Importing Values*

Marco Manacorda[†] Jacopo Ponticelli[‡] Andrea Tesei[§]

January 3, 2022
PRELIMINARY

Abstract

We propose a novel methodology to estimate migration stocks of individuals from subnational areas of origin to any possible country in the world, based on search engine data performed in destination countries. We use data obtained through this methodology to study "social remittances" from Europe to Africa and in particular to test whether exposure of migrants to discrimination at destination affects support for democracy in origin communities. Our estimates indicate that a one s.d. higher exposure to anti-immigrant sentiment at destination leads to around half a s.d. deterioration in support for democracy at origin.

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

* 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

An influential body of research suggests that outmigration affects home communities and that migrants can be important agents of change in their country of origin. As migration often happens from areas of the world that are poor, with weak institutional capacity and poorly developed democratic institutions, possibly riddled with conflict or affected by natural disasters - all established push factors - migrants can foster economic progress at origin via remittances, as well as facilitate the formation and transmission of values at home (Kapur, 2014).

While there is extensive research on the economic consequences of outmigration, we still have scarce direct empirical evidence on how migration shapes origin communities' norms, values and identities. Existing work suggests that ideas and norms flow from migrants to origin countries through the links they maintain with their communities, what is generally labeled "social remittances" (Levitt, 1998). In particular, a body of literature argues that exposure to democratic values at destination might itself promote democratic progress at home, either via transmission of information to home communities (Barsbai et al., 2017; Batista and Vicente, 2011) or via return migration (Chauvet and Mercier, 2014; Mercier, 2016; Spilimbergo, 2009). However, exposure to destination culture might also change migrants' view of the democratic system, especially when destination societies are unable to promote migrants' integration, well-being and economic success. These phenomena have become increasingly important with the spread of anti-immigrant sentiment in developed economies, as captured by the proliferation and electoral success of right wing, xenophobic populist parties (Dustmann et al., 2019; Halla et al., 2017).

In this paper, we study whether changes in attitudes towards migrants in European destination countries - which are mostly full democracies - affect the view of democracy in their communities of origin in Africa. We focus on migration from Africa as this is by and large the poorest continent in the world, and with no reliable data on outmigration at the subnational level.[1]

The main empirical challenge to identify this transmission of values from migrants to their communities of origin is the absence of systematic information on subnational 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

[1] Despite international mobility among African citizens is still modest when compared to other continents - with 2% of citizens living in country different from the one of birth (a third of which in Europe), largely due to restrictive migration policies at destination,- around a third of respondents in the Afrobarometer report having considered moving abroad and one in five individuals report depending at least a little bit on cash from abroad.

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.

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 the psychological literature on individuals' attachment to their places of origin (Lewicka, 2011), we propose to capture the presence of migrants from a specific region of Africa using internet searches for region-specific terms at destination.[2] 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 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 subnational region of origin of migrants. The majority of this data effectively captures return migration – and not current migration

[2] 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.

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 subnational 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 discrimination at destination across subnational 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 preference for democracy recorded in the communities of origin of African migrants with the exposure of such communities to anti-immigrant sentiment in destination countries. Preference for democracy at origin is measured using questions from the Afrobarometer, which covers 35 African countries in the years between 2003 and 2018, while changes attitudes towards migrants at destination are measured using the European Social Survey (ESS). 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 communities 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 a 0.23 percentage points decline in their probability of supporting democracy. This is a significant effect considering that average support for democracy among Afrobarometer respondents at origin across all years is 73%, with a standard deviation of 0.44. We find consistent effects when using alternative measures of cultural values at destination, such as the share of immigrants perceiving to be discriminated against in destination countries according to the ESS. Consistent with transmission of cultural values, information and ideas along migration networks, we find that perceived discrimination by non-immigrants in destination countries has no significant effect on support for democracy at origin.

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 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.

The rest of the paper is organized as follows. In section 2.1 we describe our methodology to estimate migrant stocks at subnational 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 subnational migration from Africa to Europe, as well as data on return migration. Then, in section 3 we propose an application of the new measure of migration stocks at sub-national level to study the role of social remittances between Europe and Africa and present the preliminary results of the paper.

2 Estimating migration stocks from Google Trends data

2.1 Methododolgy

To fix ideas, let $M_{r(o)d}$ denote the number of migrants from region $r(o)$ of country of origin o residing in destination country d and let $M_{od} = \Sigma_{r \in o} M_{r(o)d}$ be the total number of migrants from origin country o in d. We also denote the volume of Google searches for region $r(o)$ performed in country of destination d, standardized by the total volume of searches in such country, by $S_{r(o)d}$. In what follows, we use lower cases to denote logarithms.[3]

[3] We augment the argument of logarithm by one in order to account for zeros, namely $m_{r(o)d} = ln(M_{r(o)d} + 1)$ and $s_{r(o)d} = ln(S_{r(o)d} + 1)$. We also plan to experiment with the inverse hyperbolic sine transformation.

We assume the following "structural" model relating log searches for region $r(o)$ performed in country d to log migrants from region $r(o)$ residing in country d.[4] In formulas:

$$s_{r(o)d} = \beta m_{r(o)d} + f_{r(o)} + f_d + u_{r(o)d} \tag{1}$$

where β is a proportionality term that reflects migrants' propensity to search for their own place of origin. The model allows for $f_{r(o)}$, 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 to be the same across origin countries/regions and that capture overall differences in search behavior across destination countries. Clearly, it follows that:

$$s_{r(o)d} = f_{r(o)} + f_d + u_{r(o)d}, \quad M_{r(o)d} = 0 \tag{2}$$

In principle one could use model (1) to derive best linear predictions of log migrants stocks from each origin region in each destination country, $m_{r(o)d}$ based on log searches, $s_{r(o)d}$ plus estimated values for the fixed effects.[5] For this, one would need out-of-sample data or data on specific subsamples to then extrapolate to the population of interest. Such data are hardly available but we will show below how one can - under appropriate assumptions - circumvent this problem by using aggregate bilateral migration stocks at the country level.

An additional complication is that, for each query, Google Trends (GT hereafter) provides number of searches (standardized to the total volume by country) relative to the maximum value of such searches across all terms and search locations in the query. We discuss GT data in detail in Appendix section A. This means that one will not be able to recover the absolute volume of searches. However, we show that one will still be able to recover the *relative* volume of searches across search terms - e.g., regions of origin - or across places where the search was performed.

In particular, GT allow users to recover the volume of searches for each region in country o, $r(o)$, relative to a numeraire region of choice, e.g., $r^*(o)$, in each destination country, $s_{r(o)d} - s_{r^*(o)d}$, i.e., *within-country-of-origin relative* searches.[6]

[4] 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.

[5] The problem is similar in spirit to Henderson et al. (2012) who use night light density to predict measure of economic growth in small areas across Africa (as well as to improve on data of economic growth from the WDI).

[6] By performing a pairwise search for regions $r(o)$ and a numeraire (say the capital) region $r^*(o)$ in destination country d, a user can recover an estimate of $ln\frac{S_{r(o)d}/max\{S_{r(o)d}, S_{r^*(o)d}\}}{S_{r^*(o)d}/max\{S_{r(o)d}, S_{r^*(o)d}\}} = s_{r(o)d} - s_{r^*(o)d}$, i.e. searches for each origin country region relative to the capital region separately by destination country. Clearly, because of transitivity, one can always recover such measures by combining other pairwise searches. For

Similarly, GT also allows to obtain estimates of searches for each region in country d relative to searches in a numeraire country D, $s_{r(o)d} - s_{r(o)D}$, i.e., *between-country-of-destination relative* searches.

From (1), it follows that such relative searches can be easily expressed as a function of relative migration stocks plus some fixed effects. In particular:

$$s_{r(o)d} - s_{r^*(o)d} = \beta(m_{r(o)d} - m_{r^*(o)d}) + f'_{r(o)} + u'_{r(o)d} \tag{3}$$

$$s_{r(o)d} - s_{r(o)D} = \beta(m_{r(o)d} - m_{r(o)D}) + f''_{d} + u''_{r(o)d} \tag{4}$$

From (3), it follows that best linear prediction of $m_{r(o)d} - m_{r^*(o)d}$ given $s_{r(o)d} - s_{r^*(o)d}$ and $f'_{r(o)}$ is:

$$m_{r(o)}\widehat{d - m}_{r^*(o)d} = \hat{\psi}(s_{r(o)d} - s_{r^*(o)d} - \hat{f}'_{r(o)}) \tag{5}$$

If one is able to obtain estimates of $f'_{r(o)}$ and ψ, one can then use these, together with GT searches, to recover estimates of $m_{r(o)d} - m_{r^*(o)d}$. With such estimates, one can in turn obtain 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 the following expression:

$$\widehat{M}_{r(o)d} = \frac{exp(m_{r(o)}\widehat{d - m}_{r^*(o)d})}{\Sigma_{c\in o} exp(m_{r(o)}\widehat{d - m}_{r^*(o)d})} M_{od} \tag{6}$$

To operationalize our approach, we proceed in two steps. First, we exploit within-country-of-origin variation to recover estimates of $f'_{r(o)}$. In particular, based on (2), one can recover such estimates by simply taking averages of $s_{r(o)d} - s_{r^*(o)d}$ across the $\bar{N}_{od}$ countries for which $M_{od} = 0$ (and hence, a fortiori, $M_{r(o)d} = 0$). In formulas:

$$\hat{f}'_{r(o)} = \Sigma_{d:M_{od}=0} \frac{s_{r(o)d} - s_{r^*(o)d}}{\bar{N}_{od}} \tag{7}$$

Second, we derive an estimate of ψ by exploiting between-country-of-destination variation. If N_o denotes the number of regions in o and $\tilde{s}_{od} = \Sigma_{r\in o} \frac{s_{r(o)d} - s_{r(o)D}}{N_o}$ is the average relative log searches for all regions of country o in destination country d relative to the numeraire country D, it follows, aggregating (4) across regions, that:[7]

example, one can perform two separate pairwise searches for regions $r(o)$ and $r^*(o)$ in country d relative to searches for a numeraire region R of a numeraire country O in a numeraire destination country D. These searches will provide respectively $s_{r(o)d} - s_{R(O)D}$ and $s_{r^*(o)d} - s_{R(O)D}$. One will be able to recover $s_{r(o)d} - s_{r^*(o)d}$ by simple differentiation,

[7] We assume that $\Sigma_{r\in o} \frac{(m_{r(o)d} - m_{od}) - (m_{r(o)D} - m_{oD})}{N_o} \approx k_o + k_d$, where the left hand side term is a measure of entropy of country o's migrants' in terms of their subnational origin in country d relative to country D.

$$\tilde{s}_{od} = \beta m_{od} + \theta_o + \theta_d + u_{od} \tag{8}$$

One can hence recover estimates of ψ by regressing log bilateral migration stocks on average log bilateral searches, plus fixed effects:[8]

$$m_{od} = \psi \tilde{s}_{od} + \delta_o + \delta_d + e_{od} \tag{9}$$

2.2 Estimation

In the following we present empirical estimates based the procedure described below. We start by identifying, for all 50 African countries in our sample, the most populous city in each admin1 sub-national region. In total, we identify 709 cities, which in most cases correspond to the capital city of the admin1 region. Figure 1 reports a map of Africa with the precise location and name of each of the 709 identified cities.

We use the names of the 709 main admin1 sub-national cities as search terms in GT. More specifically, we perform monthly GT term searches from 2004 to 2020, across 133 destination countries. This allows to recover the terms $s_{r(o)d} - s_{r^*(o)d}$ and $s_{r(o)d} - s_{r(o)D}$. The exact procedure is described in Appendix section A. We complement these data with data at 5 years intervals (2005, 2010, 2015 and 2020) on bilateral migration stocks from the United Nations Population division.[9]

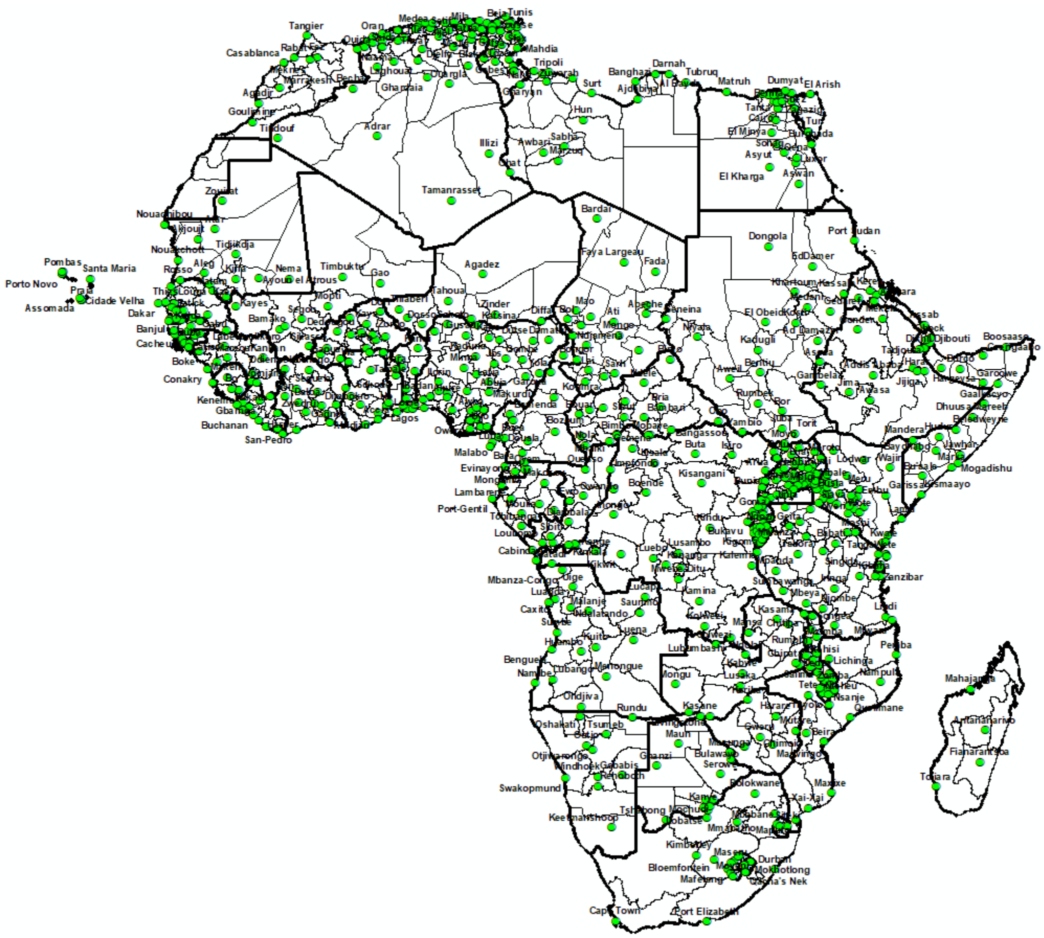
We first compute region fixed effects $f'_{r(o)}$. In order to derive such fixed effects, we restrict to country pairs for which migration stocks relative to the population at destination is zero or unreported by the United Nations, which is likely associated to a small number of migrants. On average, it appears that there are low migration stocks across two-thirds of African country of origin times country of destination dyads. We use expression (7) to derive such fixed effects.

We then turn to estimates of ψ from equation (9). 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, $\tilde{s}_{od}$. Figure 2 reports the results for Algeria, Ghana, Morocco and Nigeria - four large

[8] Formally, from (3), $\psi = \beta \frac{var(m_{r(o)d} - m_{r^*(o)d})}{var(s_{r(o)d} - s_{r^*(o)d} - f'_{r(o)})}$, while from (8) $\psi = \beta \frac{var(m_{od} - m_o - m_d)}{var(\tilde{s}_{od} - \tilde{s}_o - \tilde{s}_d)}$. Ignoring the fixed effects, for the linear prediction of m_{od} on $\tilde{s}_{od}$ to deliver a consistent estimate of ψ one requires that $\frac{var(m_{r(o)d} - m_{r^*(o)d})}{var(m_{od})} = \frac{var(s_{r(o)d} - s_{r^*(o)d})}{var(\tilde{s}_{od})}$, which is equivalent to requiring that the ratio between the within country of origin and the between country of destination variances is the same for migration and GT searches.

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

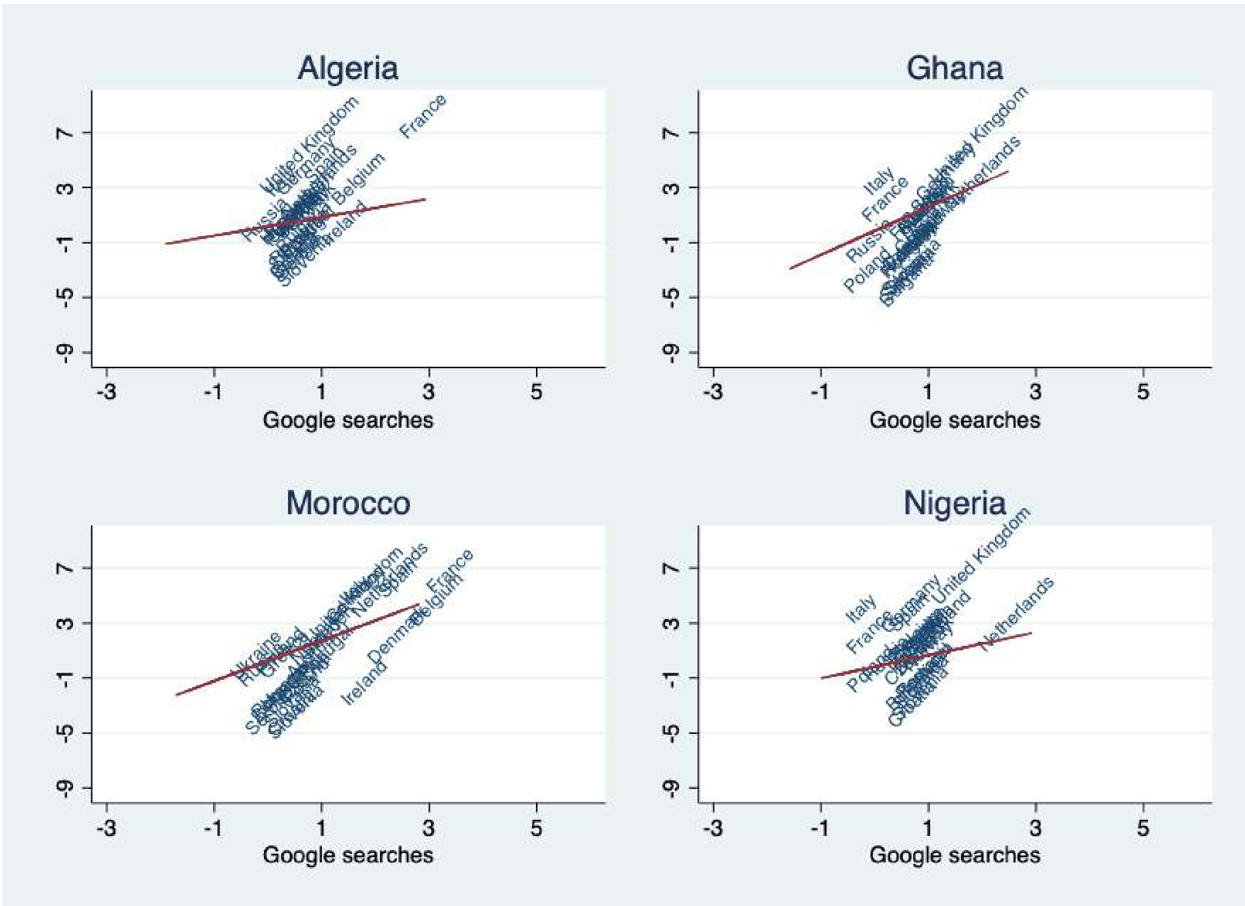
Figure 1: Largest city by Admin1 sub-national region



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.

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. We plot residuals of both variables relative to country of origin fixed effects based on population-at-destination weighted regressions. 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 GT 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 provide estimates, although likely error-ridden, of migration stocks. For completeness, Appendix Figure A.1 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.

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

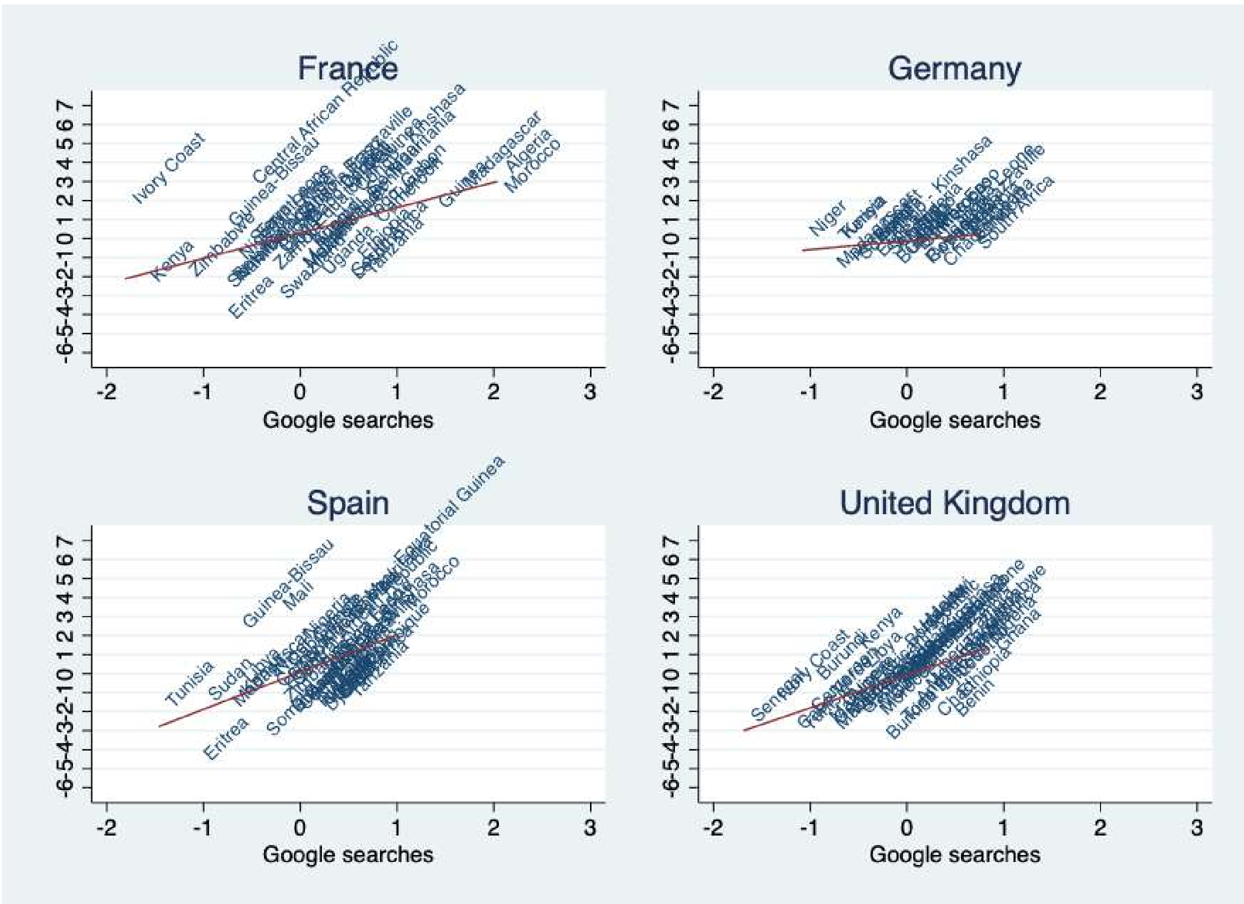


Notes: The figure reports log migration stocks from four African countries to all destinations vis a vis average log Google searches for regions of such origin countries in all destinations. Searches are standardized relative to corresponding values for Switzerland. Data are residuals from regressions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

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, for France, Germany, Spain and the United Kingdom, four large destination countries in Europe, we plot the two series across African origin countries. 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,

For completeness, Appendix Figures A.2 to A.4 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' stocksOverall, it appears that, across most destination countries, African countries with more immigrants are searched more.

Figure 3: 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.

We now turn to regression results in Table 1, which provides estimates of ψ from equation (9). Here we use all data (2005, 2010, 2015 and 2020). We present increasingly saturated models with time effects, country of origin and country of destination fixed effects, which we also interact with time fixed effects. We present regressions weighted by population at origin, although results are very similar if we weight by population at destination. We 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 0.894 and 1.254 and are statistically significant at conventional levels.[10]

Appendix Table A.1 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.2, 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, past colonial links, and common legal origin. In the most demanding specification including all bilateral controls, the magnitude of the point estimate is about 0.688 and precisely estimated.

Table 1: Country-level correlations between bilateral migration stocks and GT searches

	(1)	(2)	(3)	(4)	(5)
GT searches	1.254***	1.230***	0.894***	0.948***	1.069***
	(0.238)	(0.253)	(0.153)	(0.151)	(0.171)
Observations	8,512	8,512	8,512	8,512	8,480
Year FE	✓	✓	✓	✓	✓
Country of origin FE	×	✓	×	✓	✓
Country of dest FE	×	×	✓	✓	✓
Country of origin FE X Year	×	×	×	×	✓
Country of dest FE X year	×	×	×	×	✓

Notes: The table reports GLS estimates of equation (9), 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: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

In the last step, we use estimates of $f'_{r(o)}$ and ψ, together with data on bilateral migration stocks across country dyads, to recover estimates of the stock of migrants from each subnational region in Africa to the rest of the world, $m_{r(o)d}$.

2.3 Subnational Validation

In this section, we turn to validate our measure of predicted subnational migration stocks introduced above. As a preliminary step, Figure 4 reports the top three predicted European

[10] Estimates of β are on the order of 0.11, implying that migrants are 11% more likely than natives to search for regions in their country of origin.

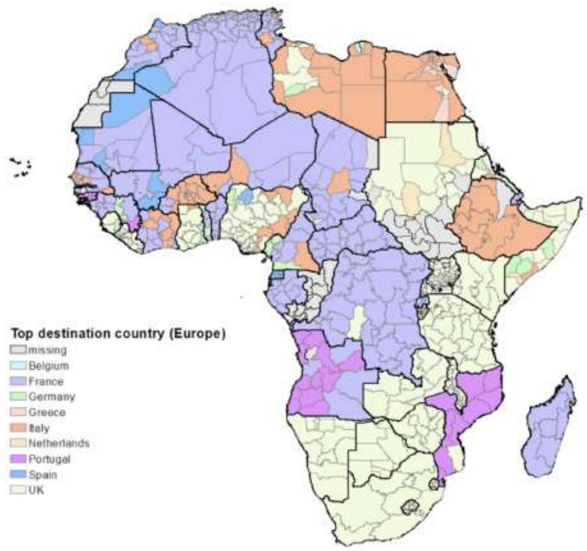
destinations for each of the 709 subnational regions of Africa. The objective of this exercise is twofold. First, to provide a preliminary assessment of the plausibility of the information content of our GT-based measure. Second, to assess the amount of variation generated by the measure.

Panel A refers to the top country of destination, which on average accounts for around 58% of total predicted migration to Europe from African regions. The map clearly hints at the importance of the European colonial legacy, with the West and Central African regions mainly linked to France and the East and Southern African regions 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 amount to 14% and 8% of total European migration, respectively. Compared to panel A, panels B and C display a more diverse set of destinations across subnational regions of the same country, with as many as seven different 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.

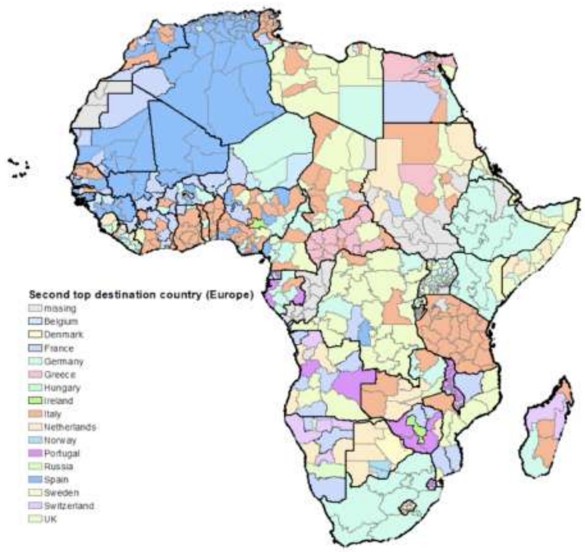
Clearly, the patterns in Figure 4 may not only reflect migration stocks from specific regions, but also broader historical links, as well as trade and cultural ties between areas. To assess the predictive content of the measure more directly for our intended purpose, we compare it with the limited available information on subnational migration stocks at this level of geographical disaggregation. Subnational data on African migration are extremely scarce and only available for a small number of countries and years. 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 obtained from small-scale, specific-purpose surveys. Appendix Table A.3 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. In the following, we focus on measures of subnational 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. As illustrative examples, in Figures 5 and 6 we provide a graphical representation of the relationship between predicted and actual migrants' stocks across subnational regions of Morocco and Cameroon. In particular, we plot log number of return migrants from 1% census samples versus log predicted stock of migrants based on

Figure 4: Top European Destinations

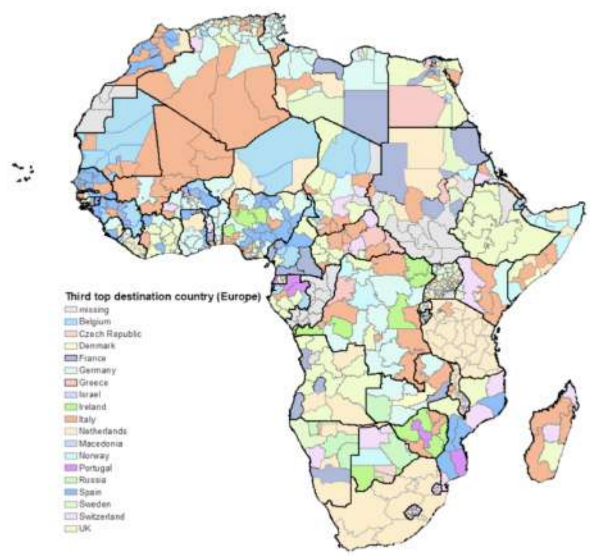
(a) Panel A



(b) Panel B



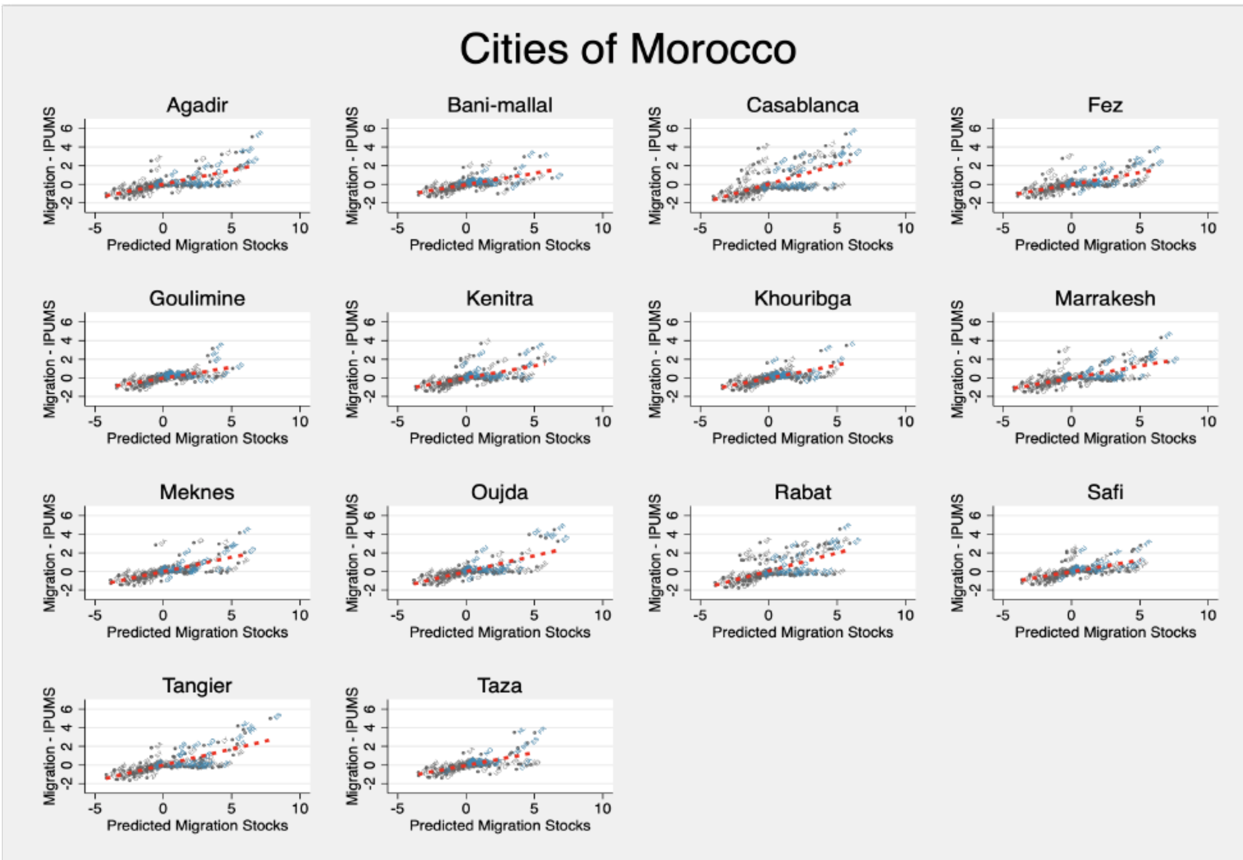
(c) Panel C



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

our methodology, (both augmented by one to account for zeros and) both 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 as well as across regions of the same country. European countries (highlighted in blue) are the main destinations for all regions of Morocco. Particularly prominent are migration stocks in Mediterranean countries such as Italy, France and Spain, but also in some Northern European countries such as Belgium, the Netherlands and Germany. More importantly, the degree of migration intensity to these countries varies between regions of origin, and is consistently captured by our predicted measure of migration.

Figure 5: Subnational Validation - Cities of Morocco

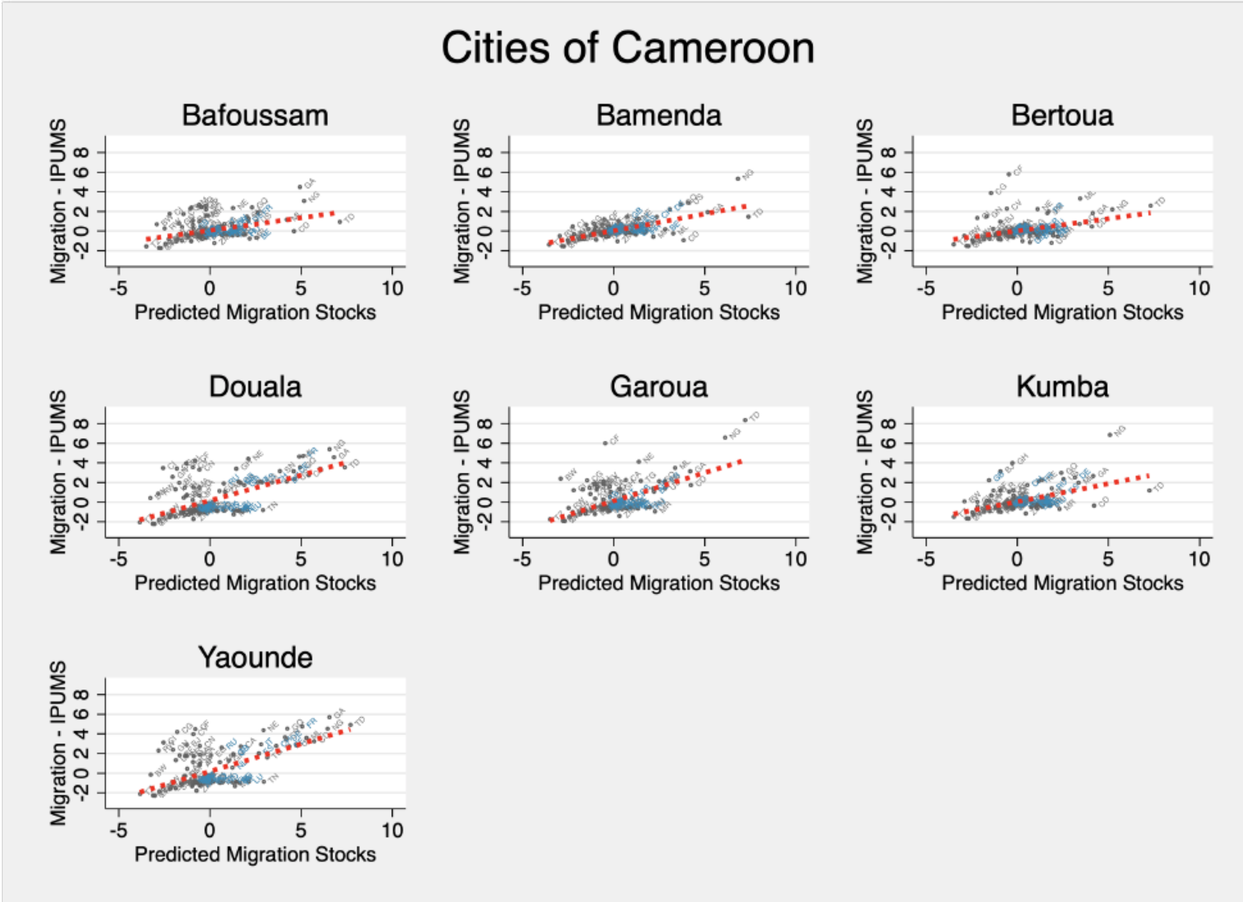


Notes: The figure reports log migration stocks from admin1 regions of Morocco to all destinations vis a vis average log Google searches for these regions in all destinations. Searches are standardized relative to corresponding values for Switzerland. Data are residuals from regressions on city fixed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

Also in the case of Cameroon, it is easy to see the close positive relationship between predicted and actual migration stocks across regions of the country. In this case, the main destinations are neighboring countries like Nigeria, Gabon and Chad. Among European destinations, it is not surprising that France features most prominently. However, it is also interesting to note the high levels of both 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 2 we turn to formally estimating the relationship between return migration from

Figure 6: Subnational 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.

the national censuses for Africa and predicted subnational migration stocks from Africa to European destinations. The units of analysis are all "region of origin-country of destination" pairs for which we have information about predicted and actual migration. We present both unweighted results and results weighted by population at origin or at destination. The odd columns of Table 2 present the regression counterparts of Figures 5 and 6, where the estimates refer to a model that includes region fixed effects and country of destination fixed effects. This allows us to controls for variation in predicted and actual migration from a given region that is due to its size or its popularity across all destinations, as well as destination-specific characteristics that may affect the level of migration (both actual and predicted) from all regions of Africa. Column (1) reports estimates across all world destinations. Point estimates vary between 0.21 and 0.29 depending on the weighting scheme adopted and are statistically significant at conventional levels. We find similar results when estimating coefficients separately across groups of destination countries by continent, except for North America and Oceania, for which, however, IPUMS data show that migration from Africa is very limited.

One concern with these estimates is that part of the variation stems from cross-country-pairs variation, which may reflect time invariant links between origin and destination pairs, such as their colonial history, common language and trade links. For this reason, in even columns of Table 2, we focus on a saturated model that replaces country of destination fixed effects with country of origin times country of destination fixed effects. This absorbs all time-invariant origin-destination links and allows us to focus on relative patterns of actual and predicted sub-national migration net of bilateral country ties. The results are remarkably similar to those in the more parsimonious model. The estimates are statistically significant at conventional levels, both overall and by continent, with the exception of North America and Oceania. If anything, when focusing on the more saturated model, the fit between predicted and actual migration stocks for European destinations improves. Overall, we take the evidence in Table 2 as strongly supportive of the hypothesis that our GT-based measure of migration contains valuable information about actual migration stocks from subnational regions of Africa, which are typically very hard to estimate systematically.[11]

[11] Appendix Tables A.4 and A.5 present results analogous to those in Tables 2 but where the actual migration stocks are obtained from available Migration Surveys (World Bank, MAFE, EUMAGINE).

Table 2: Relationship between log return migration from African National censuses and predicted migration stocks from Google trends, by African regions

	Dependent Variable: Log IPUMS													
	Any destination		Africa		Europe		Asia		Latin-America		North-America		Oceania	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Unweighted													
Log Prediction	0.291***	0.295***	0.326***	0.385***	0.084**	0.244**	0.243***	0.214***	0.097**	0.085*	0.006	0.033	-0.003	0.021
	(0.026)	(0.041)	(0.030)	(0.058)	(0.035)	(0.089)	(0.044)	(0.048)	(0.039)	(0.041)	(0.025)	(0.062)	(0.011)	(0.073)
	Weighted by Population at Destination													
Log Prediction	0.219***	0.245***	0.354***	0.388***	0.060	0.257***	0.266***	0.240***	0.178***	0.121**	0.007	0.033	-0.005	0.021
	(0.053)	(0.041)	(0.046)	(0.065)	(0.036)	(0.091)	(0.053)	(0.049)	(0.052)	(0.054)	(0.013)	(0.037)	(0.011)	(0.062)
	Weighted by Population at Origin													
Log Prediction	0.265***	0.274***	0.299***	0.361***	0.072**	0.240***	0.274***	0.212***	0.089**	0.077*	0.026	0.081	-0.017	0.041
	(0.024)	(0.040)	(0.030)	(0.064)	(0.034)	(0.083)	(0.044)	(0.048)	(0.035)	(0.039)	(0.026)	(0.054)	(0.008)	(0.084)
Observations	30,420	30,420	10,998	10,998	6,552	6,552	8,190	8,190	3,744	3,744	468	468	468	468
Region origin FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country dest 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: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

3 Importing Values via Migration Networks

3.1 Setting and empirical strategy

In this section, we propose an application of our new measure of migration at sub-national level to study the role of social remittances between Europe and Africa. We think of social remittances as "ideas, behaviors, identities, and social capital that flow from receiving- to sending-country communities" via the network of migrants (Levitt, 1998). We focus on social remittances from African migrants living in Europe to their communities of origin in Africa. In particular, we are interested in studying the following question: do attitudes towards migrants in European destination countries – which are mostly full democracies – affect the view of democracy in their communities of origin in Africa? Are communities of origin less likely to consider democracy as a desirable form of government when their migrants are more subject to discrimination in democratic regimes?

We focus on this question for several reasons. First, the last two decades have seen a large increase in migration from Africa to Europe, with immigrants being perceived in different ways in different destination countries. Second, while countries of destination in Europe are mostly consolidated democracies, countries of origin in Africa rank lower in international assessments of democratic institutions.[12] This institutional gap offers an opportunity to study potential transmission of attitudes towards democratic values along migration networks, and to explore whether anti-immigrant sentiment within Europe can affect demand for democratic institutions in countries that send migrants to Europe.[13] Finally, addressing this question requires data on migration stocks at subnational level. The methodology described in section 2.1 allows us to reconstruct estimates of migration links between all sub-national units in Africa and destination countries. We use these links to compute the exposure of communities of origin in Africa to different destination countries in Europe.

We estimate individual level regressions linking attitude towards democracy recorded in the communities of origin of African migrants with the exposure of such communities to cultural values at destination via migration networks. Attitude towards democracy is measured using questions from the Afrobarometer survey.[14] Our baseline specification is as

[12] One metric is the Polity score generated by the Center for Systemic Peace (`www.systemicpeace.org/polity/polity4.htm`), which ranges from +10 (strongly democratic) to -10 (strongly autocratic). In 2015, the average score among the top-10 European destination countries of African migrants was 9.7, while it was 3.5 among the top-10 African countries by number of migrants to Europe.

[13] See on this the related work of Barsbai et al. (2017) and Docquier et al. (2016).

[14] 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 2019, Afrobarometer conducted and completed 7 waves of this repeated cross-section survey. The last round was completed in

follows:

$$SupportDemocracy_{ir(o)t} = \alpha_{r(o)} + \alpha_{ot} + \beta \underbrace{\sum_{d \in D} \omega_{r(o)d} CulturalValues_{dt}}_{Exposure} + \varepsilon_{ir(o)t} \tag{10}$$

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 ($\alpha_{r(o)}$), and country of origin interacted with time fixed effects capturing country-level aggregate trends (α_{ot}). African regions are the 709 admin1-level regions located in 51 African countries described in section 2.1.

The outcome variable $SupportDemocracy_{ir(o)t}$ is individual i's answer to Afrobarometer questions on the level of support for democracy. In our baseline specification, we codify the outcome as a dummy variable equal to 1 if the respondent declares that democracy is preferable to any other form of government, and 0 if the respondent says that it either does not matter or that sometimes a non-democratic government can be preferable.[15]

$CulturalValues_{dt}$ is a time varying measure of cultural values in destination country d. We are interested in understanding whether attitudes towards migrants in democratic destination countries shape migrants' view of democracy. Notice that European destination countries in our sample are almost exclusively consolidated democracies. Thus, migrants might update their beliefs about the desirability of the democratic system depending on how they are treated *in* such a system. Our empirical strategy is designed to investigate if changes in these views are then transmitted via the migration network to the migrants' communities of origin. To capture attitudes towards migrants in destination countries we use country-level data from the European Social Survey (ESS). We focus on questions of the ESS capturing the degree of acceptance of immigrants as well as questions capturing whether migrants themselves perceive to be part of a discriminated group.

The weights $\omega_{r(o)d}$ capture the intensity of migration linkages between African region $r(o)$

2018 and reports information on 34 African countries. Each country has a sample size of either 1200 or 2400 individuals. Samples are designed to generate a representative cross-section of all citizens of voting age in a given country.

[15] The potential answers in Wave I are: "Democracy is preferable to any other form of government", "To people like me, it doesn't matter what form of government", "In certain situations, a non-democratic government can be preferable", or "Don't know". The question has been asked in a consistent way across the seven waves of the Afrobarometer, always with the same 4 options. Our outcome variable is based on the following questions (wave number reported in parenthesis): Q38 (Wave II), Q37 (Wave III), Q30 (Wave IV), Q32 (Wave V), Q30 (Wave VI), Q28 (Wave VII).

and destination country d, and are defined as follows:

$$\omega_{r(o)d} = \frac{Migrants_{r(o)d}}{Migrants_{r(o)}} \tag{11}$$

That is, $\omega_{r(o)d}$ captures the share of migrants from African region $r(o)$ that are currently living in destination country d. The *Migrant* variables are constructed using predicted number of migrants from Google Search data and the methodology described in section 2.1. The methodology produces estimates of number of migrants from African region $r(o)$ to destination d for three benchmark years: 2005, 2010, 2015 and 2020. The weights are constructed by predicted migrants in the baseline year 2005 in both the numerator and the denominator, to avoid that endogenous changes in the geography of the migrant network affect our results.

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. Additionally, we present results with an augmented specification that controls for observable characteristics of the location (primary sampling unit, or PSU) of the respondent. These characteristics include: access to electricity, access to running water, presence of a school, a police station, a hospital, or a local market.

One potential concern in the model described by equation (10) is that shocks to attitude towards migrants at destination might be correlated with other shocks at destination (e.g. a recession), which are also transmitted to the communities of origin via the migrant network (e.g. via lower economic remittances), affecting the outcomes of interest. To deal with this issue, and at the cost of over-controlling, we include in our specification a measure of exposure to changes in economic conditions at destination as follows:

$$\text{Exposure to economic conditions at destination} = \sum_{d \in D} \omega_{r(o)d} \Delta GDP_{dt} \tag{12}$$

where $\omega_{r(o)d}$ are the weights described in equation (11) capturing connections between communities of origin and destination countries via migrant networks, and ΔGDP is GDP growth (annual %) in destination country d from the World Development Indicators dataset.

Finally, we study whether variation in the ability to communicate with destination regions across individuals within the same community of origin affect the strength of social remittances. 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.[16] Figure 7 shows a visual example of our dataset for countries in West Africa. As shown, the data also allow us to compare the impact of social remittances on individuals with vs without access to the mobile phone network within the same community of origin in Africa. We test for heterogeneous effects by mobile phone coverage at origin by estimating the following specification:

$$\begin{aligned} SupportDemocracy_{ir(o)t} &= \alpha_{r(o)t} + \beta_1 \sum_{d \in D} \omega_{r(o)d} CulturalValues_{dt} \times 1(Covered)_{ir(o)t} \\ &+ \beta_2 \sum_{d \in D} 1(Covered)_{ir(o)t} + \varepsilon_{ir(o)t} \end{aligned} \tag{13}$$

where $1(Covered)_{ir(o)t}$ is a dummy equal to one if individual i location was covered by the mobile phone network in the year of the Afrobarometer survey. Notice that in this specification we can also control for city of origin times year fixed effects ($\alpha_{r(o)t}$) which effectively allows us to compare individuals within the same community of origin but with different exposure to cultural values at destination due to differences in mobile phone coverage at origin. Although mobile phone based money transfers were scarcely used during the period under study in equation (13) we always control also for an interaction of exposure to economic conditions at destination with mobile phone access.

3.2 Results

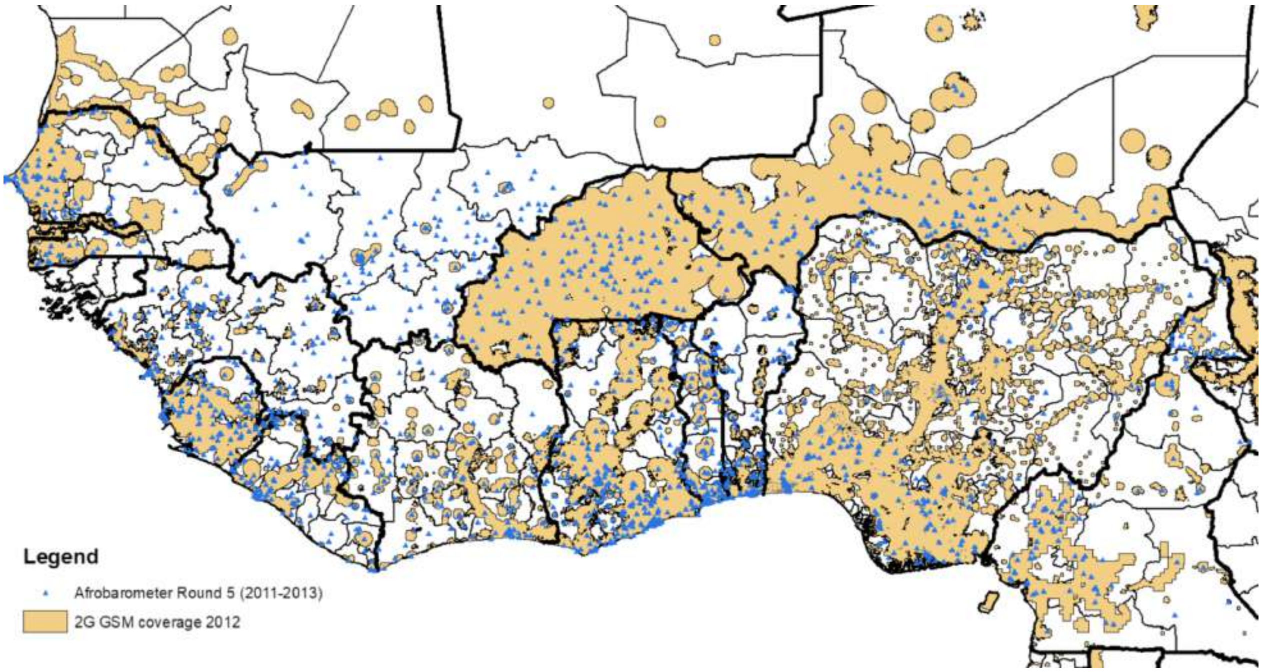
3.2.1 *Attitudes towards immigrants and their view of democracy*

Before presenting our core results studying how attitudes towards migrants in European destination countries affect the view of democracy in their communities of origin in Africa, we provide some descriptive evidence of how such attitudes might affect the view of democracy of the migrants themselves. To this end, we use data from the European Social Survey (ESS), which allows us to observe, for the same individual, both a measure of perceived discrimination and a measure of satisfaction with democratic institutions. For this exercise, which is purely descriptive, we restrict our attention to respondents of the ESS that are immigrants.

Table 3 reports a set of correlations between perceived discrimination and satisfaction with democratic institutions among immigrants. The outcome variable captures satisfaction with how democracy works in the destination country on a scale from 0 to 1 (at 0.1 intervals),

[16] 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.

Figure 7: Location of Afrobarometer respondents and mobile coverage in West Africa



Notes: The maps zooms onto Western Africa to show the location of Afrobarometer respondents in Round 5 along with the extent of 2G mobile phone coverage over the region in 2012.

while perceived discrimination is a dummy equal to 1 if a respondent perceives their group as discriminated against on the basis of one of the following characteristics: race, country of origin, ethnicity, religion or language.

As shown, we find that immigrants that perceive to be discriminated are on average less satisfied with the workings of democracy in destination countries. The estimate remains negative and precisely estimated when controlling for country of destination times year fixed effects, as well as controlling for the set of individual characteristics available in the ESS. These characteristics include: age, gender, education level, a variable capturing employment status (unemployed, employed, not active or retired) and a dummy capturing whether the respondent lives in a large urban center. The magnitude of the estimates suggests that immigrants that perceive to be discriminated against are about 15% of a standard deviation less satisfied with democratic institutions at destination. Although these estimates likely suffer from severe omitted variable bias, they are at least consistent with attitudes towards migrants in destination countries influencing migrants' view of democracy. In the next section, we turn to investigating more systematically whether changes in attitudes towards migrants at destination affect support for democracy in their communities of origin.

Table 3: Perceived discrimination and satisfaction with democracy among immigrants

	Satisfaction with democracy		
	(1)	(2)	(3)
Perceived discrimination	-0.033***	-0.033***	-0.039***
	(0.007)	(0.008)	(0.008)
Observations	23,845	23,845	23,706
R-squared	0.142	0.168	0.177
Country destination FE	✓	✓	×
Year FE	✓	✓	×
Country destination X Year FE	×	✓	✓
Individual Controls	×	×	✓

Notes: Sample restricted to respondents of the European Social Survey that were not born in the country were they are living. Standard errors clustered at country of destination reported in parenthesis. Individual controls include: age, gender, education level, employment status and living in large urban area fixed effects. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$..

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

Our analysis focuses on migration from Africa to Europe for the years 2002 to 2018, a period for which we have data from both the European Social Survey – used to capture anti-immigrant sentiment at destination – and the Afrobarometer – used to capture support for democracy in communities of origin. In Table 4 we use a measure of anti-immigrant sentiment from the ESS to capture $CulturalValues$ in destination countries. The measure is based on the ESS survey question "to what extent do you think [respondent's country] should allow people from the poorer countries outside Europe to come and live here?". ESS provides four potential answers, which capture the degree of agreement with allowing immigrants in the destination country on a scale from 1 to 4.[17] We construct our measure of anti-immigrant sentiment in the destination country by taking a weighted average across respondents at destination, using sampling probabilities provided by the ESS as weights.[18]

Table 4 reports the results of estimating equation (10) and using the measure of anti-immigrant sentiment described above as $CulturalValues$ at destination. The estimated β

[17] The 4 potential answers are: 1 "Allow many to come and live here", 2 "Allow some", 3 "Allow a few", and 4 "Allow none".

[18] We also construct an alternative measure of anti-immigrant sentiment at destination based on the question "to what extent do you think [respondent's country] should allow people of a different race or ethnic group from most [respondent's country] people to come and live here?". The question offers the same four potential answers. The two measures of anti-immigrant sentiment have a correlation coefficient of 0.97. The point estimates obtained when estimating equation 10 with this alternative measure are very similar in size and significance to the ones reported in Table 4.

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 standardized the exposure to anti-immigrant sentiment so that coefficients can be read as the effect of one standard deviation higher exposure relative to the mean. Thus, the coefficient in column (1) indicates that a one standard deviation increase in exposure to anti-immigrant sentiment at destination translates into a 0.23 percentage points 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 a large set of individual and location controls, as well as our measure of exposure to GDP growth at destination. As shown, the point estimate on exposure to anti-immigrant sentiment remains of similar magnitude (0.215) when including these controls.

Next, we estimate equation (13), which relies on variation in mobile phone access across individuals within the same city. First, in column (3), we report the results obtained with the same controls as column (2), but without including city of origin times year fixed effects. Next, in column (4), we include city of origin times year fixed effects, so that only coefficients on variables that across individuals within the same region can be estimated. We find that, for a given exposure to anti-immigrant sentiment, individuals with access to mobile phone coverage experience a 1 percent larger decline in their support for democracy relative to individuals in the same city but without access to mobile coverage, a small but statistically significant effect. Notice that we additionally control for exposure to GDP growth at destination interacted with mobile phone coverage, to account for potential differential effects of being connected via mobile phones to destination countries in different stages of their business cycle. We find no differential effect of economic conditions at destination on support for democracy at origin.

3.2.3 Effects of exposure to perceived discrimination of immigrants at destination on origin communities

In the second part of this analysis, we use survey data on perceived discrimination directly reported by respondents to the European Social Survey to measure changes in $CultureValues$ at destination. The ESS reports the immigration status of each respondent, which allow us to create different measures of perceived discrimination for migrants and for non-migrants.[19]

[19] More specifically, the ESS asks whether respondents were born in the country in which they are currently living. We categorize as immigrants all respondents born abroad. Based on this classification, out of the 343,218 respondents to the ESS between 2002 and 2019, 8.5% are immigrants.

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

	Support for democracy			
	(1)	(2)	(3)	(4)
Exposure to anti-immigrant sentiment at destination	-0.229***	-0.215***	-0.237***	
	(0.043)	(0.049)	(0.046)	
Exposure to anti-immigrant sentiment at destination × Coverage			-0.010	-0.010**
			(0.005)	(0.003)
Exposure to GDP growth at destination		0.001	0.001	
		(0.003)	(0.003)	
Exposure to GDP growth at destination × Coverage			-0.001	0.001
			(0.005)	(0.003)
Coverage			0.005	-0.006
			(0.018)	(0.008)
Observations	203,886	200,968	200,968	200,966
R-squared	0.074	0.084	0.084	0.113
Region origin FE	✓	✓	✓	×
Country origin X Year FE	✓	✓	✓	×
Region origin X Year FE	×	×	×	✓
Individual Controls	×	✓	✓	✓
PSU Controls	×	✓	✓	✓

Notes: The table reports OLS estimates of equations 10 (columns 1 and 2) and 13 (columns 3 and 4). Standard errors clustered at city-country of origin and AfroBarometer round reported in parenthesis. Individual controls include: age, gender, education level fixed effects, living in urban area fixed effect. PSU controls include: access to electricity, access to running water, local presence of a school, a police station, a hospital and a market. 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. All regressions are weighted by sampling weights from Afrobarometer. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$..

Overall, about 15% of respondents that are immigrants describe themselves as "being a member of a group that is discriminated" in their country of residence, against 5% of respondents among non-immigrants.

We estimate equation 10 using as measure of $CulturalValues$ at destination the degree of perceived discrimination against immigrants. We define this measure as the share of immigrants in a given destination country and year that report in the ESS to perceive their group as discriminated against on the basis of race, country of origin, ethnicity, religion or language.[20] It is important to notice here that the European Social Survey is not designed to be representative of the immigrant population within a given country. In this sense, the

[20] The vast majority (87%) of immigrants declare that discrimination against their group is based on these characteristics.

results presented in this section should be taken as only suggestive evidence. Still, we do find that the perceived discrimination measure is positively and significantly correlated with the measure of anti-immigrant sentiment at country level used in the analysis of Table 4.

Overall, the results presented in columns (1) to (3) of Table 5 are consistent with those presented in Table 4, although of smaller magnitude. The coefficients in columns (1) and (2) indicate that a standard deviation increase in exposure to immigrant discrimination at destination translates into about 0.022 to 0.025 percentage points decline in the probability to support democracy at origin. As in the previous analysis, we find that this effect is larger for individuals with access to the mobile phone network, as shown in columns (3) and (4). Finally, in columns (5) to (8), we replicate this analysis using as measure of $CulturalValues$ at destination the degree of perceived discrimination as reported by non-immigrants, focusing again on perceived discrimination on the basis of race, country of origin, ethnicity, religion or language. Although discrimination against immigrants is positively correlated with discrimination against non-immigrants at the country level, we expect the transmission of the latter to be weaker along migration networks. Indeed, we find no effect of exposure to discrimination perceived by non-immigrants at destination on support for democracy at origin.

Table 5: The effect of exposure to perceived discrimination against migrants at destination on support for democracy at origin

	Support for democracy							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure to migrant discrimination at destination	-0.025**	-0.022***	-0.021***					
	(0.007)	(0.003)	(0.003)					
Exposure to migrant discrimination at destination × Coverage			-0.009	-0.011***				
			(0.008)	(0.002)				
Exposure to non-migrant discrimination at destination					0.044	0.045	0.049	
					(0.028)	(0.026)	(0.025)	
Exposure to non-migrant discrimination at destination × Coverage							-0.006	-0.008**
							(0.006)	(0.002)
Exposure to GDP growth at destination		0.002	0.002			0.002	0.002	
		(0.003)	(0.003)			(0.003)	(0.003)	
Exposure to GDP growth at destination × Coverage			-0.001	0.000			0.000	0.001
			(0.005)	(0.004)			(0.005)	(0.004)
Coverage			0.006	-0.004			0.003	-0.006
			(0.019)	(0.008)			(0.019)	(0.008)
Observations	203,886	200,968	200,968	200,966	203,886	200,968	200,968	200,966
R-squared	0.074	0.084	0.084	0.113	0.074	0.084	0.084	0.113
Region origin FE	✓	✓	✓	×	✓	✓	✓	×
Country origin X Year FE	✓	✓	✓	×	✓	✓	✓	×
Region origin X Year FE	×	×	×	✓	×	×	×	✓
Individual Controls	×	✓	✓	✓	×	✓	✓	✓
PSU Controls	×	✓	✓	✓	×	✓	✓	✓

Notes: The table reports OLS estimates of equations 10 (columns 1 and 2) and 13 (columns 3 and 4). Standard errors clustered at city-country of origin and AfroBarometer round reported in parenthesis. Individual controls include: age, gender, education level fixed effects, living in urban area fixed effect. PSU controls include: access to electricity, access to running water, local presence of a school, a police station, a hospital and a market. 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. All regressions are weighted by sampling weights from Afrobarometer. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$..

4 Summary and concluding remarks

In this paper we propose a novel methodology to estimate subnational migration stocks from poor origin countries to rich destination countries using Google Trends data. We postulate that - and find strong evidence in favor of migrants disproportionately perform(ing) Google searches for their place of origin, hence searches proxying for their presence.

We use these estimates to derive a measure of exposure to migrants' discrimination at destination across 709 subnational areas in Africa, obtained 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. Our preliminary results show evidence of increased anti-immigrant feelings in Europe spilling over to origin communities in the form of reduced support for democratic institutions. These findings suggest that large immigration inflows that others have suggested having led to a right-wing, populist backslash might have had political consequences even outside in Europe.

References

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.

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.

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.

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.

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.

Kapur, D. (2014). Political effects of international migration. *Annual Review of Political Science 17*, 479–502.

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

Lewicka, M. (2011). Place attachment: How far have we come in the last 40 years? *Journal of environmental psychology 31*(3), 207–230.

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

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.

A Appendix

A Google Trends Data

In this section, we detail the process of identifying, extracting and normalizing the Google Trends (GT henceforth) search terms that we use in our analysis.

Google trend queries require users to specify the search term(s), the search location(s) and the time period (spanning from four hours to multiple years). Given a specified time period, queries can be performed for a combination of up to five different terms and locations. The output is provided by sub-periods (i.e. hours, days, weeks or months depending on the length of the period searched). Importantly, GT data refer to *relative* popularity of search terms, rather than to their *absolute* volume of searches. In particular, the data are expressed relative to the total volume of searches in the specified location in each sub-period and further standardized to the maximum value of relative searches across all specified terms, locations and sub-periods in the query. The relative importance of a search term in a given period is thus expressed as an integer value from 0 to 100. Searches below an (undisclosed) volume threshold appear as zeros.[21]

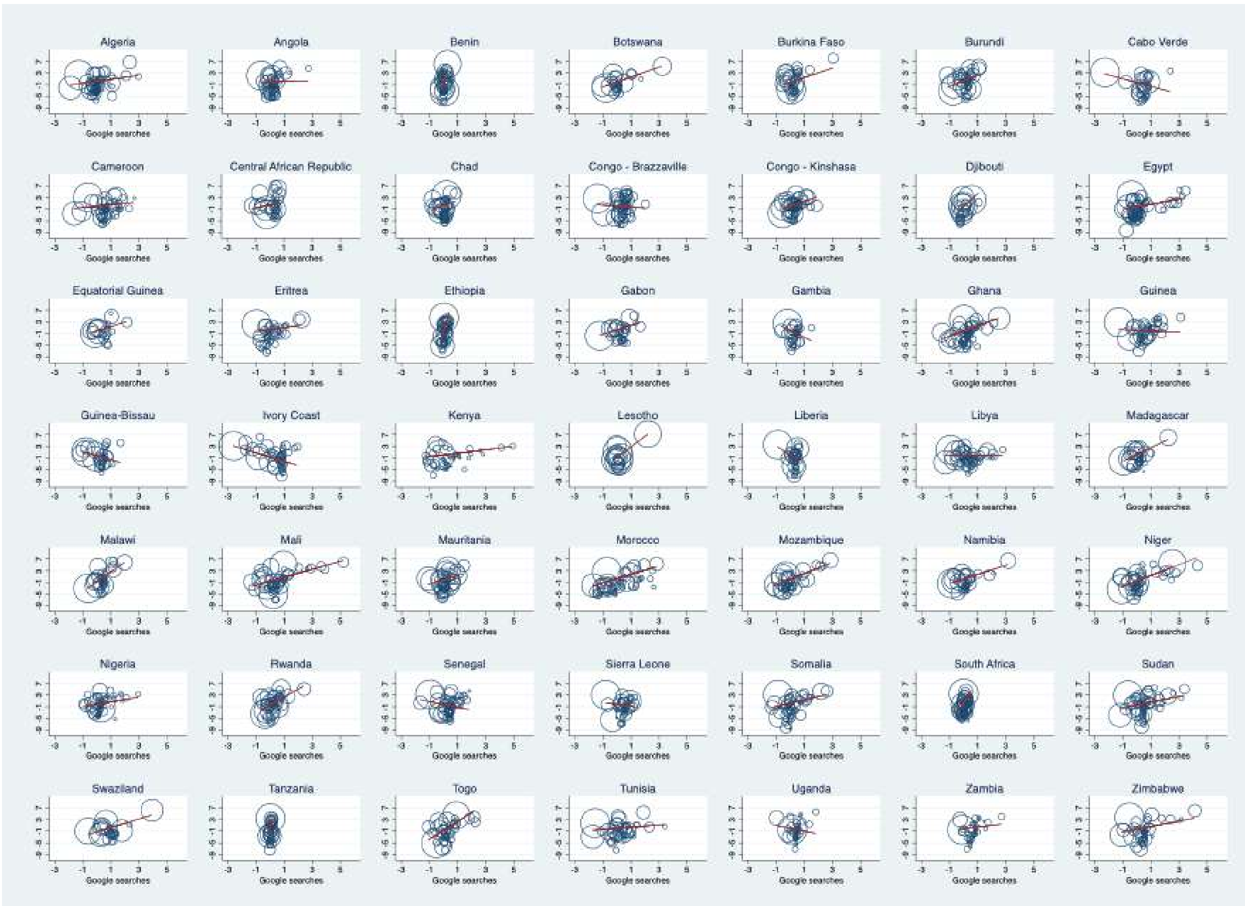
In order to compare consistently the relative popularity of search terms across time and destination countries, we normalize all searches to the same numeraire. In order to do so, we adopt the following routine, which consist of three steps.[22] First, for each of the 50 African countries of origin, we perform pairwise GT searches for the capital city against any other city of the country. We repeat the process for all 133 destination countries. This allows us to normalize GT searches across all cities of a given origin country within any given destination country relative to the country's capital city. Second, for each capital city in the 50 countries of origin, we perform pairwise extractions of its GT searches in Switzerland against its searches in any other destination country. This allows us to normalize GT searches across all cities of a given origin country across all destination countries. Finally, for each origin country capital, we compare its GT searches in Switzerland against searches for the capital of Ghana, Accra, in Switzerland. This finally allows us to normalize GT searches across all cities of origin in all destination countries in terms of a common numeraire, the number of GT searches for Accra in Switzerland. The process results in around 280k extractions. We iterate the procedure 10 times to reduce the extent of measurement error (as the relative popularity of a given search term in a single extraction is calculated on a random sample of all Google searches performed in a given location and a specified time period) and we average searches across these 10 iterations. This average is then used in the analysis.

[21] GT extractions also present computational challenges. The process of extraction is automatised using both *R* and *Python*, specifically the ad-hoc APIs *gtrendsR* and *pytrends*. These APIs facilitate but do not fully solve the extraction problem, as the number of extractions that can be downloaded is somehow limited by an undisclosed quantity. To avoid this, we use a commercial VPN. However, the process of switching IPs cannot be automatised, which implies that human interaction is needed when the server reaches the extraction threshold.

[22] We repeat each step for three, overlapping, time periods: 2004-2010, 2010-2015 and 2015-2020

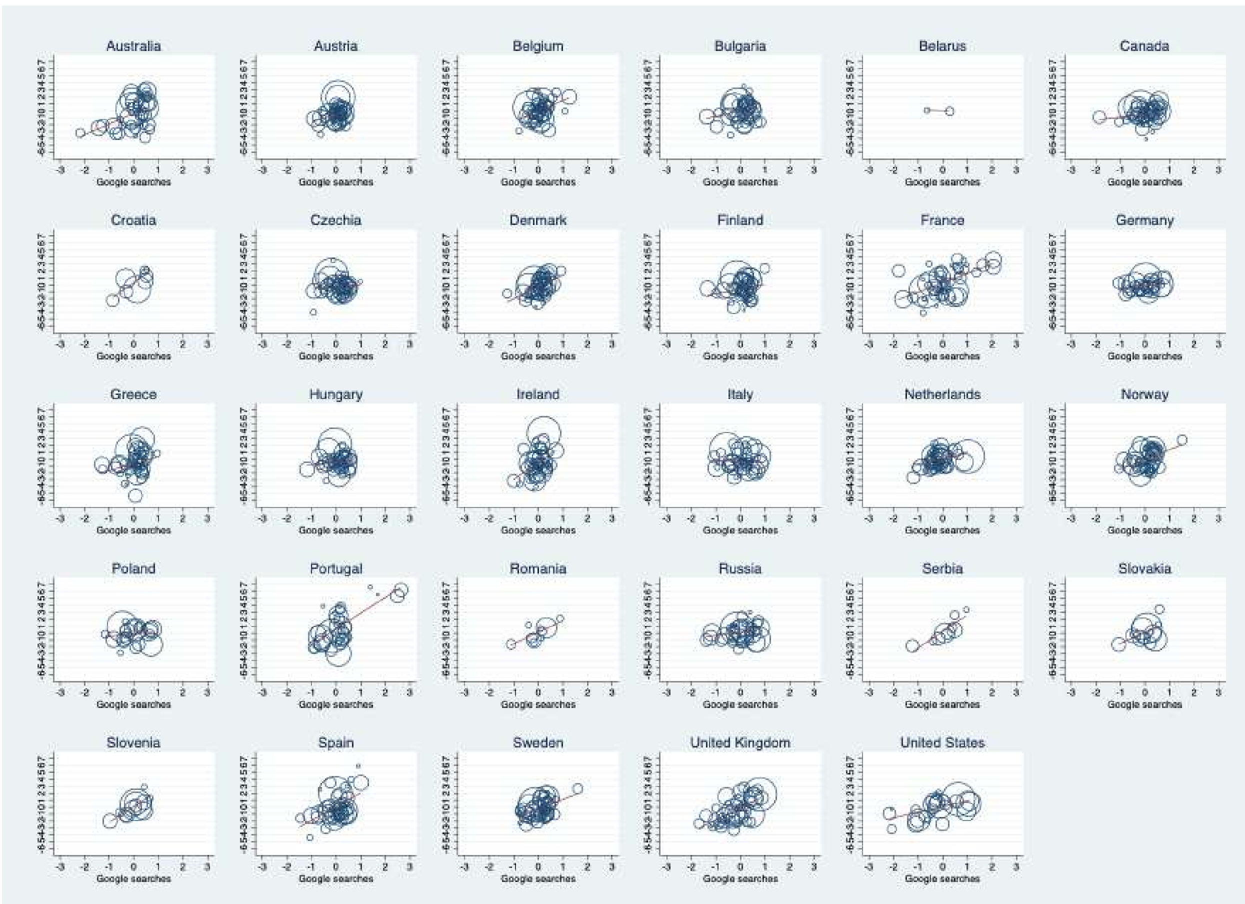
B Appendix Tables and Figures

Figure A.1: Outmigration versus Google searches across countries of destination



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

Figure A.2: Outmigration versus Google searches across country of origin - Destination countries in Europe, North America and Oceania

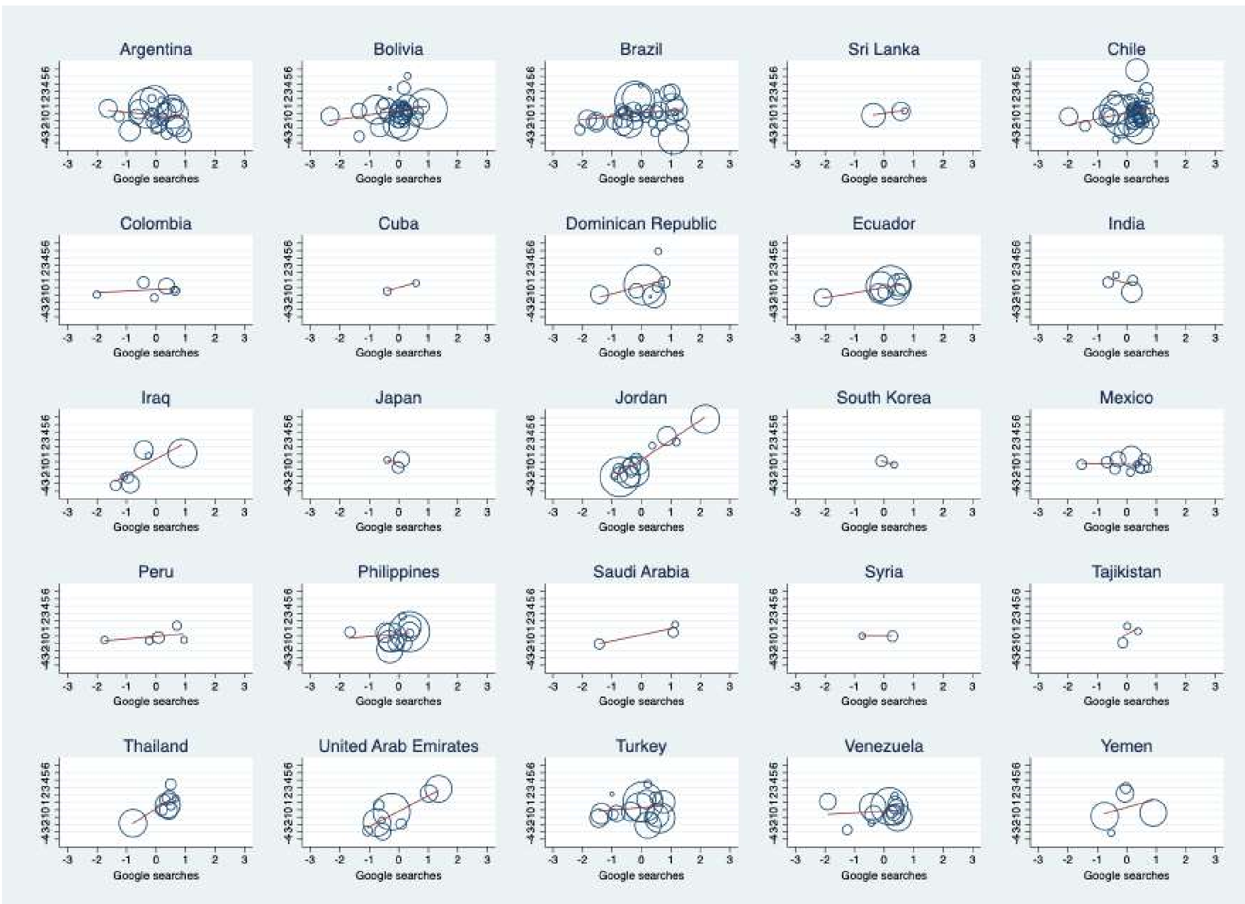


Notes: See notes to Figure 3. Larger dots correspond to larger countries of origin. Data only refer to countries of destination with population greater than one million.

Figure A.3: Outmigration versus Google searches across country of origin - Destination countries in Africa

Notes: See notes to Figure A.2.

Figure A.4: Outmigration versus Google searches across country of origin - Destination countries in Asia and Latin America



Notes: See notes to Figure A.2.

Table A.1: 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 1 separately by destination continent.

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

	(1)	(2)	(3)	(4)	(5)
GT searches	1.069***	0.990***	0.771***	0.706***	0.688***
	(0.171)	(0.175)	(0.143)	(0.145)	(0.138)
common language		0.630**	0.764***	0.616***	0.609**
		(0.274)	(0.226)	(0.203)	(0.228)
distance			-0.422***	-0.422***	-0.424***
			(0.079)	(0.078)	(0.078)
colony of destination ever				1.937***	1.406***
				(0.284)	(0.486)
colony of origin ever				0.911***	0.854*
				(0.328)	(0.477)
common colonizer					-0.537**
					(0.229)
common legal origin					0.591
					(0.423)
Observations	8,480	8,460	8,460	8,460	8,460
Country of origin FE X Year	✓	✓	✓	✓	✓
Country of dest FE X Year	✓	✓	✓	✓	✓

Notes: The table reports GLS estimates of equation (9), 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.3: 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

Table A.4: Relationship between log return migration from African Migrant Surveys and predicted migration stocks from Google trends, by African regions

	Dependent Variable: Log Migrant Surveys (WB, MAFE, EUMAGINE)						
	Any destination	Africa	Europe	Asia	Latin-America	North-America	Oceania
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Unweighted			
Log Prediction	0.069***	0.066***	0.028	0.105***	–	0.056	0.087
	(0.010)	(0.011)	(0.021)	(0.028)		(0.023)	(0.027)
				Weighted by Population at Destination			
Log Prediction	0.092***	0.088***	0.068***	0.159***	–	0.056	0.085
	(0.013)	(0.020)	(0.019)	(0.024)		(0.014)	(0.023)
				Weighted by Population at Origin			
Log Prediction	0.051***	0.047***	0.030	0.084***	–	0.021	0.083
	(0.009)	(0.008)	(0.027)	(0.026)		(0.042)	(0.030)
Observations	23,140	8,366	4,984	6,230	2,848	356	356
Region of origin FE	✓	✓	✓	✓	✓	✓	✓
Country of dest FE	✓	✓	✓	✓	✓	✓	✓

Notes: The table 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 A.5: RELATIONSHIP BETWEEN LOG RETURN MIGRATION FROM AFRICAN MIGRANT SURVEYS AND PREDICTED MIGRATION STOCKS FROM GOOGLE TRENDS, BY AFRICAN REGIONS

	Dependent Variable: Log Migrant Surveys (WB, MAFE, EUMAGINE)						
	Any destination	Africa	Europe	Asia	Latin-America	North-America	Oceania
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Unweighted			
Log Prediction	0.126***	0.131***	0.044	0.125***	–	-0.014	0.057
	(0.024)	(0.030)	(0.032)	(0.032)		(0.051)	(0.044)
				Weighted by Population at Destination			
Log Prediction	0.144***	0.149***	0.124***	0.155***	–	-0.014	0.057
	(0.027)	(0.034)	(0.039)	(0.029)		(0.030)	(0.037)
				Weighted by Population at Origin			
Log Prediction	0.123***	0.132**	0.016	0.127***	–	-0.041	0.045
	(0.036)	(0.050)	(0.014)	(0.042)		(0.091)	(0.032)
Observations	23,140	8,366	4,984	6,230	2,848	356	356
Region of origin FE	✓	✓	✓	✓	✓	✓	✓
Country of orig X Country of dest FE	✓	✓	✓	✓	✓	✓	✓

Notes: The table 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 A.6: The effect of exposure to anti-immigrant sentiment at destination on support for democracy at origin - Top Democracies

	(1)	(2)	(3)	(4)
Exposure to migrant discrimination	-0.164***	-0.157***	-0.171***	
	(0.031)	(0.035)	(0.041)	
Exposure to migrant discrimination × Coverage			-0.009	-0.006
			(0.008)	(0.007)
Exposure to GDP growth at destination		0.002	0.001	
		(0.003)	(0.003)	
Exposure to GDP growth at destination × Coverage			0.000	0.002
			(0.004)	(0.002)
Coverage			0.003	-0.011
			(0.014)	(0.006)
Observations	203,886	200,968	200,968	200,966
R-squared	0.074	0.084	0.084	0.113
Region f.e.	✓	✓	✓	×
Country X Year f.e.	✓	✓	✓	×
City X Year f.e.	×	×	×	✓
Individual Controls	×	✓	✓	✓
PSU Controls	×	✓	✓	✓

Notes: Standard errors clustered at Region of origin and Afrobarometer round reported in parenthesis. Individual controls include: age, gender, education level fixed effects, living in urban area fixed effect. PSU controls include: access to electricity, access to running water, local presence of a school, a police station, a hospital and a market. 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. All regressions are weighted by sampling weights from Afrobarometer. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.