will happen if migrants disproportionately settle in countries with higher Google searches for African countries for reasons other than migration. To shed some light on this, in Appendix Figure A.4 we plot the two series across African origin countries separately for four large destination countries in Europe, namely France, Germany, Spain and the United Kingdom. Data are obtained as residuals from regressions on country of destination fixed effects and are weighted by population at origin. Again, there is a very clear positive gradient between searches and immigrants' stocks.[7]

[5] The data are available at `www.un.org/development/desa/pd/content/international-migrant-stock`).

[6] For completeness, Appendix Figure A.3 reports the same data, separately for each origin country in Africa. Larger dots correspond to larger countries (both in Europe and elsewhere). One can see that this relationships holds across most origin countries.

[7] For completeness, Appendix Figures A.5 to A.7 report the two series for all destination countries in the world. We report separate graphs for destination countries in Africa, Europe plus North America and Oceania, and Asia plus Latin America and the Caribbean (LAC). We focus only on destination countries with population as of 2015 greater than 1 million. Again, across most destination countries, there is a very clear positive gradient between searches and immigrants' stocks.

We now turn to estimates of equation (3) in Table 2. We present increasingly saturated models with different combinations of country of origin and country of destination fixed effects. Regressions are weighted by population at origin, although results are very similar if we weight by population at destination. We also present two-way clustered standard errors, by country of origin and destination. Results are only mildly sensitive to the specification. Point estimates for ψ vary between 1.504 and 1.287 and are statistically significant at conventional levels.

Appendix Table A.3 also reports coefficients separately across groups of destination countries, using the most saturated specification with the interaction of country of origin and country of destination fixed effects with time effects. Results hold true across most continents except the Americas.

As one might be concerned that such correlations capture the strength of bilateral ties between country pairs, which might be correlated with both migration stocks and Google searches, in Appendix Table A.4, we also present regressions where we control for a large array of country of origin times country of destination specific variables, including: common language, distance between capitals, (log) trade volume, past colonial links, and common legal origin. In the most demanding specification including all bilateral controls, the point estimate is still positive (about 0.781) and precisely estimated.

As explained in the previous section and more formally in Appendix A, we use the estimates of ψ together with relative Google searches to reapportion migrants from each to country to their respective region of origin. Before doing so, though, we purge relative searches from the relative popularity of each city relative to the country capital. In order to do so, we use data on country pairs for which migration stocks relative to the population at destination is zero or unreported by the United Nations, which is likely associated to a small number of migrants. On average, there are low migration stocks across two-thirds of African country of origin times country of destination dyads.

Figure 3 reports the results of our estimation procedure. In particular, for each of the 709 sub-national regions of Africa the figure reports the top three predicted European destinations . The objective of this exercise is twofold. First, to provide a preliminary assessment of the plausibility of the information content of our GT-based measure. Second, to assess the amount of variation generated by this measure within country of origin.

For each African region, Panel A reports to the top country of destination, which on average accounts for around 58% of total predicted migration to Europe from Africa. The map clearly hints at the importance of the European colonial legacy, with regions in West and Central Africa mainly linked to France and regions in East and Southern Africa linked to the UK, as well as more localized enclaves of Portuguese and Italian influence in, respectively, Mozambique and Angola and Libya and Ethiopia. Panels B and C refer to predicted migration to the second and third main European destinations, which account for 14% and 8% of total European migration, respectively. Compared to panel A, panels B and C display a more diverse set of destinations across regions of the same country, with as many as seven different top-2 European destinations in the same country of origin. Indeed, an Herfindahl index for the dispersion in the second (third) ranked European destination across regions of the same country delivers a value of 0.46 (0.38), reflecting the high degree of dispersion in European destinations within African countries.

3.3 Sub-national Validation

To further probe the predictive content of our measure of subnational migration stocks, we compare it with the limited available information on outmigration stocks from sub-national areas of origin in Africa by country of destination. Such data are extremely scarce and only available for a small number of countries. [8] In the following, we focus on measures of sub-national return migration calculated from census samples of 15 African countries from IPUMS. While the results of the analysis are overall invariant to the measure of migration used, census samples are likely to provide the best available approximation of the true extent of migration from given regions.

Figures 4 and A.8 depict the relationship between predicted and actual migrant stocks across regions of Morocco and Cameroon. Specifically, we plot the log number of return migrants from 1% census samples against the log-predicted stock of migrants using our methodology. Both values are augmented by one to address the presence of zeros and are residualized with respect to origin region fixed effects and country of destination fixed effects, obtained by pooling all origin-destination pairs. A clear positive gradient between predicted and actual migration stocks can be observed in all regions. Migration patterns vary across countries and across regions of the same country. European countries, highlighted in blue, constitute the primary destinations for all regions of Morocco, with notable migration stocks in Southern European countries like Italy, France, and Spain, as well as in Northern European countries such as Belgium, the Netherlands, and Germany. The intensity of migration to these countries varies among regions of origin and is consistently captured by our predicted migration measure. Similarly, in Cameroon, a close positive relationship between predicted and actual

[8] The nature of the data also varies: it ranges from measures of return migration calculated from 1% samples of national censuses for Africa, to current migration of former household members obtained from small-scale, specific-purpose surveys. Appendix Table A.6 reports, for each data source, the number of regions of origin, countries of destination and number of migrants available. In total, across all sources, we have information on the region of origin and country of destination of approximately 500,000 migrants from 21 countries of origin.

migration stocks is apparent across regions, with neighboring countries like Nigeria, Gabon, and Chad serving as the main destination. Among European destinations, France features prominently, while it is interesting to note substantial actual and predicted migration to Germany from the South-Western regions of Kumba and Bamenda, respectively the capital city and an important military station of German Kamerun (a colony of the German empire until 1916).

In Table 3, we formally estimate the relationship between return migration from national censuses in Africa and predicted sub-national migration stocks from Africa to European destinations. Our units of analysis are all "region of origin-country of destination" pairs for which we have information on predicted and actual migration. While unweighted results are presented, weighted results by population at origin or at destination (available upon request) are remarkably similar. We focus on a saturated model incorporating region of origin and country of origin X country of destination fixed effects. This approach allows us to control for variation in predicted and actual migration from a given region that is due to its size or popularity across all destinations. It also absorbs all time-invariant origin-destination links, allowing us to abstract from bilateral country ties, such as colonial history, common language, and trade links. Column (1) presents estimates across all world destinations, which delivers a positive and statistically significant point estimate (0.29). Similar results are found when estimating coefficients separately across destination continents, except for North America and Oceania, where IPUMS data reveal very limited migration from Africa. Overall, the evidence in this and in the previous section speaks strongly in favor of our GT-based measure of migration providing valuable information about actual migration stocks from sub-national regions of Africa, which are typically challenging to systematically estimate.

4 Empirical strategy

In this section, we present the empirical specifications we use to study the international transmission of support for democracy via migrant networks. The focus of our empirical analysis is the network of African migrants to Europe. We construct this network based on the measure of migration from sub-national regions of Africa to different European countries starting from Google searches data described in Section 3.

We are interested in understanding whether attitudes towards migrants in European destination countries – which are mostly full democracies – affect the view of democracy in their communities of origin in Africa. In particular, we ask whether communities of origin are less likely to consider democracy as a desirable form of government when their migrants are more subject to discrimination in democratic regimes. We measure preference for democracy in origin communities of Africa using data from the Afrobarometer survey.[9] This survey includes three set of information that are crucial for our analysis. First, it includes questions asking respondents about their view of democracy, electoral competition, rule of law, and active civic participation. Second, the data are geo-located, enabling us to identify the sub-national region of residence for each respondent and associate it with changes in the democratic experience of migrants from that region in Europe. Third, the data contains information on respondents' socio-economic characteristics, which allows us to control for individual-level other determinants of their view of democracy.

To construct measures of individual-level preferences for democracy, we focus on ten questions that have been asked consistently through waves 3 to 8 of the Afrobarometer. Our first measure, which we name "Preference for democracy", is based on the question asking respondents whether they think democracy is preferable to any other form of government. Our second measure, "Preference for electoral competition", captures respondents support for specific features of consolidated democracies, such as free and fair elections and multi-party competition. Our third measure, "Preference for rule of law" captures respondents support for constraints on the president (whether president should be bound by law and whether it should be monitored) and presence of a term limit. Finally, our last measure, "Civic participation", captures respondents active civic participation as captured by voting in the last elections, and participation in protests, community meetings or other forms of association to raise a specific issue. In our main analysis we construct respondents' preference for electoral competition, the rule of law and civic participation using a principal component of the answers to the questions underlying each measure, but we also report results for individual variables separately in Appendix.

Our baseline specification at the individual-respondent level is as follows:

$$y_{ir_ot} = \alpha_{r_o} + \alpha_{ot} + \beta \sum_{d \in D} \omega_{r_od}\, \textit{Anti-Imm}_{dt} + \lambda X_{ir_ot} + \gamma \Xi_{r_ot} + \varepsilon_{ir_ot} \quad (4)$$

where i indexes individual respondents to the Afrobarometer, r_o indexes region r in origin country o in Africa, d indexes destination countries, and t indexes years in which the survey was conducted. Our baseline specification includes region fixed effects capturing time-invariant regional characteristics (α_{r_o}), and country of origin interacted with time fixed effects

[9] The Afrobarometer is a pan-African institution conducting a public attitude survey. The first round of the survey was completed in 2001 and includes 12 African countries: Botswana, Ghana, Lesotho, Malawi, Mali, Namibia, Nigeria, South Africa, Tanzania, Uganda, Zambia, Zimbabwe. As per 2021, Afrobarometer conducted and completed 8 waves of this repeated cross-section survey. The last round was completed in 2021 and reports information on 34 African countries. Each country has a sample size of either 1,200 or 2,400 individuals. Samples are designed to generate a representative cross-section of all citizens of voting age in a given country.

capturing country-level aggregate trends (α_{ot}). African regions are the 709 admin1-level regions located in 51 African countries described in Section C. Standard errors are two-way clustered at country of origin and year to allow for correlation in the error term across respondents located in the same country and across respondents answering the Afrobarometer questionnaire in the same year.

The outcome variable y_{ir_ot} is individual i's answer to Afrobarometer questions on the level of support for democracy and preference for electoral competition, the rule of law and civic participation. We codify responses to the questions underlying our four measures as dummy variables equal to 1 if the respondent agrees or strongly agrees with statements about the desirability of these democratic features. For example, the outcome "Preference for democracy" is constructed as a dummy equal to 1 if the respondent answers that democracy is "preferable to any other form of government", and 0 if the respondent answers the other two potential responses, i.e. either that "it does not matter" or that "sometimes a non-democratic government can be preferable".

$Anti\text{-}Imm_{dt}$ is a time varying measure of anti-immigrant sentiment in destination country d. We use several definitions of anti-immigrant sentiment, as discussed in Section 2. First, we use the electoral success of parties that sponsor anti-immigration policies in destination countries and support closed society politics. We construct these measures by combining CHES data on policies promoted by different parties with electoral results in national elections. Second, we calculate the share of native respondents that are highly contrary to immigration from outside the European Union as recorded in the ESS. Third, we use a measure of perceived discrimination by immigrants themselves in destination countries as recorded in the ESS.

The weights ω_{r_od} capture the intensity of migration linkages between African region r_o and destination country d, and are defined as follows:

$$\omega_{r(o)d} = \frac{M_{r_od}}{M_{r_o}} \tag{5}$$

That is, ω_{r_od} captures the share of migrants from African region r_o that are currently living in destination country d. The M variables are the number of migrants constructed using predicted number of migrants from Google Search data and the methodology described in Section C. The methodology produces estimates of number of migrants from African region r_o to destination d for four benchmark years: 2005, 2010, 2015 and 2020. The weights are constructed by taking the average of predicted migrants over the four benchmark years in both the numerator and the denominator.

The Afrobarometer data also reports a large set of individual characteristics for each respondent. In all specifications, we control for the respondent's age, gender, level of education and a dummy capturing individuals living in urban areas (X_{ir_ot}). Importantly, we also control for a set of time-varying characteristics of the respondent's region (Xi_{r_ot}). These controls include the share of regional migrants over total population, the regional share of migrants to Europe over total migrants from that region, as well as other observable characteristics of the primary sampling unit of the respondent reported in the Afrobarometer such as access to electricity, access to running water, presence of a school, a police station, a hospital, or a local market.

One potential concern with the model described by equation (4) is that shocks to attitude towards migrants at destination might be correlated with other shocks at destination, which are also transmitted to the communities of origin via the migrant network, affecting the outcomes of interest. One example of this issue would be if anti-immigrant sentiment grows during economic recessions in European countries, leading to lower economic remittances to the communities of origin. If communities of origin respond to this income shock by lowering their support for democracy, that would constitute an alternative mechanism for our results.

To deal with this issue, we include in our specification a measure of exposure to changes in economic conditions at destination, which follows the same structure as our main explanatory variable. That is, for each region of origin and year, we construct a measure of exposure to changes in economic conditions in Europe via migrant networks as the weighted average of GDP growth at destination, where the weights are represented by the share of migrants from a given region of origin in that destination country. Notice that augmenting our specification with this control might absorb some of the variation we are most interested in, especially to the extent that anti-immigration sentiment is enhanced during periods of low economic growth and high unemployment because migrants are depicted as scapegoats. However, the results reported in Section 5 show that the estimated β in equation (4) remains stable when adding this control. This suggests that exposure to anti-immigrant sentiment and exposure to recessions at destination capture different variation in the data.

Finally, we study whether variation in the ability to communicate with destination regions across individuals within the same community of origin affect the international transmission of preferences for democracy. We use access to the mobile phone network as a proxy for communication costs. We use geographical coordinates of Afrobaromenter respondents' location matched with fine geographical data on the diffusion of mobile phone coverage to construct an individual-level measure of access to the mobile phone network.[10]

Figure 5 shows a visual example of our dataset for countries in West Africa. As shown, the

[10] Data on mobile phone coverage is sourced from the Global System for Mobile Communication Association (GSMA). GSMA is the association representing the interests of the mobile phone operators worldwide. The data is collected by GSMA directly from mobile operators and refers to the 3G GSM network.

data allow us to compare individuals with and without access to the mobile phone network and that are located within the same community of origin in Africa. We test for heterogeneous effects by mobile phone coverage at origin by estimating the following specification:

$$\begin{aligned} y_{ir_ot} &= \alpha_{r_o} + \alpha_{ot} + \beta_1 \sum \omega_{r_od}\ \textit{Anti-Imm}_{dt} + \beta_2 \sum \omega_{r_od} \textit{Anti-Imm}_{dt} \times 3G_{ir_ot} \\ &+ \lambda X_{ir_ot} + \gamma \Xi_{r_ot} + \varepsilon_{ir_ot} \end{aligned} \tag{6}$$

where $3G_{ir_ot}$ is a dummy equal to one if individual i location was covered by the mobile phone network in the year of the Afrobarometer survey. Although mobile phone based money transfers were scarcely used during the period under study in equation (6) we always control also for an interaction of exposure to economic conditions at destination with mobile phone access.

5 Results

5.1 Effects of exposure to anti-immigrant sentiment at destination on origin communities

Our analysis focuses on migration from Africa to Europe for the years 2005 to 2021. As described in Section 4, we estimate equations (4) and (6) using different measures of anti-immigrant sentiment in destination countries, that range from political support for a closed society to public sentiment towards immigrants as well as self-reported migrants' perceived discrimination in the host countries. In total we have four main outcomes variables and eight measures of shocks at destination. For presentation purposes, we present regression results in table format for only one outcome and one measure of shock at destination, and we then move to present point estimates for the coefficient of interest across all outcomes and shocks at destination in graphical format.

Columns (1) and (2) of Table 4 report the results of estimating equation (4) where the outcome variable is individual-level "Preference for democracy" and the measure of anti-immigrant sentiment at destination is the political support for parties advocating for strict immigration policy. All specifications include region of origin fixed effects, country of origin times year fixed effects, individual controls and location controls.

The estimated β coefficient in column (1) is negative and statistically significant, indicating that individuals in communities of origin exposed to an increase in anti-immigrant sentiment at destination experience a relative decline in their support for democracy. To facilitate the interpretation of magnitudes, we report standardized coefficients in all tables.

Thus, the magnitude of the estimated coefficient in column (1) indicates that a one standard deviation increase in exposure to anti-immigrant sentiment at destination translates into a 0.1 of a standard deviation decline in the probability to support democracy at origin. This effect is large when considering that the average support for democracy among Afrobarometer respondents at origin across all years is 0.73, with a standard deviation of 0.44.

In column (2), we include our measure of exposure to GDP growth at destination via migrant networks. This measure is designed to capture changes in attitudes towards democracy driven by changes in economic conditions in the European regions where local individuals have migrated. As shown, the point estimate on exposure to anti-immigrant sentiment remains of similar magnitude, indicating the two measures of exposure capture independent variation. In addition, we find that exposure to economic conditions at destination to have no significant effect on preference for democracy in the regions of origin.

Next, in Figure 6, we report estimates obtained by estimating the specification in column (2) of Table 4 for all four outcome variables and all eight measures of anti-immigrant sentiment in destination countries.[11] In panel (a), upper left panel of Figure 6, we focus on preference for democracy as outcome. As shown, we find consistently negative effects of increases in different measures of anti-immigrant sentiment at destination on preference for democracy of respondents at origin. The magnitude of the estimated effects is also similar across the eight definitions of anti-immigrant sentiment. Regardless of the measure used, a one standard deviation increase in anti-immigrant sentiment in the destination country is associated with a 0.1 to 0.15 standard deviation decline in preference for democracy among respondents at origin.

In panel (b), upper right panel of Figure Figure 6, we focus instead on the outcome "preference for electoral competition". This outcome is defined as the principal component of two variables: support for free and fair elections and preference for multi-party system. Also in this case, increases in different measures of anti-immigrant sentiment at destination generate a statistically significant decline in preference for electoral competition among respondents at origin, with magnitudes ranging between -0.1 and -0.2. One exception is when we use migrants' perceived discrimination at destination as a measure of anti-immigrant sentiment. In this case, the estimate is close to zero and not statistically significant.

In panels (c) and (d) we focus on two additional outcomes: preference for the rule of law and civic participation. For both outcomes, estimates tend to be more noisy and we typically cannot reject that the overall effect of negative immigrants' experience at destination has no statistically significant impact on origin communities. Nonetheless, point estimates across different measures of anti-immigrant sentiment indicate a negative impact such sentiment on

[11] Appendix Table A.5 reports the corresponding estimates.

both outcomes. In particular, when focusing on the effect of restrictive immigration policies, we find negative estimates of similar magnitudes as the ones in panels (a) and (b), though less precisely estimated.[12]

5.2 Heterogeneous effects by mobile phone coverage

In the last part of the paper we study the heterogeneous effects of international transmission of preference for democracy via migrant networks for respondents with different access to the mobile phone network. To this end, we estimate equation (6), which relies on variation in mobile phone access across individuals within the same region in Africa.

The results are reported in columns (3) and (4) of Table 4. The point estimates on the interaction coefficient β_2 indicate that, for a given exposure to anti-immigrant sentiment, individuals with access to mobile phone coverage experience a 23-25% larger decline in their support for democracy relative to individuals in the same region but without access to mobile coverage. In column (4) we augment the estimating equation by controlling for exposure to GDP growth at destination interacted with mobile phone coverage. This accounts for potential differential effects of being connected via mobile phones to destination countries in different stages of their business cycle. As shown, the estimated coefficients β_1 and β_2 remain stable in magnitude and of similar precision after including this control.

Figure 7 reports the point estimates and confidence intervals for the interaction term β_2 when using all our four main outcome variables and all measures of anti-immigrant sentiment at destination.[13] As shown, access to the mobile phone network seem to amplify the negative effect of exposure to anti-immigrant sentiment via migrant networks on individuals' preference for democracy, preference for electoral competition and civic participation. On the other hand, we find estimates that are close to zero and not statistically significant when focusing on preference for rule of law as an outcome. However, by and large, the results indicate an incremental decline in democratic support at origin among individuals who have easier access to information about potentially negative experiences of individuals from their region that migrated to democratic countries in Europe.

[12] Appendix Figure A.9 reports the same results for all individual outcome variables. While the results are largely in line with those reported in Figure 6, it is worth noting that the coefficients are precisely identified on key variables such as voting in the last elections and individual support for multi-party systems.

[13] Appendix Figure A.10 reports estimates separately for each component of the four main outcome variables.

6 Summary and concluding remarks

In this paper we propose a novel methodology to estimate sub-national migration stocks using Google Trends data. We postulate and find strong evidence in favor of migrants disproportionately performing Google searches for their place of origin, making these searches a good proxy for their presence. We use these estimates to derive a measure of exposure to anti-immigrant sentiment at destination across 709 sub-national areas in Africa, obtained as a weighted average of migrants' baseline composition in terms of destination countries and measures of actual or perceived anti-immigrant sentiments across European destination countries. We apply this measure to the study of the international transmission of support for democracy and preference for characteristics of democratic systems. Our results show that increased anti-immigrant sentiments in Europe spill over to communities of origin in the form of reduced support for democratic institutions.

As public support for democracy is considered a fundamental prerequisite for a legitimate and lasting democratic structure in countries with weak institutional capacity, these findings suggest that the consequences of negative anti-immigrant sentiment in Europe may be far-reaching and have important political consequences for the future path of countries at the crossroad in their democratic journey. While one of the long-term efforts of Western foreign policy has been to export democracy through the creation of liberal democratic institutions in countries with limited institutional capacity, often with poor results, our results suggest that a key factor in consolidating democracy abroad is setting an example at home.

References

Acemoglu, D., N. Ajzenman, C. G. Aksoy, M. Fiszbein, and C. A. Molina (2021). (successful) democracies breed their own support. Technical report, National Bureau of Economic Research.

Algan, Y. and P. Cahuc (2010). Inherited trust and growth. *American Economic Review 100*(5), 2060–2092.

Barsbai, T., H. Rapoport, A. Steinmayr, and C. Trebesch (2017). The effect of labor migration on the diffusion of democracy: evidence from a former soviet republic. *American Economic Journal: Applied Economics 9*(3), 36–69.

Batista, C. and P. C. Vicente (2011). Do migrants improve governance at home? evidence from a voting experiment. *The World Bank Economic Review 25*(1), 77–104.

Billari, F. C., V. Rotondi, and J. Trinitapoli (2020). Mobile phones, digital inequality, and fertility. *Demographic Research 42*, 1057–1096.

Blumenstock, J., O. Dube, and K. Hussain (2022). Can secular media create religious backlash? evidence from pakistan's media liberalization.

Boese, V. A., M. Lundstedt, K. Morrison, Y. Sato, and S. I. Lindberg (2022). State of the world 2021: autocratization changing its nature? *Democratization 29*(6), 983–1013.

Böhme, M. H., A. Gröger, and T. Stöhr (2020). Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics 142*, 102347.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies 86*(4), 1448–1486.

Chauvet, L. and M. Mercier (2014). Do return migrants transfer political norms to their origin country? evidence from mali. *Journal of Comparative Economics 42*(3), 630–651.

Choi, H. and H. Varian (2012). Predicting the present with google trends. *Economic record 88*, 2–9.

Couttenier, M., V. Petrencu, D. Rohner, and M. Thoenig (2019). The violent legacy of conflict: evidence on asylum seekers, crime, and public policy in switzerland. *American Economic Review 109*(12), 4378–4425.

Docquier, F., E. Lodigiani, H. Rapoport, and M. Schiff (2016). Emigration and democracy. *Journal of Development Economics 120*, 209–223.

Dustmann, C., K. Vasiljeva, and A. Piil Damm (2019). Refugee migration and electoral outcomes. *The Review of Economic Studies 86*(5), 2035–2091.

Easton, D. (1965). A systems analysis of political life.

Fernández, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics 1*(1), 146–177.

Fuchs-Schündeln, N. and M. Schündeln (2015). On the endogeneity of political preferences: Evidence from individual experience with democracy. *Science 347*(6226), 1145–1148.

Giuliano, P. and M. Tabellini (2020). The seeds of ideology: Historical immigration and political preferences in the united states. Technical report, National Bureau of Economic Research.

Guiso, L., P. Sapienza, and L. Zingales (2004). The role of social capital in financial development. *American economic review 94*(3), 526–556.

Guriev, S. and E. Papaioannou (2022). The political economy of populism. *Journal of Economic Literature 60*(3), 753–832.

Halla, M., A. F. Wagner, and J. Zweimüller (2017). Immigration and voting for the far right. *Journal of the European Economic Association 15*(6), 1341–1385.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012). Measuring economic growth from outer space. *American economic review 102*(2), 994–1028.

Hernández, B., M. C. Hidalgo, M. E. Salazar-Laplace, and S. Hess (2007). Place attachment and place identity in natives and non-natives. *Journal of environmental psychology 27*(4), 310–319.

Jensen, R. and E. Oster (2009). The power of tv: Cable television and women's status in india. *The Quarterly Journal of Economics 124*(3), 1057–1094.

Jolly, S., R. Bakker, L. Hooghe, G. Marks, J. Polk, J. Rovny, M. Steenbergen, and M. A. Vachudova (2022). Chapel hill expert survey trend file, 1999–2019. *Electoral studies 75*, 102420.

Karadja, M. and E. Prawitz (2019). Exit, voice, and political change: Evidence from swedish mass migration to the united states. *Journal of Political Economy 127*(4), 1864–1925.

Levitt, P. (1998). Social remittances: Migration driven local-level forms of cultural diffusion. *International migration review 32*(4), 926–948.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American political science review 53*(1), 69–105.

Manacorda, M., G. E. Tabellini, and A. Tesei (2023). *Mobile Internet and the Rise of Communitarian Politics*. Centre for Economic Policy Research.

Mercier, M. (2016). The return of the prodigy son: Do return migrants make better leaders? *Journal of Development Economics 122*, 76–91.

Nuti, S. V., B. Wayda, I. Ranasinghe, S. Wang, R. P. Dreyer, S. I. Chen, and K. Murugiah (2014). The use of google trends in health care research: a systematic review. *PloS one 9*(10), e109583.

Persson, T. and G. Tabellini (2009). Democratic capital: The nexus of political and economic change. *American Economic Journal: Macroeconomics 1*(2), 88–126.

Spilimbergo, A. (2009). Democracy and foreign education. *American economic review 99*(1), 528–43.

Spyratos, S., M. Vespe, F. Natale, I. Weber, E. Zagheni, and M. Rango (2019). Quantifying international human mobility patterns using facebook network data. *PloS one 14*(10), e0224134.

Tuan, Y.-F. (1990). *Topophilia: A study of environmental perception, attitudes, and values*. Columbia University Press.

Yarkin, A. (2023). Learning from the origins.

Figure 1: Anti-immigrant sentiment at destination

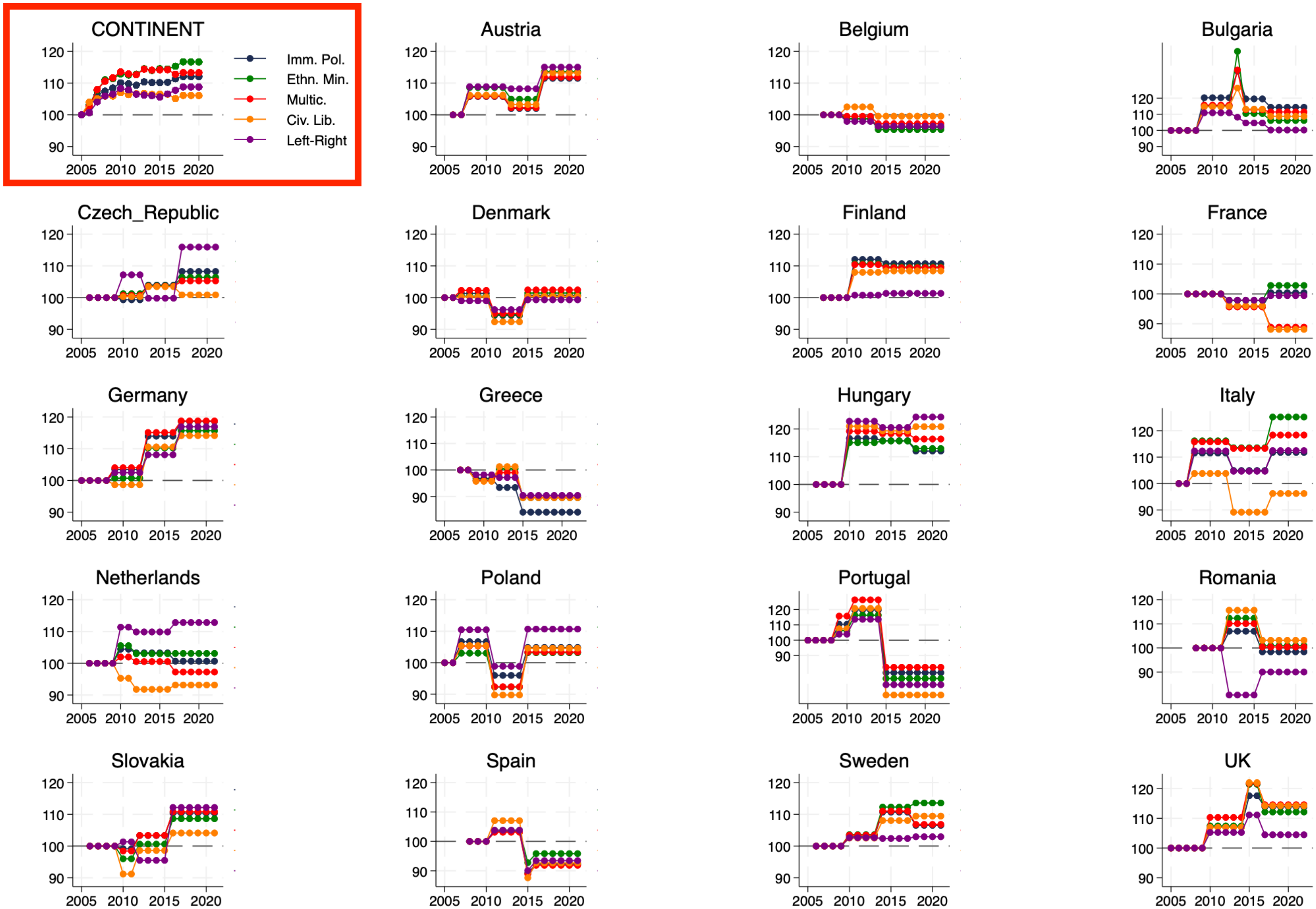
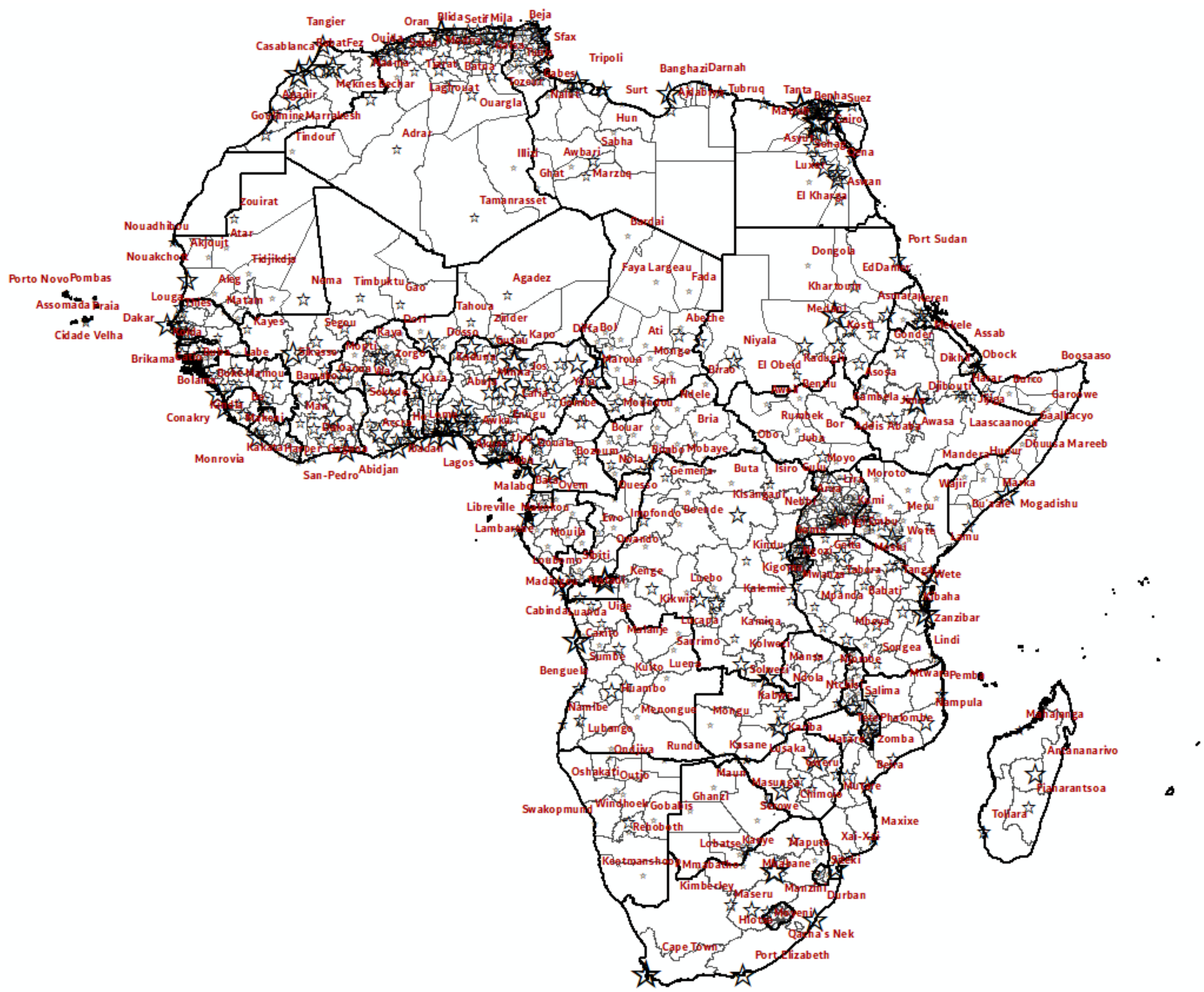


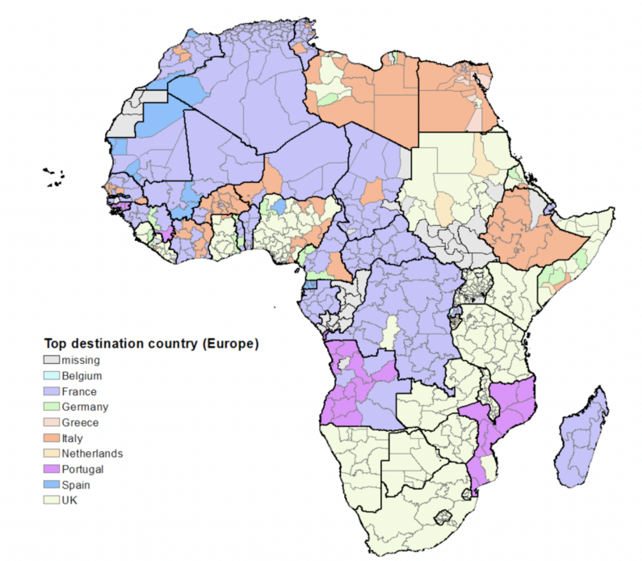
Figure 2: Largest city by Admin1 sub-national region



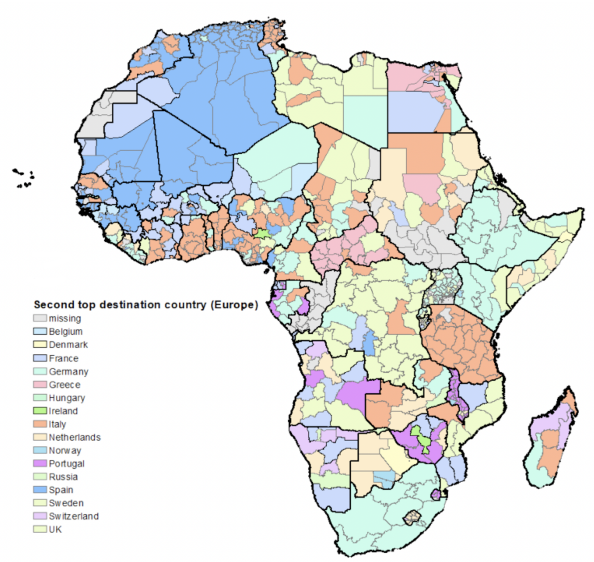
Notes: The map reports the name and location of the largest city in each of the 709 Admin1 sub-national regions in our sample. We use the names of these cities as search terms in Google Trends, as proxies for the interest in the corresponding Admin1 region of reference.

Figure 3: Top European Destinations

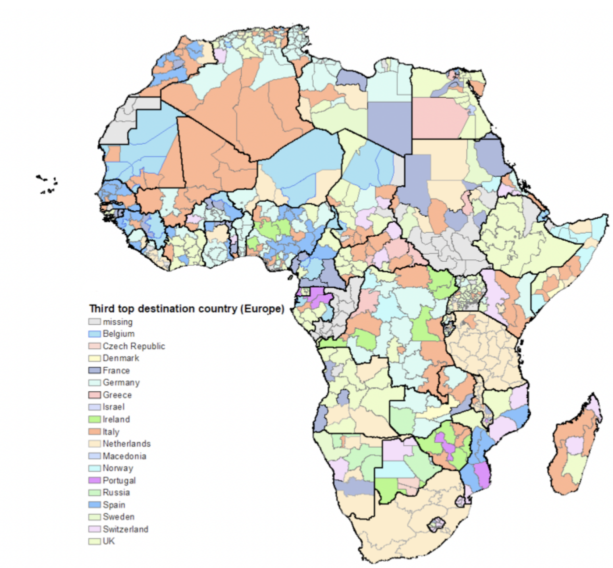
(a) Panel A



(b) Panel B

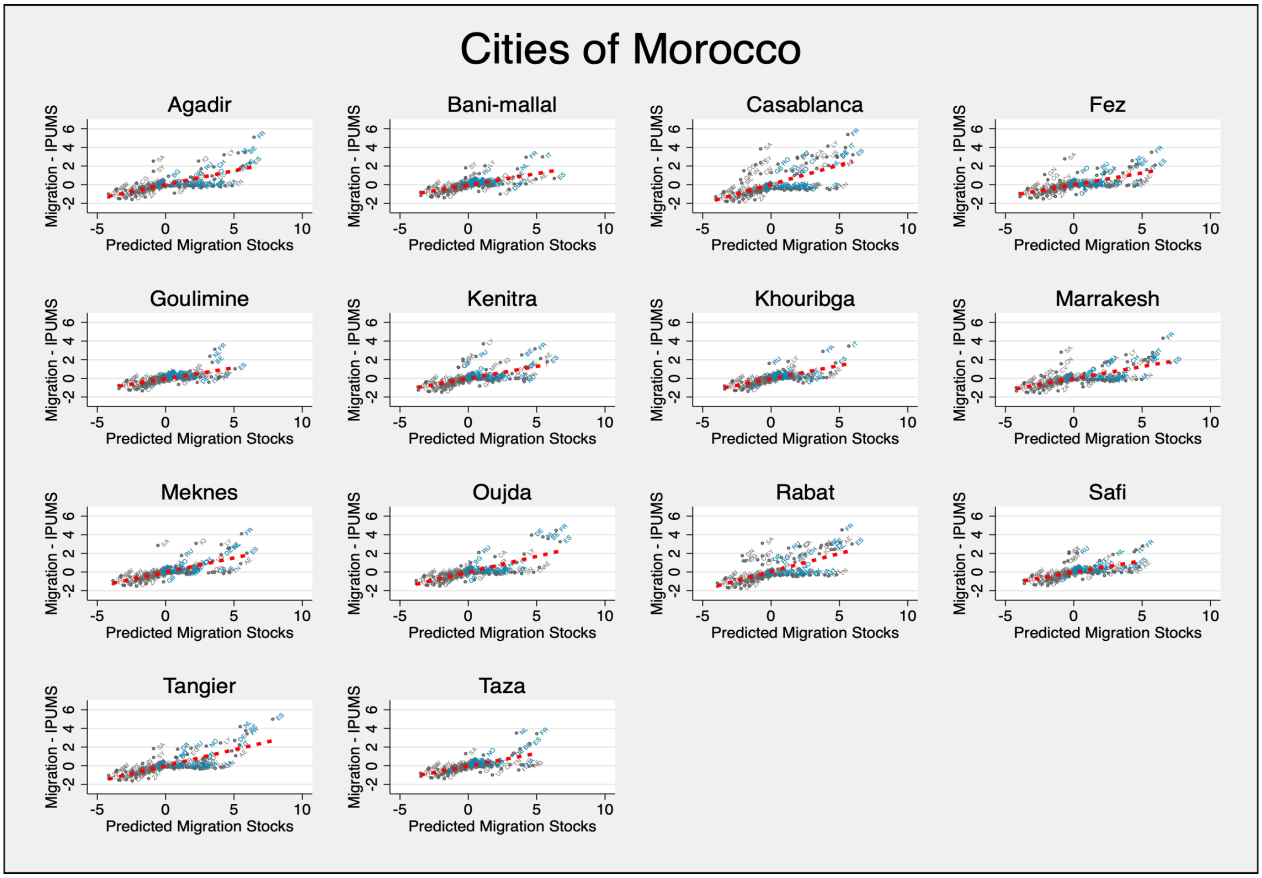


(c) Panel C



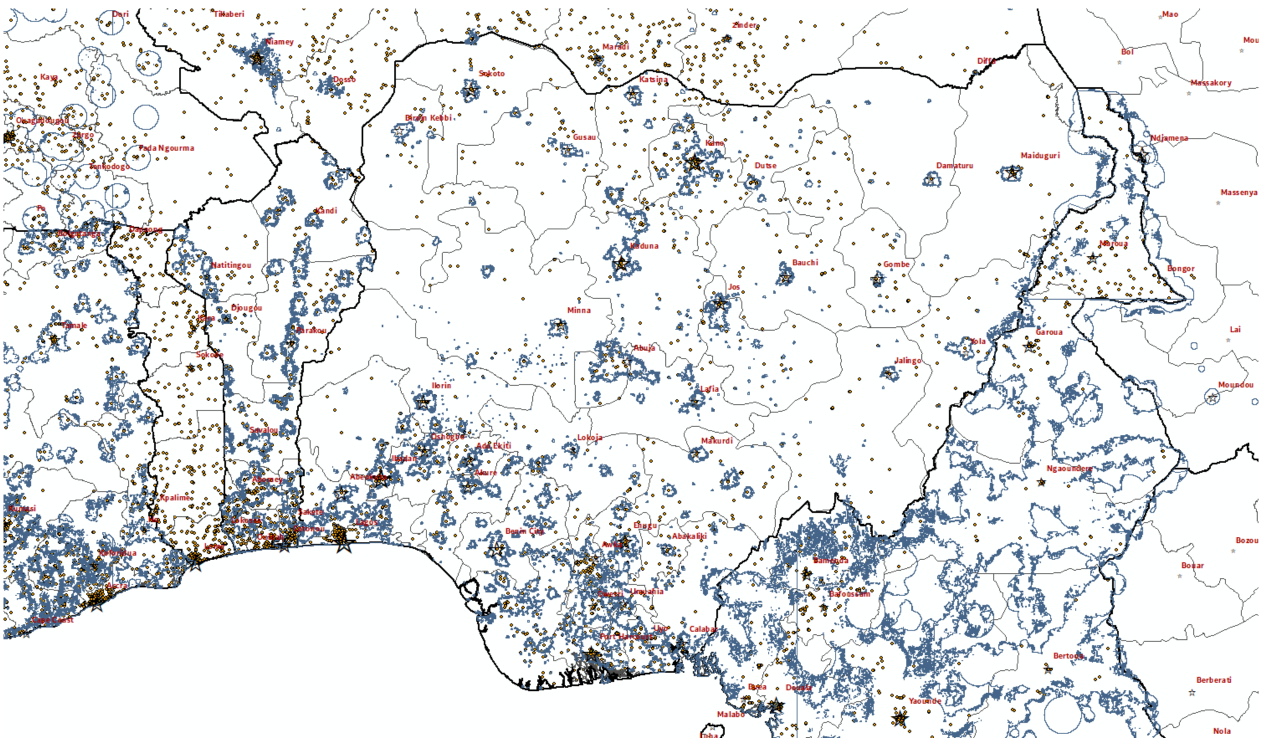
Notes: The figure reports the top three European destination countries across the 709 admin-regions in Africa.

Figure 4: Sub-national Validation - Cities of Morocco



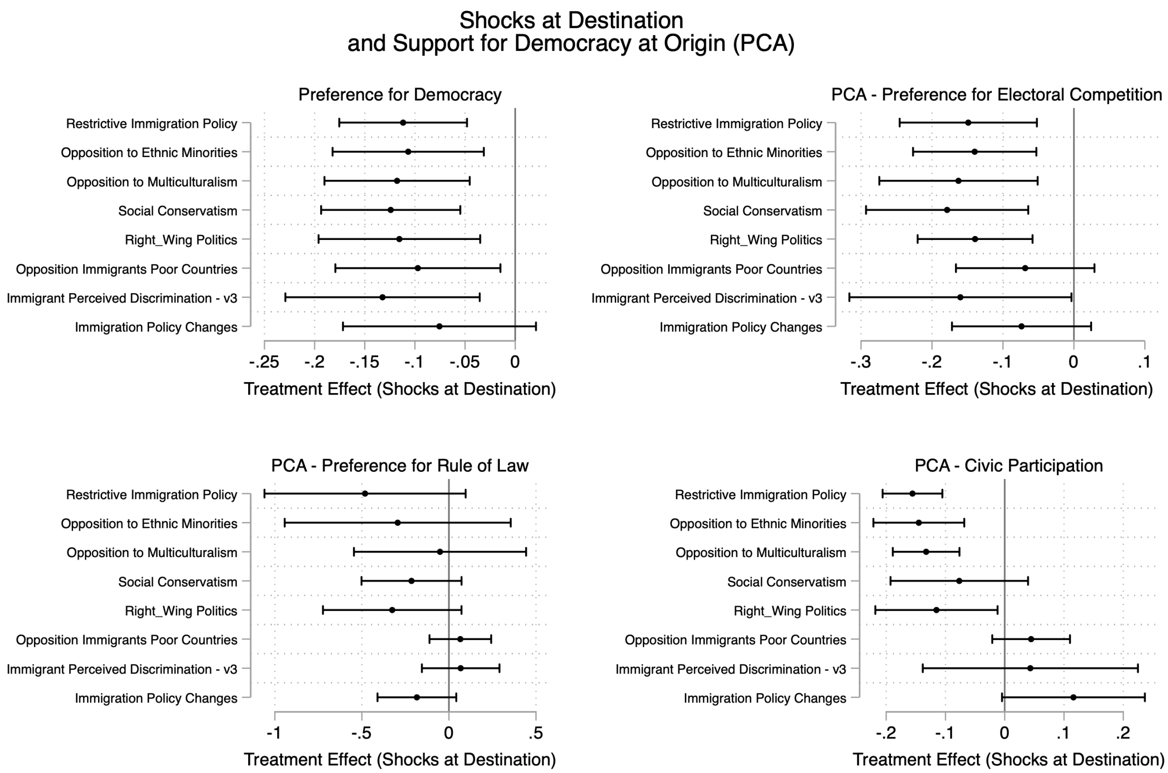
Notes: The figure reports log migration stocks from admin1 regions of Morocco to all destinations vis a vis average log Google searches for these regions in all destinations. Searches are standardized relative to corresponding values for Switzerland. Data are residuals from regressions on city fixed effects and country of destination fixed effects.

Figure 5: Location of Afrobarometer respondents and mobile coverage in West Africa



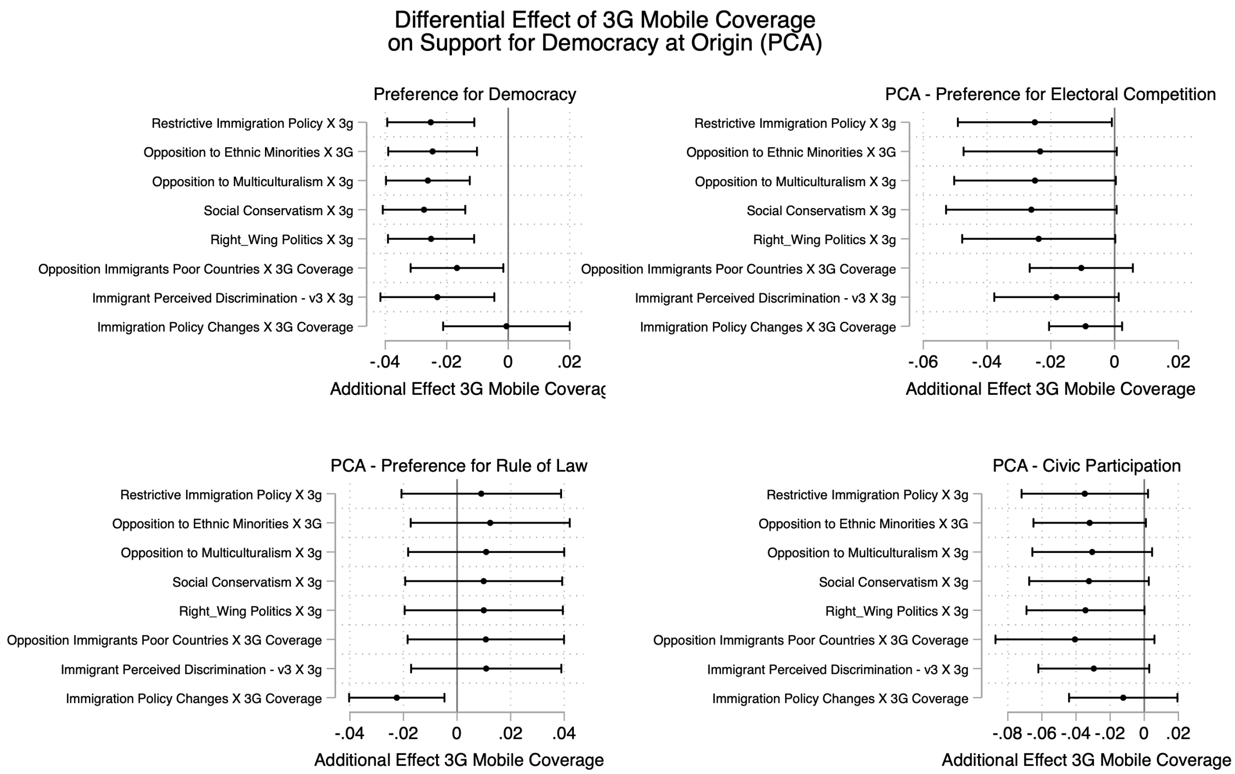
Notes: The maps zooms onto Western Africa to show the location of Afrobarometer respondents (blue triangles) in Round 7 (carried out between September 2016 and September 2018) along with the extent of 3G mobile phone coverage (in orange) over the region in 2018.

Figure 6: The effect of exposure to anti-immigration sentiment via migrant networks on Afrobarometer respondents



Notes: The Figure reports the estimated coefficients β with 95 percent confidence intervals from equation (4). We report the coefficients separately in four panels which correspond to the four main outcomes variables. For each outcome variable, we report the results for the eight measures of exposure to anti-immigration sentiment in destination countries.

Figure 7: The effect of exposure to anti-immigration sentiment via migrant networks on Afrobarometer respondents: Heterogeneous effect by mobile phone coverage at origin.



Notes: The Figure reports the estimated coefficients β_2 with 95 percent confidence intervals from equation (6). We report the coefficients separately in four panels which correspond to the four main outcomes variables. For each outcome variable, we report the results for the eight measures of exposure to anti-immigration sentiment in destination countries.

Table 1: Politics at Destination: Discrimination and Satisfaction with Democracy

Trait:	Right-Wing Politics	Social Conservatism	Restrictive Immigration Policy	Opposition to Ethnic Minorities	Opposition to Multiculturalism
	(1)	(2)	(3)	(4)	(5)
	Dependent variable: Perceived discrimination				
Non-EU immigrant	0.152*	0.162**	0.248***	0.076	0.143**
	(0.081)	(0.074)	(0.065)	(0.087)	(0.071)
Trait × Non-EU immigrant	0.304*	0.293*	0.112	0.470***	0.293**
	(0.156)	(0.150)	(0.117)	(0.167)	(0.123)
Observations	174,428	174,428	174,428	174,428	174,428
(Non-EU immigrants)	(6,297)	(6,297)	(6,297)	(6,297)	(6,297)
	Dependent variable: Satisfaction with democracy				
Non-EU immigrant	0.057	-0.028	-0.060	-0.044	-0.066
	(0.067)	(0.055)	(0.055)	(0.053)	(0.053)
Trait × Non-EU immigrant	-0.371***	-0.214**	-0.137	-0.183*	-0.121
	(0.126)	(0.104)	(0.095)	(0.105)	(0.092)
Observations	169,021	169,021	169,021	169,021	169,021
(Non-EU immigrants)	(6,029)	(6,029)	(6,029)	(6,029)	(6,029)
Region FE	✓	✓	✓	✓	✓
Country X Year FE	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓

Notes: Higher values of the trait indicate more Communitarian politics. The specifications include NUTS2 FE, country X year FE, and the following individual controls: age, education, dummies for marital status, unemployed, big city, gender, religion, country of birth. Source: European Social Survey. ***,**,*: statistically significant at 1%, 5% and 10%, respectively.

Table 2: Country-level correlations between bilateral migration stocks and GT searches

	Dependent Variable: Country-level migration			
	(1)	(2)	(3)	(4)
Aggregate GT searches	1.470***	1.504***	1.119***	1.287***
	(0.284)	(0.306)	(0.190)	(0.201)
Observations	2,070	2,070	2,062	2,062
Country of origin FE	×	✓	×	✓
Country of dest FE	×	×	✓	✓

Notes: The table reports GLS estimates of equation (3), with weights equal to population at origin. Searches are standardized to corresponding values for Switzerland. Two-way standard errors clustered by country of origin and destination reported in parenthesis. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 3: Relationship between log return migration from African National censuses and predicted migration stocks from Google trends, by African regions

	Dependent Variable: Log IPUMS return migrants						
	Any destination (1)	Africa (2)	Europe (3)	Asia (4)	Latin-America (5)	North-America (6)	Oceania (7)
Log predicted migration stock from GT	0.296***	0.382***	0.244**	0.221***	0.086*	0.031	0.021
	(0.041)	(0.056)	(0.089)	(0.050)	(0.041)	(0.062)	(0.073)
Observations	30,420	10,998	6,552	8,190	3,744	468	468
Region origin FE	✓	✓	✓	✓	✓	✓	✓
Country orig X Country dest FE	✓	✓	✓	✓	✓	✓	✓

Notes: The tables reports the coefficient from a regression of log return migrants stocks from each country to each region in Africa based on IPUMS data on log migrants stocks estimated based on GT data. Regressions include region of origin plus country of origin times country of destination fixed effects. The top panel reports unweighted estimates, the middle panel estimates weighted by population at destination and the bottom panel estimates weighted by population at origin. Two-way standard errors clustered by country of origin and destination reported in parenthesis. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 4: The effect of exposure to anti-immigrant sentiment at destination on support for democracy at origin

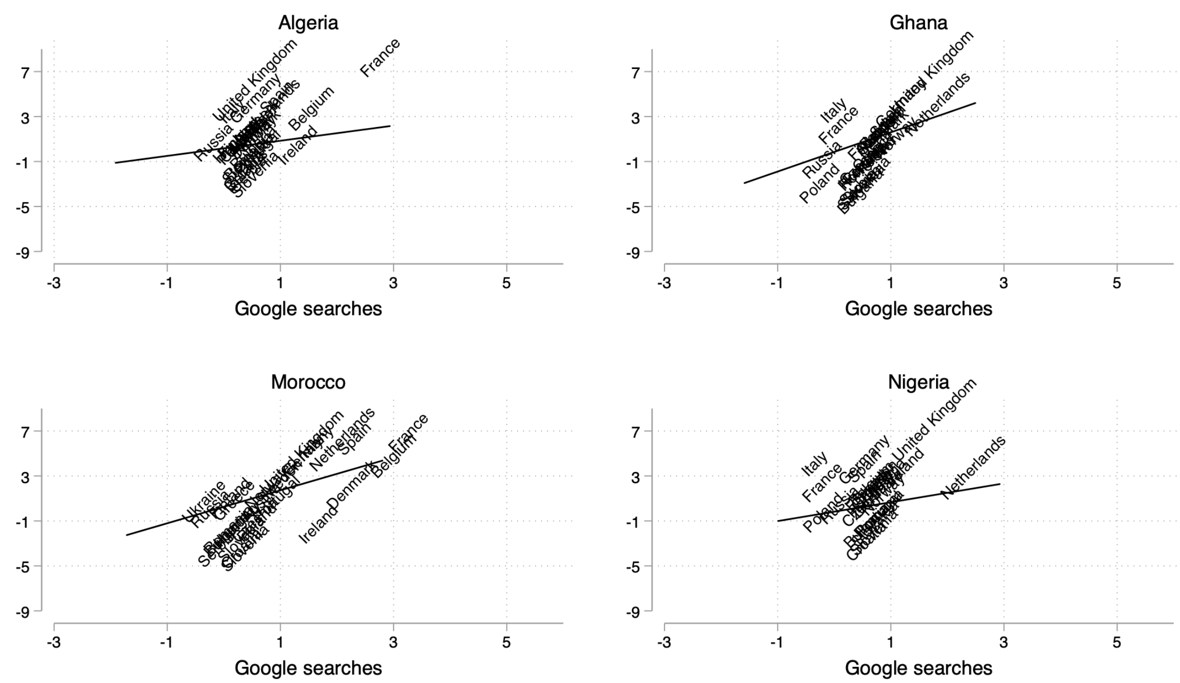
	Preference for democracy			
	(1)	(2)	(3)	(4)
Restrictive Immigration Policy at destination	-0.104**	-0.099**	-0.103**	-0.099**
	(0.046)	(0.045)	(0.046)	(0.045)
Restrictive Immigration Policy at destination × 3G Coverage			-0.025**	-0.023**
			(0.010)	(0.010)
Exposure to GDP growth at destination		0.011		0.010
		(0.012)		(0.016)
Exposure to GDP growth at destination × 3G Coverage				0.011
				(0.014)
Coverage			0.010	0.011
			(0.013)	(0.012)
Observations	215,647	215,647	215,647	215,647
Region origin FE	✓	✓	✓	✓
Country origin X Year FE	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓
Region Controls	✓	✓	✓	✓

Notes: The table reports OLS standardized coefficients of equation (4) in columns 1-2 and equation (6) in columns 3-4. Coverage is a dummy equal to 1 if the reported geographical location of the respondent is covered by 3G mobile network in the Afrobarometer wave year. Individual controls include respondent's age, gender, level of education and a dummy capturing individuals living in urban areas. Regional (or PSU, primary sampling unit) controls include: share of regional migrants over total population, the regional share of migrants to Europe over total migrants from that region, access to electricity, access to running water, presence of a school, a police station, a hospital, or a local market. Standard errors are two-way clustered at country of origin and year. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$..

Appendix Tables and Figures

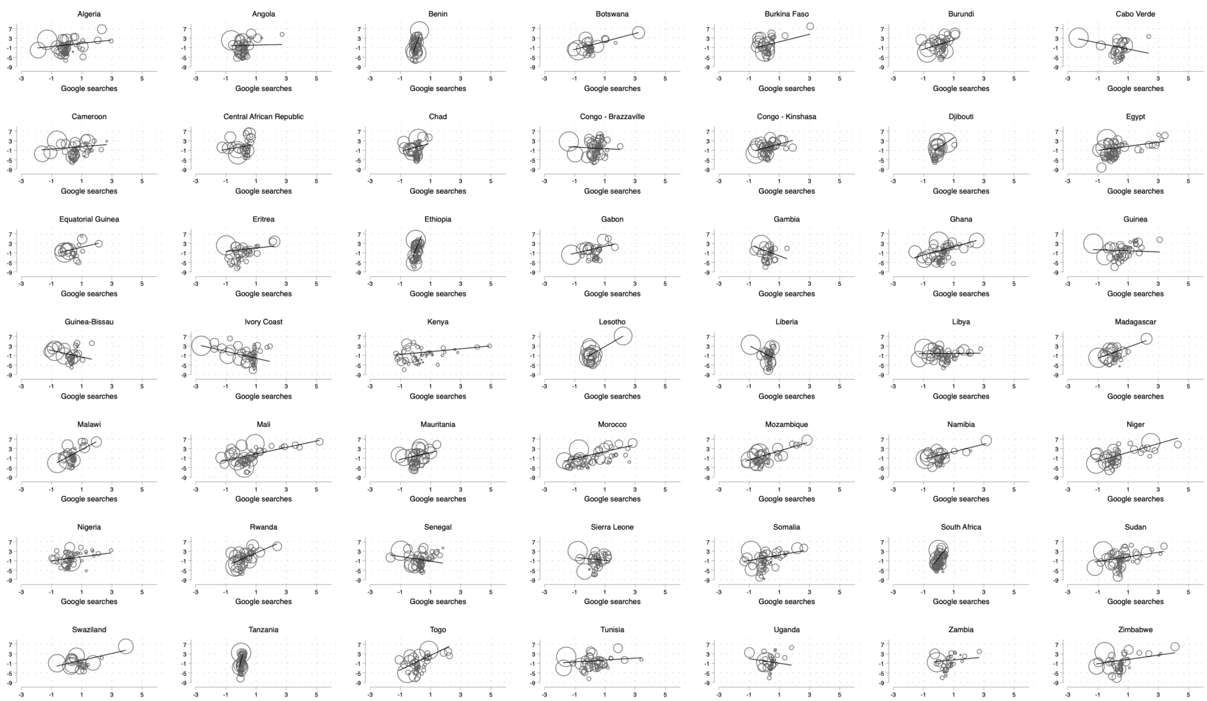
Figure A.1: Zoom west africa

Figure A.2: Outmigration versus Google searches across country of destination - Selected African countries of origin



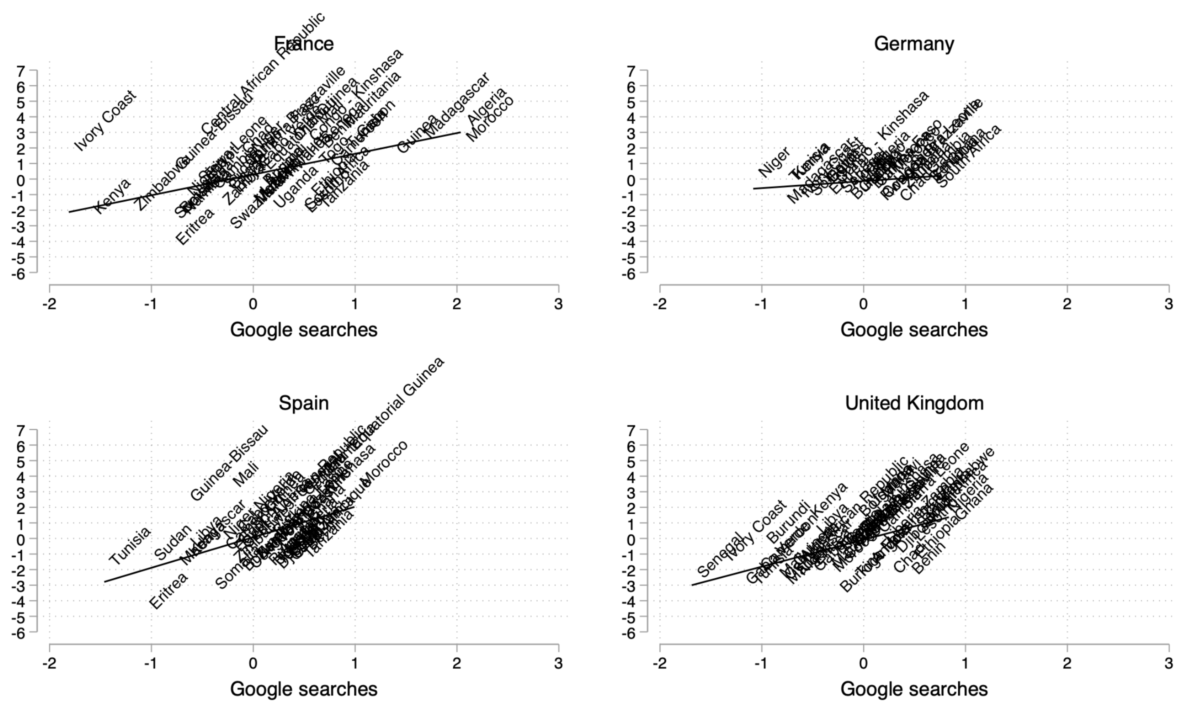
Notes: The figure reports log migration stocks from four African countries to all destinations vis a vis average log Google searches for regions of such origin countries in all destinations. Searches are standardized relative to corresponding values for Switzerland. Data are residuals from regressions on country of destination fixed effects and are weighted by population at origin. Data refer to the year 2015.

Figure A.3: Outmigration versus Google searches across countries of destination



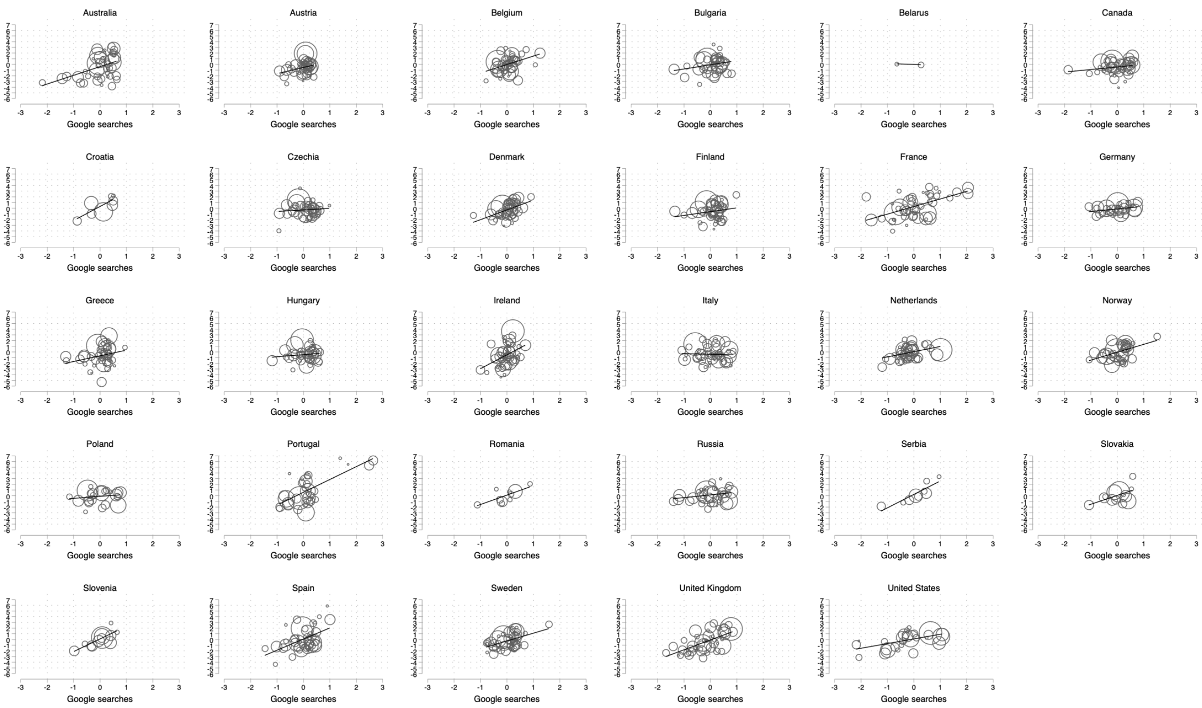
Notes: See notes to Figure A.2. Larger dots correspond to larger countries of destination.

Figure A.4: Outmigration versus Google searches across country of origin - selected European countries of destination



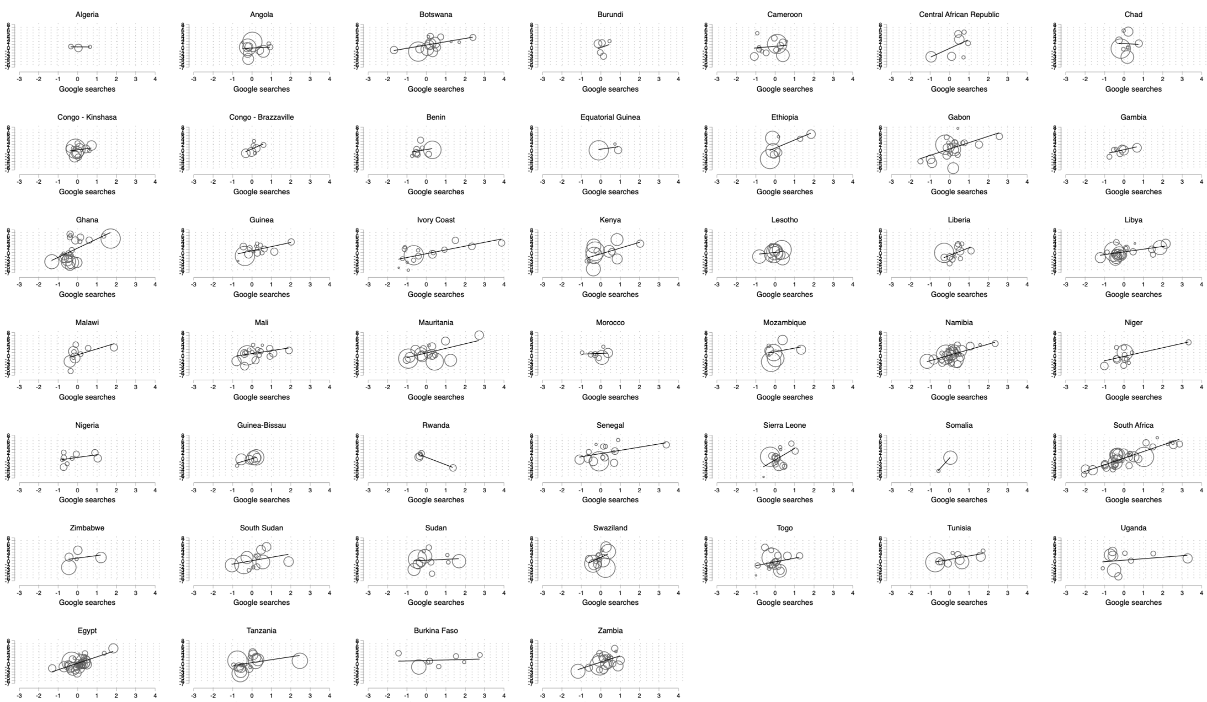
Notes: The figure reports log migration stocks from all African countries to four European destinations vis a vis average log Google searches for regions of origin countries in such destinations. Searches are standardized relative to corresponding values for Switzerland. Data are residuals from regressions on country of origin fixed effects and are weighted by population at destination. Data refer to the year 2015.

Figure A.5: Outmigration versus Google searches across country of origin - Destination countries in Europe, North America and Oceania



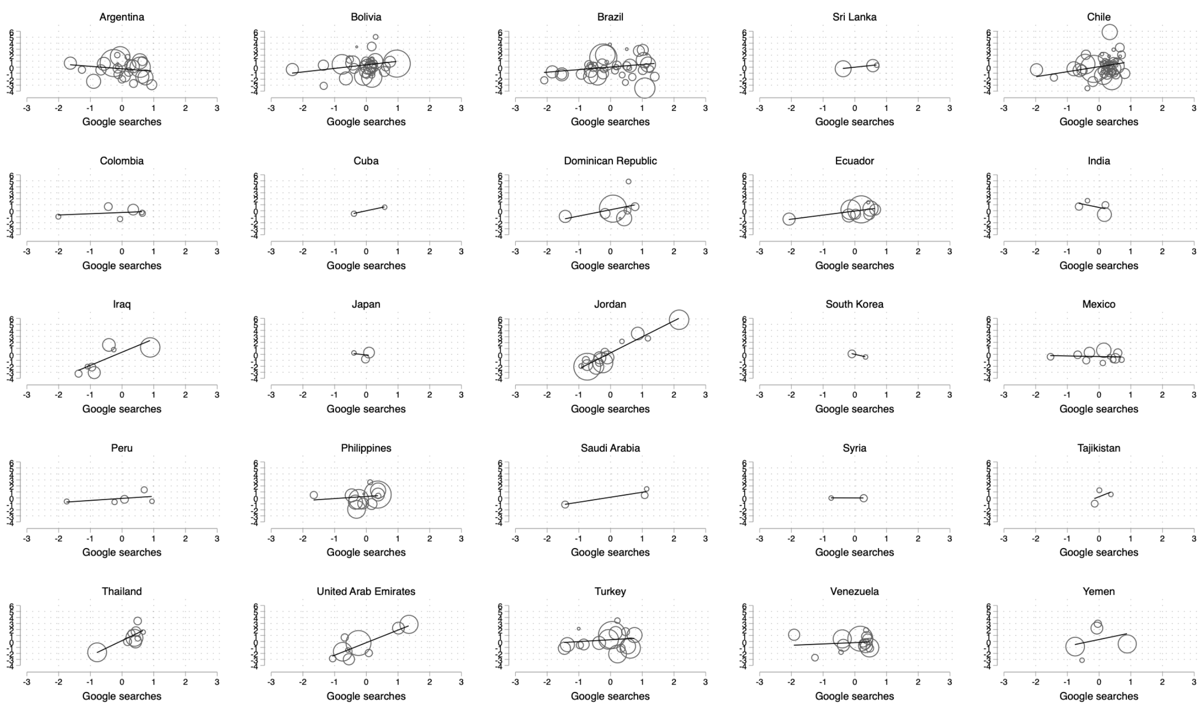
Notes: See notes to Figure A.4. Larger dots correspond to larger countries of origin. Data only refer to countries of destination with population greater than one million.

Figure A.6: Outmigration versus Google searches across country of origin - Destination countries in Africa



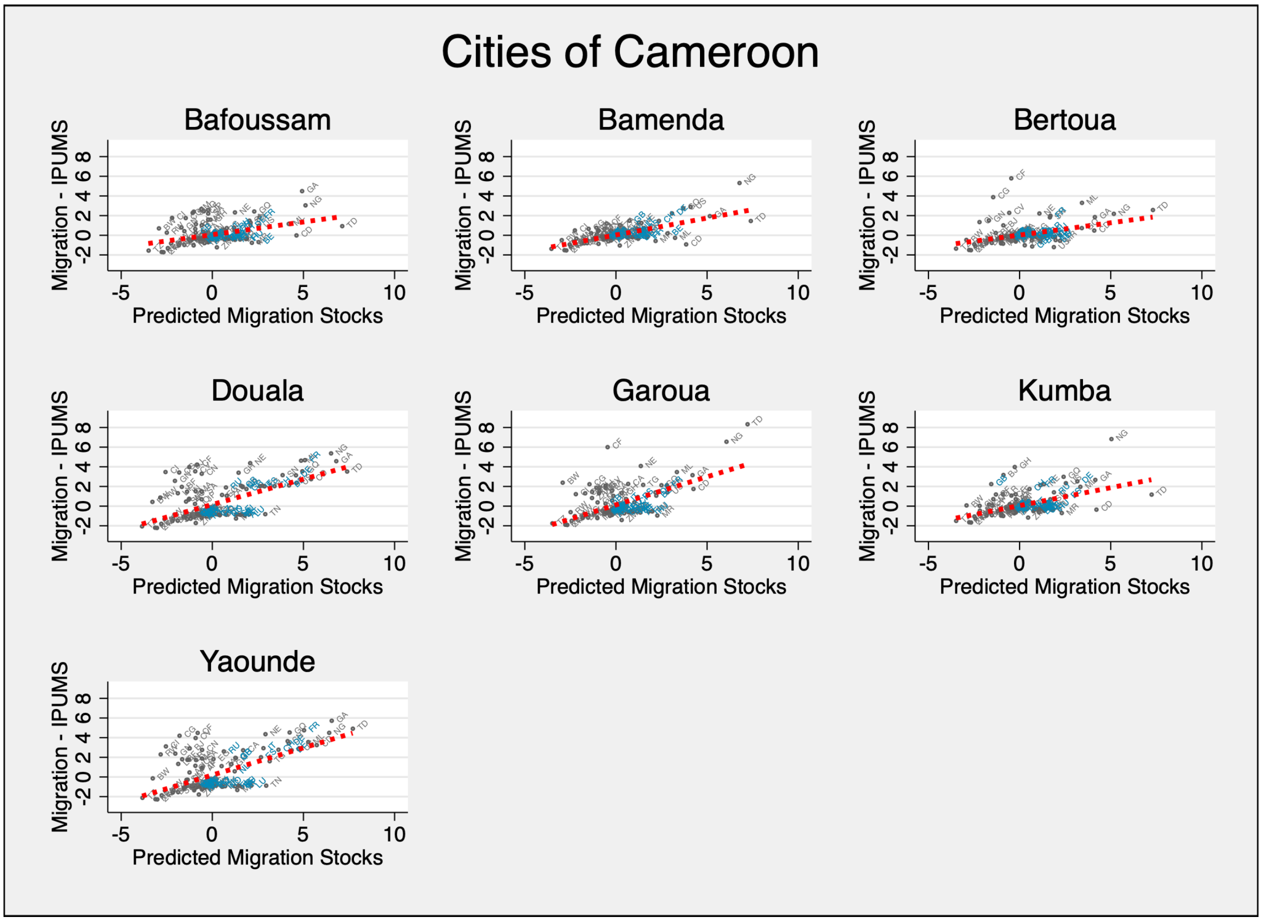
Notes: See notes to Figure A.5.

Figure A.7: Outmigration versus Google searches across country of origin - Destination countries in Asia and Latin America



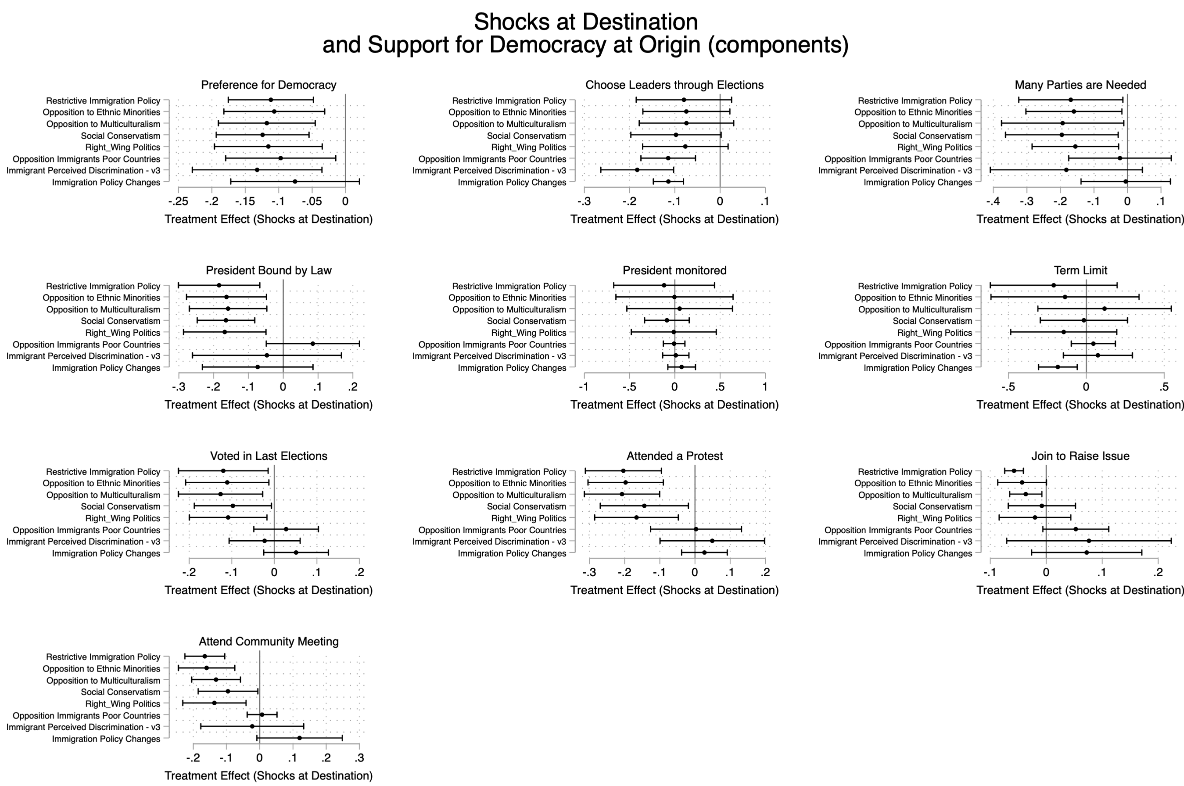
Notes: See notes to Figure A.5.

Figure A.8: Sub-national Validation - Cities of Cameroon



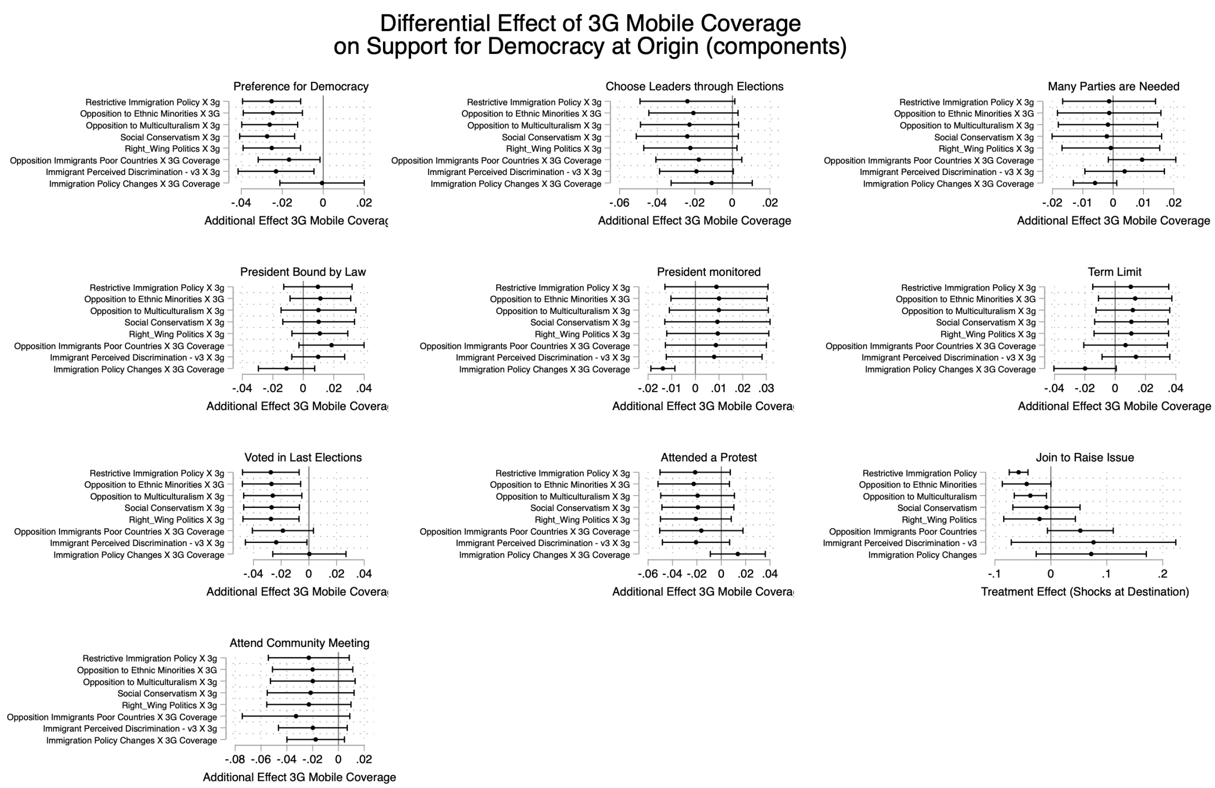
Notes: The figure reports log migration stocks from admin1 regions of Cameroon to all destinations vis a vis average log Google searches for these regions in all destinations. Searches are standardized relative to corresponding values for Switzerland. Data are residuals from regressions on city fixed effects and country of destination fixed effects.

Figure A.9: The effect of exposure to anti-immigration sentiment via migrant networks on Afrobarometer respondents: Individual components



Notes: The Figure reports the estimated coefficients β with 95 percent confidence intervals from equation (4). We report the coefficients separately for each component of the main outcomes variables. For each component, we report the results for the eight measures of exposure to anti-immigration sentiment in destination countries.

Figure A.10: The effect of exposure to anti-immigration sentiment via migrant networks on Afrobarometer respondents: Heterogeneous effect by mobile phone coverage at origin: Individual components



Notes: The Figure reports the estimated coefficients β_2 with 95 percent confidence intervals from equation (6). We report the coefficients separately for each component of the main outcomes variables. For each component, we report the results for the eight measures of exposure to anti-immigration sentiment in destination countries.

Table A.1: Politics at Destination: Discrimination and Satisfaction with Democracy - EU immigrants

Trait:	Right-Wing Politics (1)	Social Conservatism (2)	Restrictive Immigration Policy (3)	Opposition to Ethnic Minorities (4)	Opposition to Multiculturalism (5)
	Dependent variable: Perceived discrimination				
EU immigrant	0.052	0.115*	0.074	0.082	0.134**
	(0.067)	(0.065)	(0.068)	(0.068)	(0.056)
Trait × EU immigrant	0.012	-0.114	-0.025	-0.049	-0.130
	(0.111)	(0.105)	(0.111)	(0.111)	(0.085)
Observations	174,428	174,428	174,428	174,428	174,428
(EU immigrants)	(6,205)	(6,205)	(6,205)	(6,205)	(6,205)
	Dependent variable: Satisfaction with democracy				
EU immigrant	0.049	0.205***	-0.003	0.120*	0.081
	(0.077)	(0.057)	(0.069)	(0.063)	(0.062)
Trait × EU immigrant	0.102	-0.201**	0.170*	-0.034	0.040
	(0.127)	(0.089)	(0.094)	(0.100)	(0.088)
Observations	169,021	169,021	169,021	169,021	169,021
(EU immigrants)	(5,981)	(5,981)	(5,981)	(5,981)	(5,981)
Region FE	✓	✓	✓	✓	✓
Country X Year FE	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓

Notes: Higher values of the trait indicate more Communitarian politics. The specifications include NUTS2 FE, country X year FE, and the following individual controls: age, education, dummies for marital status, unemployed, big city, gender, religion, country of birth. Source: European Social Survey. ***,**,*: statistically significant at 1%, 5% and 10%, respectively.

Table A.2: Contact with relatives abroad

	EUmagine (1)	MAFE (2)
	Have you been in contact during the last 12 months?	
Yes	94%	96%
No	6%	4%
	How often have you been in contact?	
Every day	5%	–
Every week	28%	33%
At least once a month	77%	71%
Observations	1,452	4,723

Notes: The exact wording of the question is the following. EUmagine: "During the last 12 months, how often have you had contact ((spoken, written, sms) with this person [your relative abroad]"; MAFE: "How often have you or anyone else in the household been in contact with this person [your relative abroad] over the last 12 months, by phone, mail, Internet?"

Table A.3: Country-level correlations between bilateral migration stocks and GT searches - separately by area of destination.

	(1) Europe	(2) Asia	(3) Latin-America	(4) North-America	(5) Oceania
GT searches	1.675***	1.182***	0.568	-0.240	1.046***
	(0.248)	(0.125)	(0.374)	(0.231)	(0.002)
Observations	1,883	2,999	237	162	133
Country of origin FE X Year	✓	✓	✓	✓	✓
Country of dest FE X year	✓	✓	✓	✓	✓

Notes: The table reports the same estimates as in column 5 of Table 2 separately by destination continent.

Table A.4: Country-level correlations between bilateral migration stocks and GT searches

	(1)	(2)	(3)	(4)	(5)	(6)
GT searches	1.287***	1.187***	0.942***	0.878***	0.808***	0.781***
	(0.201)	(0.206)	(0.174)	(0.157)	(0.162)	(0.154)
Common language		0.663***	0.782***	0.779***	0.664***	0.612***
		(0.245)	(0.206)	(0.188)	(0.161)	(0.200)
Distance			-0.365***	-0.292***	-0.296***	-0.299***
			(0.063)	(0.054)	(0.053)	(0.054)
Log Trade Volume				0.146***	0.142***	0.140***
				(0.022)	(0.021)	(0.021)
Colony of destination ever					1.504***	0.934**
					(0.238)	(0.391)
Colony of origin ever					0.493	0.252
					(0.302)	(0.254)
Common colonizer						-0.379*
						(0.215)
Common legal origin						0.691*
						(0.362)
Observations	2,062	2,062	2,062	2,015	2,015	2,015
Country of origin FE	✓	✓	✓	✓	✓	✓
Country of dest FE	✓	✓	✓	✓	✓	✓

Notes: The table reports GLS estimates of equation (3), with weights equal to population at origin. Searches are standardized to corresponding values for Switzerland. Two-way standard errors clustered by country of origin and destination reported in parenthesis. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table A.5: Estimates - All Shocks at Destination on Main Outcomes

	Preference for Democracy		Preference for Electoral Competition (PCA)		Preference for Rule of Law (PCA)		Civic Participation (PCA)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Shock at destination: Restrictive Immigration Policy								
Shock	-0.138**	-0.114**	-0.154*	-0.139**	-0.483	-0.420	-0.179*	-0.169
	(0.053)	(0.048)	(0.080)	(0.060)	(0.330)	(0.341)	(0.095)	(0.099)
Shock × 3G Coverage		-0.026**		-0.024*		0.009		-0.033**
		(0.010)		(0.014)		(0.017)		(0.015)
Panel B. Shock at destination: Opposition to Ethnic Minorities								
Shock	-0.129**	-0.108*	-0.142*	-0.130**	-0.271	-0.231	-0.168*	-0.160
	(0.054)	(0.052)	(0.074)	(0.056)	(0.329)	(0.346)	(0.091)	(0.094)
Shock × 3G Coverage		-0.026**		-0.022		0.013		-0.030*
		(0.010)		(0.014)		(0.018)		(0.014)
Panel C. Shock at destination: Opposition to Multiculturalism								
Shock	-0.140**	-0.121**	-0.172*	-0.154**	-0.065	-0.016	-0.142	-0.141
	(0.058)	(0.056)	(0.089)	(0.068)	(0.338)	(0.346)	(0.091)	(0.097)
Shock × 3G Coverage		-0.027**		-0.024		0.011		-0.029*
		(0.009)		(0.014)		(0.017)		(0.015)
Panel D. Shock at destination: Social Conservatism								
Shock	-0.151**	-0.129**	-0.196**	-0.173**	-0.222	-0.191	-0.092	-0.087
	(0.056)	(0.050)	(0.088)	(0.069)	(0.198)	(0.201)	(0.105)	(0.112)
Shock × 3G Coverage		-0.029***		-0.025		0.010		-0.030*
		(0.009)		(0.015)		(0.017)		(0.015)
Panel E. Shock at destination: Right-Wing Politics								
Shock	-0.138**	-0.119**	-0.144*	-0.133**	-0.305	-0.269	-0.137	-0.133
	(0.057)	(0.055)	(0.072)	(0.055)	(0.250)	(0.261)	(0.095)	(0.099)
Shock × 3G Coverage		-0.026**		-0.023		0.010		-0.033**
		(0.010)		(0.014)		(0.017)		(0.015)
Panel F. Shock at destination: Opposition to Migrants Poorer Countries								
Shock	-0.107**	-0.104*	-0.094*	-0.072	0.071	0.073	0.018	0.029
	(0.043)	(0.052)	(0.052)	(0.059)	(0.095)	(0.088)	(0.039)	(0.035)
Shock × 3G Coverage		-0.018		-0.011		0.010		-0.040*
		(0.013)		(0.011)		(0.014)		(0.019)
Panel G. Shock at destination: Immigrant Perceived Discrimination								
Shock	-0.070	-0.077	-0.003	-0.021	-0.090	-0.077	-0.084	-0.077
	(0.063)	(0.067)	(0.081)	(0.080)	(0.103)	(0.102)	(0.100)	(0.104)
Shock × 3G Coverage		-0.026**		-0.022		0.006		-0.027*
		(0.009)		(0.013)		(0.017)		(0.014)
Observations	215,648	215,648	223,873	223,873	191,537	191,537	213,694	213,694
Region origin FE	✓	✓	✓	✓	✓	✓	✓	✓
Country orig X Country dest FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A.6: Migration Survey and Census Data by Source

Panel A: World Bank MRS

Country	# Regions	# Destinations	# Migrants
Burkina Faso	7	14	1704
Ethiopia	8	11	123
Kenya	8	35	1089
Nigeria	13	16	779
Senegal	11	24	1292
Uganda	44	18	381
Total	91	59	5368

Panel B: IPUMS1

Country	# Regions	# Destinations	# Migrants1 (Migrants5)
Botswana	13	19	3831 (5504)
Burkina Faso	13	11	6186
Cameroon	7	33	(10271)
Kenya	8	20	4103
Mozambique	10	24	5686 (6215)
Tanzania	22	12	2271
Zambia	8	16	4565
Total	81	64	26642 (21989)

Panel C: IPUMS2

Country	# Regions	# Destinations	# Migrants
Benin	11	21	59074
Cameroon	7	40	26438
Malawi	26	38	17620
Mali	8	57	46635
Morocco	14	33	6836
Rwanda	5	19	27325
Togo	3	18	32226
Uganda	35	10	43455
Total	109	97	259698

Panel D: IPUMS South Africa

Country	# Regions	# Destinations	# Migrants
South Africa	9	51	27842

Panel E: MAFE

Country	# Regions	# Destinations	# Migrants
DR Congo	26	50	1059
Ghana	10	60	958
Senegal	13	48	1158
Total	49	85	3175

Panel F: EUMAGINE

Country	# Regions	# Destinations	# Migrants
Morocco	3	17	1184
Senegal	4	36	865
Total	7	41	2049

A Appendix

In this section we provide details on the GT data, the data extraction procedure and the methodology used to predict migration stocks from subnational areas of origin in Africa to destination countries.

A Google Trends Data

GT queries return the interest for a certain topic based on Google searches. A query requires a user to specify one or more search terms, one or more search countries and the reference time period (spanning from four hours to multiple years). Queries can be performed for a combination of up to five different terms and countries.[14] The output is provided by sub-periods (i.e. hours, days, weeks or months depending on the length of the period searched).

As mentioned, GT data provide *relative* rather than *absolute* volumes of searches. In particular, for each query, the data are expressed relative to the total volume of searches in the respective country(ies) and further standardized to the maximum value of searches across all specified terms, locations and sub-periods in the query (set to 100). The relative importance of a search term is thus expressed as an integer value from 0 to 100.

Formally let S_{jdt} denote the volume of Google searches for term j performed in country d in subperiod t relative to the total volume of Google searches performed in that country and subperiod. A GT query for the term j in country d over period $T = \{t_1, t_2, ..t_k\}$ will return one value for each subperiod t, namely $\frac{S_{jdt}}{Max\{S_{jdt_1}, S_{jdt_2}, ...S_{jdt_k}\}}$.

Figure A.11 for example shows the results of a query for the term "Accra", the capital of Ghana in the UK over the period January 1st 2015 to December 31st 2020. The resulting data are at the monthly level. The figure shows that the maximum number of relative searches for this term was performed in JAN 2020. Compared to this base period, relative searches for the same term, say, in January 2015 were 63%.

As said, GT also allows comparing two or more search terms in the same country. A GT query for the terms j and z over period T in country d will return two values for each subperiod, respectively $\frac{S_{jdt}}{\tilde{S}_{jzdT}}$ and $\frac{S_{zdt}}{\tilde{S}_{jzdT}}$ where $\tilde{S}_{jzdT} = Max\{S_{jdt_1}, S_{jdt_2}, ...S_{jdt_k}, S_{zdt_1}, S_{zdt_2}, ...S_{zdt_k}\}$. Taking the ratio between the two, one will hence be able to recover relative searches for term j relative to term z in country d in each subperiod, namely $\frac{S_{jdt}}{S_{zdt}}$.

For example, Figure A.12 provides the results of a search for the search term "Accra" relative to the search for the term "Kumasi", the regional capital of the Ashanti region, also in Ghana, performed in the UK over the same period. One can see that relative searches for the term "Kumasi" in February 2020 are 11% of those for the term "Accra". And that this number is 8% in January 2015.

Finally, GT also allows to compare searches for the same term across two or more countries. A GT query for term j in countries d_1 and d_2 will return two values for each subperiod t, respectively $\frac{S_{jd_1t}}{\tilde{S}_{jd_1d_2T}}$ and $\frac{S_{jd_2t}}{\tilde{S}_{jd_1d_2t}}$, where $\tilde{S}_{jd_1d_2t} = Max\{S_{jd_1t_1}, S_{jd_1t_2}, ...S_{jd_1t_k}, S_{jd_2t_1}, S_{jd_2t_2}, ...S_{j_2dt_k}\}$. Taking the ratio between the two, one will hence be able to recover relative searches for the

[14] Note if the absolute number of searches for a term is below an undisclosed threshold this appears as a zero in the data. This implies that we are typically unable to examine searches performed at subnational level (i.e. regions) in countries of destination.

term j across the two countries in each subperiod, namely $\frac{S_{jd_1t}}{S_{jd_2t}}$.

For example, Figure A.13 provides the results of a search for the term "Accra" in the UK and Switzerland. The Figure illustrates that in February 2020 searches for the term Accra in Switzerland (relative to the overall number of searches performed in Switzerland) - were 54% of searches performed in the UK (again standardized to the total number of searches in the UK).

B Measuring relative interest in African regions

In order to extract Google searches for African capital regions for the purpose of predicting subnational migration stocks we proceed in two steps.

First, for each of the 50 African countries we perform pairwise GT queries for any regional capital against the country capital, what we label within-country-of-destination relative searches. We do this separately for all 133 destination countries in the world (not only Europe). If S_{r_od} denote the number of searches for region r_o of country of origin o performed in country d relative to total searches in d, and $x = ln(X + 1)$, this allows us to identify $s_{r_od} - s_{r_o^*d}$, where r_o^* denotes the capital region of country o.

Second, we perform pairwise extractions of GT searches for any region of Africa performed in any destination country against searches for the same region performed in a numeraire country, Switzerland.[15] This allows us to identify $s_{r_od} - s_{r_oD}$, where D denotes the numeraire country, what we label between-country-of-destination relative searches.

We extract data for the period January 1st 2004 to December 31st 2020.[16] As the relative popularity of a given search term in a single extraction is calculated on a random sample of Google searches we also iterate the procedure ten times to reduce the extent of measurement error. The process results in around 280k extractions. We average relative searches from each of the two steps across these ten iterations and all months to derive average relative searches across the entire period.

C Predicting migration stocks from subnational areas of origin in Africa

Armed with Google searches, we use these to predict migration stocks. In particular, let M_{r_od} denote the number of migrants from region r_o of country of origin o residing in destination country d and let $M_{od} = \Sigma_{r\in o} M_{r_od}$ be the total number of migrants from origin country o in d.

We assume the following "structural" model relating log searches for region r_o performed

[15] In practice we split this into two sub-steps. We do so because, when the volume of searches for one the terms or country in the data is comparatively low, GT approximates this to zero. In particular, for each national capital city in the 50 countries of origin, we perform pairwise extractions of its GT searches in any country against its searches in Switzerland. We combine this with information from the first step to derive relative searches for a certain regional capital in a given country relative to its searches in Switzerland.

[16] In practice, we perform three separate extractions over thee partially overlapping time periods: 2004-2010, 2010-2015 and 2015-2020 and we link these three data sets to express all searches relative to the maximum over the period and search term.

in country d to log migrants from region r_o residing in country d:[17]

$$s_{r_o d} = \beta m_{r_o d} + f_{r_o} + f_d + u_{r_o d} \tag{7}$$

where β reflects migrants' propensity to search for their own place of origin. The model includes f_{r_o}, i.e., region of origin fixed effects, which capture the relative popularity of a region across all countries, irrespective of the destination country, and f_d, i.e., country of destination fixed effects, which we assume are the same across origin countries/regions and that capture overall differences in search behavior across destination countries.

In order to estimate subnational migration stocks we proceed in three steps.

First, we derive an estimate of the best linear prediction of migration on Google searches by aggregating data across at the level of country of origin X country of destination. If N_o denotes the number of regions in o and $\tilde{s}_{od} = \Sigma_{r \in o} \frac{s_{r_o d}}{N_o}$ denotes average log searches for all regions of country o in destination country d, aggregating (8) across regions, and assuming that $\Sigma_{r \in o} \frac{s_{r_o d} - s_{od}}{N_o} \approx k_o + k_d$, where the left hand side term is a measure of entropy of country o's migrants' in terms of their sub-national origin in country d, it follows that:

$$\tilde{s}_{od} = \beta m_{od} + f_o + f_d + u_{od} \tag{8}$$

And the best linear prediction of m_{od} given $\tilde{s}_{od}$ is:

$$m_{od} = \psi \tilde{s}_{od} + \kappa_o + \kappa_d + e_{od} \tag{9}$$

Although equation (9) cannot be estimated directly on GT data, as we have no information on absolute searches, we can recover an estimate of ψ using between-country-of-destination variation in searches. In practice, as long as we include country of origin fixed effects, rescaling the right hand side variable for searches performed in the numeraire country, will leave the model specification unchanged. In particular:

$$m_{od} = \psi(\tilde{s}_{od} - \tilde{s}_{oD}) + \delta_o + \delta_d + e_{od} \tag{10}$$

where D denotes a numeraire country (Switzerland).

Second, we exploit within-country-of-origin variation to recover estimates of region of origin fixed effects (a place "popularity"). In particular, from equation (7) it follows that :

$$s_{r_o d} - s_{r_o^* d} = \beta(m_{r_o d} - m_{r_o^* d}) + f'_{r_o} + u'_{r_o d} \tag{11}$$

and

$$s_{r_o d} - s_{r_o^* d} = f'_{r_o} + f_d + u_{r_o d}, \quad M_{r_o d} = 0, M_{r_o^* d} = 0 \tag{12}$$

from which:

$$\hat{f}'_{r_o} = \Sigma_{d:M_{od}=0} \frac{s_{r_o d} - s_{r_o^* d}}{\bar{N}_{od}} \tag{13}$$

In practice, one can obtain estimates of f'_{r_o} by taking averages of within-country-of-destination relative searches, $s_{r_o d} - s_{r_o^* d}$ (step 1 in the previous section), across the $\bar{N}_{od}$

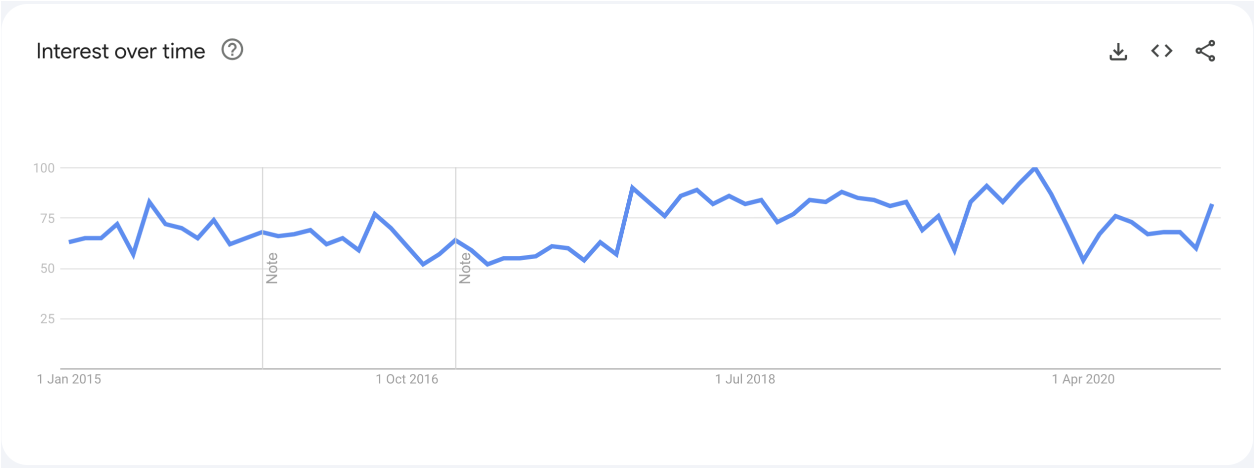
[17] One can derive this model from an underlying model of behavior where users maximize the utility from online searches, subject to a time constraint and where migrants derive a greater utility than other users from searching for their region of origin.

countries for which $M_{od} = 0$ (and hence, a fortiori, $M_{r_o d} = M_{r_o^* d} = 0$).

With estimates of f'_{r_o} and ψ, one finally can recover estimates of the number of migrants from region of origin r_o to destination country d, $M_{r_o d}$. These can simply be obtained by re-apportioning migrants from country o to country d, which one can derive from aggregate statistics, based on within-cunty-of-destination relative searches. In formulas, this is:

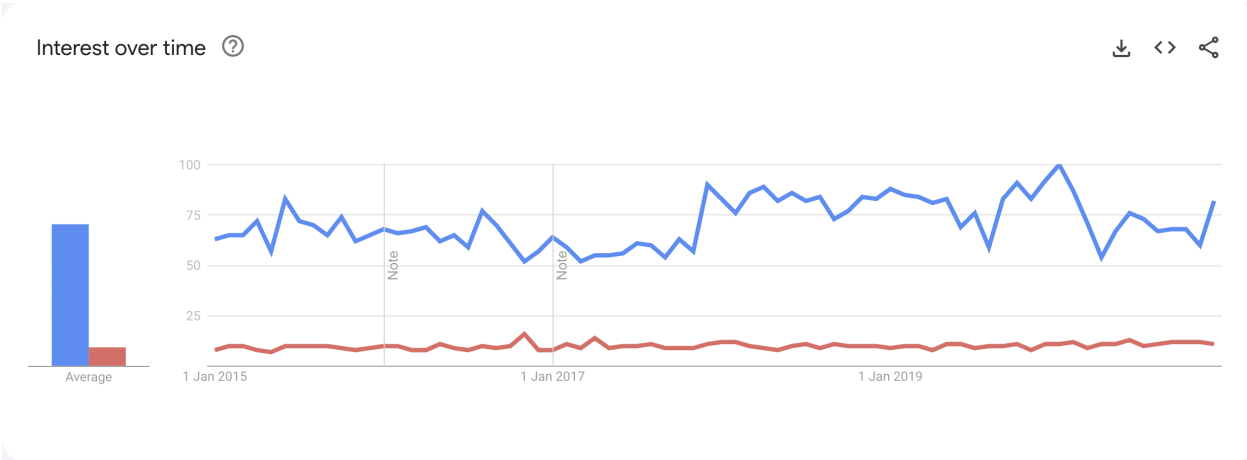
$$\widehat{M}_{r_o d} = \frac{exp(s_{r_o d} - s_{r_o^* d} - \hat{f}'_{r_o})^{\hat{\psi}}}{\Sigma_{r \in o} exp(s_{r_o d} - s_{r_o^* d} - \hat{f}'_{r_o})^{\hat{\psi}}} M_{od} \tag{14}$$

Figure A.11: Google Searches for the Term "Accra" in the UK



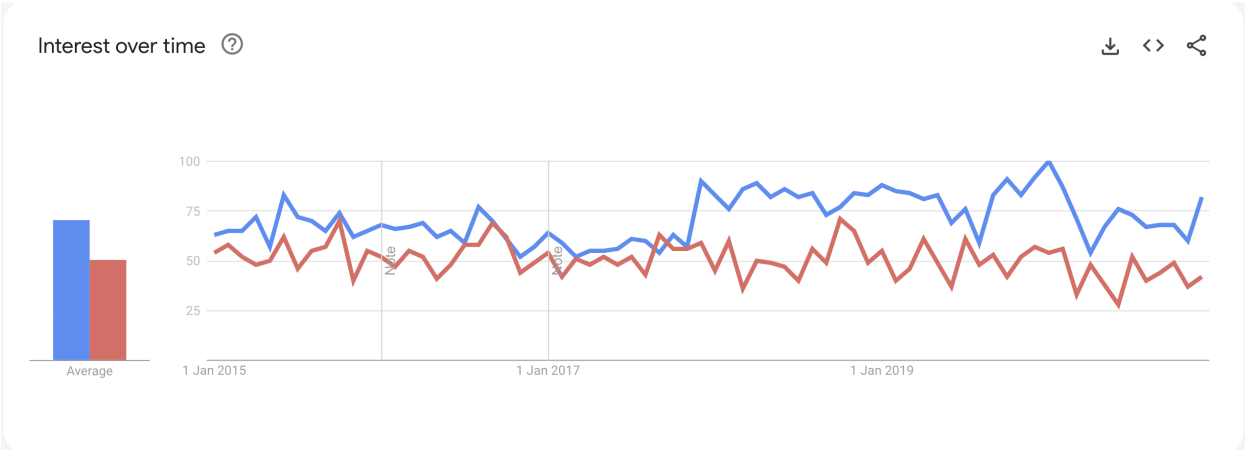
Notes: The figure reports the results of a Google Trends query for searches of the term "Accra (capital of Ghana)" performed in the UK between January 1st 2015 and December 31st 2020.

FIGURE A.12: GOOGLE SEARCHES FOR THE TERMS "ACCRA" AND "KUMASI" IN THE UK



Notes: The figure reports the results of a Google Trends query for searches of the terms "Accra (capital of Ghana)" and "Kumasi (city in Ghana)" performed between January 1st 2015 and December 31st 2020.

Figure A.13: Google Searches for the Term "Accra" in the UK and Switzerland



Notes: The figure reports the results of a Google Trends query for searches of the term "Accra (capital of Ghana)" performed in the UK and Switzerland between January 1st 2015 and December 31st 2020.