World Development 162 (2023) 106134

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Aid and internal migration in Malawi

Mauro Lanati [a,*], Marco Sanfilippo [b], Filippo Santi [c]

[a] University of Milano-Bicocca, Department of Economics, Management and Statistics (DEMS) and European University Institute, Italy
[b] University of Turin, Department of Economics and Statistics "Cognetti de Martiis" and Collegio Carlo Alberto, Italy
[c] University of Turin, Department of Economics and Statistics "Cognetti de Martiis" and European University Institute, Italy

ARTICLE INFO

Article history:
Accepted 24 October 2022

JEL classification:
F35
R2
O55

Keywords:
Foreign Aid
Malawi
Sub-Saharan Africa
Internal Migration

ABSTRACT

This paper uses geographically disaggregated data to investigate the role of foreign aid as a *pull factor* for internal migration in Malawi over the period 1998–2008. Employing a standard gravity model of migration, we show a positive relationship between the volume of foreign assistance a district receives and the number of immigrants. While aid makes districts more attractive as migrant destinations, there is no evidence of a counterbalancing *push factor* effect on internal mobility. We also dig deeper into the mechanisms through which foreign aid can shape internal migration decisions. According to our results, the positive welfare effects of foreign assistance manifest themselves not only through a rise in economic opportunities, but also in improved access to public services in recipient districts.

1. Introduction

Differentials in economic opportunities and in the availability and quality of public services are typically among the main factors influencing the decision to migrate (Lucas, 2015). Migrants tend to move to areas where employment and income opportunities are larger (Harris & Todaro, 1970; Young, 2013), or in which the supply of public services, such as health care and education, is more abundant and of better quality (e.g. Clark et al., 2003; Gollin et al., 2021).

In developing countries, gaps in public services' provision and income opportunities across geographical areas are often related to the spatial distribution of foreign aid. Especially in poor and fragile contexts, governments have been relying on development aid to provide those social and economic infrastructures that are crucial for local population's subsistence and wellbeing. Indeed, the volume of ODA flows for several least developed countries (e.g. Burundi, Liberia, Malawi) accounts for large shares of gross national income and represents more than those countries can collect through taxes (OECD 2014). Recent studies focusing on geo-localized aid show that foreign development assistance is positively associated with healthcare quality (e.g. Kotsadam et al., 2018; Odokonyero et al., 2018), education outcomes (e.g. De & Becker, 2015; Martorano et al., 2020), the provision of public services (Pickbourn et al., 2022) and economic growth (e.g. Dreher & Lohmann, 2015; Khomba & Trew, 2022).[1]

This paper investigates the role of ODA as a pull factor for internal migration. We argue that the presence of aid projects, particularly in poor and aid-dependent countries, positively influences both monetary as well as non-monetary dimensions of wellbeing at local level. This in turn shapes the incentives to migrate internally and drives population movements. While several recent studies focused on international emigrant flows (e.g. Berthélemy et al., 2009; Lanati & Thiele, 2018; Clist & Restelli, 2021), the impact of development assistance on internal migration remains substantially unexplored. Yet, much of the population movements, especially in developing countries, occur internally rather than internationally. Globally, 1 in 7 people are internal migrants (UNDP, 2009), which is three times as many as international migrants. Furthermore, internal migration is one of the driving forces underlying the rapid demographic change occurring in most developing countries, particularly in Sub-Saharan Africa (Lagakos, 2020). The challenges imposed by rapid urbanization (Henderson

* Corresponding authors.
E-mail addresses: mauro.lanati@unimib.it (M. Lanati), marco.sanfilippo@unito.it (M. Sanfilippo), filippo.santi@eui.eu (F. Santi).

[1] These results corroborate previous empirical research at the macro level based on cross-country analysis, which found a positive impact of aid disaggregated along various lines on a range of economic and social indicators (e.g. Mishra & Newhouse, 2009; Clemens et al., 2012).

https://doi.org/10.1016/j.worlddev.2022.106134
0305-750X/

& Turner, 2020) call for a better understanding of the factors that influence migration decisions and the role of international donors in shaping the forces that drive population movements.

Our work focuses on the case of Malawi, which presents some desirable characteristics for this type of analysis. On the one hand, internal migration in Malawi is far more relevant than international migration among both rural and urban households and represents the main driver of urbanization in the country (FAO, 2018). While most of the country's population still predominantly resides in rural areas, Malawi has one of the world's highest rates of urban population growth (Anglewicz et al., 2019) and recent estimates show that internal migrants account for over half of the annual population growth in urban areas (World Bank, 2016). On the other hand, social and economic infrastructures in Malawi are highly dependent on external financial resources. Aid represents approximately 20 % of the country's GNI, and it is estimated to account for over three quarters of the country's total development expenditures (Khomba & Trew, 2022).[2] This is especially true as far as the provision of social services is concerned. For instance, recent studies showed that foreign aid accounts for 81 % of Malawi's total health expenditure (CHAI, 2015). It is also positively related to the quality of services proxied by a series of health and educational outcomes (De & Becker, 2015; Dolan, 2018). In such a context, foreign aid is likely to significantly shape the spatial differences in the provision of public services and economic opportunities.

Our empirical analysis relies on a standard gravity model of migration (e.g. Ortega & Peri, 2013) where internal bilateral migration flows are regressed on foreign aid volumes at destination. We construct a dyadic matrix over the period 1998–2008, combining information on district-to-district bilateral migration flows (source: IPUMS, 2019, Malawi population census 2008) with geo-localized data on foreign aid (source: AidData). In accordance with previous gravity model applications (e.g., Beine & Parsons, 2015; Bertoli & Fernández-Huertas Moraga, 2015), our model is estimated via the Poisson Pseudo-Maximum Likelihood (PPML) estimator. To reduce the risk of model mis-specification and the potential omitted variable bias, we include origin-time and district-pair fixed effects to account for the so-called multilateral resistance to migration (Bertoli & Fernández-Huertas Moraga, 2013). This also fully controls for origin specific and district-pair (time invariant) unobserved heterogeneity.

Our identification strategy cannot completely rule out measurement errors and potential endogeneity concerns. To address measurement related issues, we provide a series of robustness tests in which we account for different definitions of the variables of interest. This includes for instance specifying aid both in volumes and in numbers, and using different lags of both measures. We also replace data on migration flows from the census with data on the probability to migrate from several (and more recent) waves of the Demographic and Health Surveys (DHS). Endogeneity is addressed with an instrumental variable (IV) approach. The latter combines a two-step strategy along the lines of Eaton and Kortum (2002) and Head and Mayer (2014) with an IV that exploits the exogenous variation in the supply of ODA weighted by the district's probability of receiving aid (as in Nunn & Qian, 2014; Chauvet & Ehrhart, 2018; Dreher et al., 2019).

Our results reveal a positive impact of foreign aid as a *pull factor* for internal migration in Malawi. A simple back-of-the-envelope calculation based on our baseline estimates shows that moving from zero to positive aid inflows (which roughly corresponds to the 55th percentile of the aid distribution), leads to about 30 more migrants per dyad. This roughly corresponds to an additional 900 immigrants per district, which is almost 11 % of the average number of migrants per district in 2008. The use of the IV estimation leads to a significant (upward) correction of our coefficient of interest, suggesting that our baseline statistics ought to be interpreted as a lower-bound estimate of the true aid effect. Nonetheless, even our most conservative estimate of the aid elasticity leads to an impact which is not only statistically significant, but also economically relevant.

Next, we also show that the effect of foreign assistance on within-country migration (a) does not vary significantly between men and women; (b) is stronger for younger cohorts of emigrants; (c) is more likely to explain migration to urban areas; (d) is predominantly driven by economic-oriented aid projects; and (e) operates mainly as a pull, rather than as a push factor for internal migrants.

In the final part of the paper, we identify some of the potential channels at work. We assume that migration choices are mostly driven by economic opportunities and differentials in the provision of public services (Lagakos, 2020). We test for this hypothesis using additional data from different sources. The results suggest that economic development in Malawian districts, which we proxy using variation in nightlight density, is positively associated with volumes of foreign aid. Furthermore, by exploiting survey information available from Afrobarometer, we show that Malawian districts that received more development assistance are also those exhibiting improved individual access to various public services, including to education or health facilities, as well as to other types of utilities.

Our contribution to the literature is threefold. First, we complement an existing (but rather small) literature on the link between aid and internal migration, which is almost exclusively confined to the impact of cash transfer or credit access programs (e.g. Ardington et al., 2009; Bryan et al., 2014; Cai, 2020). A common pattern emerging from those studies is that access to such programs favors internal migration by relaxing household's liquidity constraints in the presence of substantial upfront migration costs. The focus on cash transfer and credit access programs is particularly suitable to investigate the budget constraint channel of aid. However, it limits the scope of the analysis to very specific types of assistance, specifically designed to relax liquidity-related constraints. By including the provision of other types of aid, our analysis may capture alternative forces driving the decision to migrate which go beyond the better capacity of would-be-emigrants to finance their moving costs. We thereby complement previous research which shows that the quality of amenities and public services is a rather important determinant of migration decisions (e.g. Dustmann & Okatenko, 2014; Gollin et al., 2021; Henderson & Turner, 2020).

Second, the paper sheds some light on the potential mechanisms linking aid to internal migration. In particular, we test whether the volume of foreign assistance in recipient districts is associated to development outcomes that are likely to shape internal migration flows. This links our paper to the growing and recent literature that uses geo-localized data to evaluate the impact of aid on both economic and social welfare indicators (e.g. Dreher & Lohmann, 2015; Kotsadam et al., 2018).

Finally, to the best of our knowledge, no other works look at the role of official development assistance as a *pull* factor for internal migration in a developing country. Instead, scholars mostly investigated the controversial role of foreign aid as a *push* factor for international migration (e.g. Dreher et al., 2019; Clist & Restelli, 2021) given the salience in the policy debate around the issue of how to deal with the rising South-North migration following the so-called refugee crises. Our findings suggest that the welfare enhancing effects of aid-supported projects not only make districts more appealing as internal migrant destinations, but also seem to

[2] This refers to the period 1998–2008. Data retrieved from the World Bank World Development Indicators.

create more incentives for households to stay in their district of origin, rather than leaving. In other words, we find no evidence of a correspondent *push factor* effect of foreign assistance which either significantly affects net internal migration patterns or positively influence international emigration decisions.

The remainder of the paper is structured as follows. Section 2 describes the method and data employed in the econometric analysis and provides some descriptive statistics. Section 3 reports and discusses the empirical findings, including a series of robustness checks as well as the replication of the whole set of results with the 2SLS estimator. Section 4 provides several extensions to the main results, while Section 5 digs deeper into the mechanisms through which foreign aid influences migration decisions. Section 6 concludes.

2. Data and empirical specification

2.1. Data on aid projects

We use data on the precise geographical location of aid-supported projects in Malawi from AidData. The dataset includes donor-reported information on ODA projects, totaling an estimated value of $5.3 billion, and covering approximately 80 % of total foreign aid reported by the government of Malawi during the period 2000–2011. For each reported aid project, AidData provides a rich set of information, including the volume of foreign assistance (committed and disbursed), sectoral and purpose codes (based on OECD's CRS classification), type of assistance, donor and year in which each project was originally agreed as well as the date of its completion and the degree of geo-localization precision.[3]

In this paper we only consider projects that were completed by the end of 2008 – the latest year for which information on internal migration is available, and whose geographic location is defined with a minimum level of precision. In line with the approach of some recent studies (e.g. Briggs, 2018), we select projects with a precision code of up to 4, which makes it easier to associate the exact location of the project within a specific district, our spatial unit of interest. While this strategy reduces the uncertainty regarding the exact geo-localization of aid-supported projects and attenuates potential issues due to measurement errors, it lowers the number of observations by about 18 %.[4]

Our baseline sample includes 65 projects, covering a total of 411 project-locations.[5] Panel (a) of Fig. 1 maps the distribution of aid projects in Malawi. The projects are spread almost equally over all districts, with a slight majority of them based in urban areas such as the capital city, Lilongwe (accounting for about 10 % of the total), and Zomba (8.8 %). Conversely, the aid volumes (Panel b of Fig. 1), demonstrate that the larger ODA flows are concentrated in the districts of Karonga, Mangochi, and in Lilongwe district. The largest share of aid disbursements in our sample takes the form of grants (around 70 %) and comes from a restricted group of multilateral agencies (African Development Bank, the European Commission, World Bank and FAO) and bilateral donors, namely the US, Norway and Germany. The top seven donors accounts for about 90 % of the total number of projects in the period considered.

As far as the sectoral composition is concerned, aid-supported projects in Malawi are highly concentrated in the agricultural sector, and are almost evenly distributed across the other groups. When looking at the size of these projects, however, aid disbursements in rural development and roads, public works and transport make about 60 % of the total volume of ODA (Fig. 2, Panel a and b).

2.2. Migration data

We employ the 2008 Population Census of Malawi to construct a retrospective panel of district-to-district migration over the period 1998–2008.[6] The Census collects demographic and social characteristics of the Malawian population, including information on individuals current and previous places of residence. We use this information to track internal mobility over time. We define as migrants all those individuals who moved to the current district of residence from any other district in a given year.[7] Starting from the year 2008 and going backward until 1998, we aggregate over origin and destination districts to build a dyadic matrix of migration flows. We compute annual district-to-district migration flows as follows:

$$Mig_{ij,t} = \sum_{p=1}^{P} I_{pi,t}[district_{t-1} = j] i \neq j$$

where I takes value 1 if the individual p currently residing in district i, moved from district j ($with\ i \neq j$) in a given year $t = 1998, \cdots, 2008$.

This approach allows us to cover internal migration flows and in a dyadic panel setup, which is particularly suitable for gravity estimations. Nonetheless, this approach and the data used introduce three potential issues of concern for our analysis. First, there might be people that migrate more than once over the course of their life. Since census data only report the current and the last declared registered residence, they do not allow us to keep track of any intermediate movement, either temporary or not. It also implies that the bilateral flows computed using this procedure include both returning migrants and first-time emigrants, without a proper distinction among the two. This first issue is likely to be more relevant for most recent years than for the less recent ones.[8] Second, the flows we reconstructed do not account for those people who might have migrated during some of the previous years, but that were not alive in 2008. As a consequence, our measure of migration flows can be biased downward but, unlike the previous case, this potential measurement issue is likely be more pronounced in the earlier years of the panel. Third, the 2008 census did not distinguish between internally displaced people (IDPs) and voluntary migration. This issue could represent a potential threat to our identification, as the routes (and the motivations) followed by IDPs might diverge from those of other migrants. The fact that our data do not exhibit any relevant/sudden surge in district-level outflows (which might have

[3] The information on foreign aid projects has been geo-localized using the UCDP/AidData methodology (see Tierney et al., 2011). Further details and information on the dataset are available at the following webpage (accessed on Dec 2019): https://www.aiddata.org/data/malawi-release-17-April-2012.

[4] Projects with a precision code higher than 4 are in almost all cases projects that have not been geolocated at all. These includes for the largest part grants directed to the central government in sectors related to Governance.

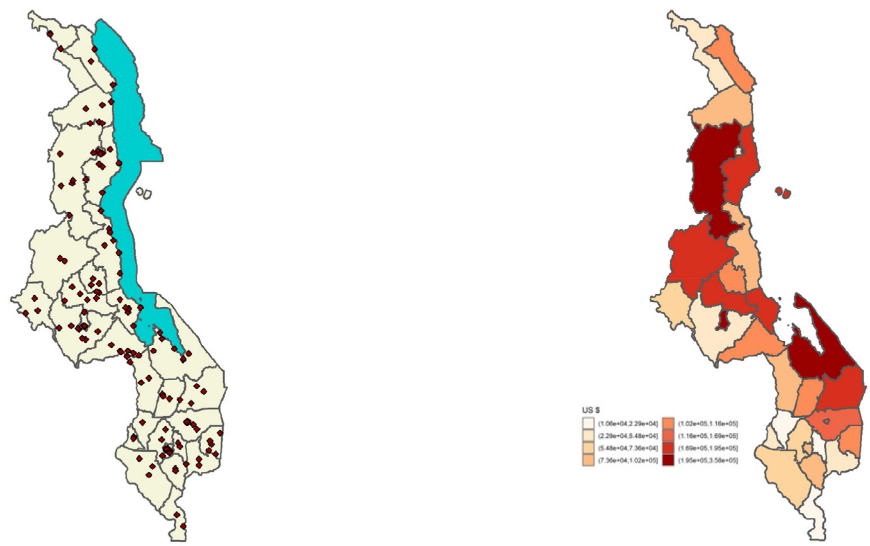
[5] As the information on financial disbursements is only available at the main project level – i.e. reported as the cumulative disbursement of all project-location entries that share the same project code – we evenly split the value of each multi-location project across the different sites involved (as it is standard practice in this literature, see for instance Dreher & Lohmann, 2015). All financial disbursements are reported in constant US $.

[6] The census tracks the current place of residence down to Traditional Authorities level, Malawi's lowest administrative division. Nonetheless, we cannot go more granular since the IPUMS anonymization procedure allows to track the previous residence of migrants only at the district level. This limitation also prevents us from considering *within* district movements (for instance, from rural to urban areas), which is itself an important component of internal migration (Becerra-Valbuena & Millock, 2021).

[7] As we identify migrants based on individual respondents, our yearly figures might be affected by some degree of recall bias. Even though this issue is more likely to affect the older periods than the more recent ones, we assume that 10 years is a short enough period to exclude substantial recall errors.

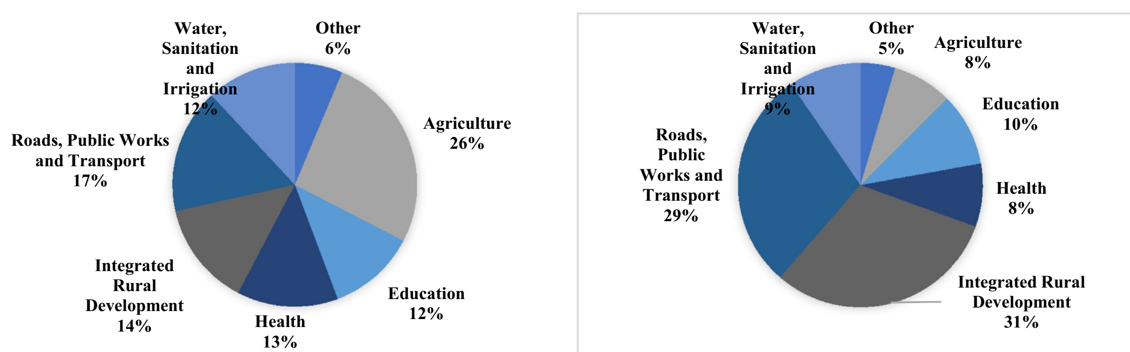
[8] An example helps clarifying this point: if an individual declares to have moved to its current location in 1998, then we can safely assume that he/she did not move a second time in the subsequent years. We cannot make the same assumption concerning an individual who declare to have moved to his/her current location in 2008: it could be that this individual had migrated at least another time between 1998 and 2008.

Panel a: Location of Aid Projects *Panel b: Disbursements for Concluded Projects*



Notes: The graph includes only completed projects concluded in the period 1998-2008.
Source: Authors' elaboration on AidData.

Fig. 1. Aid Projects in Malawi, 1998–2008.



Source: Authors' elaboration on AidData.

Fig. 2. Sectoral distribution of Aid Projects (1998–2008).

been caused by an adverse displacing event) in the period considered suggests however that this issue might not be a major concern for our analysis.

Overall, Malawi exhibits relatively high internal migration rates. About 45 % of the total Malawian population in 2008 (i.e., around 5.2 million people) declared to have changed their district of residence at least once over the course of their life. Among them, slightly less than 3 million people changed their district of residence during the decade 1998–2008.

Fig. 3 (left panel) reports the stock of migrants and shows that the southern districts are the most attractive destinations for internal migrants. Such an uneven distribution of immigration flows across geographical areas can be explained by the presence of important economic activities in the south, which attracted people from other districts. This includes, for example, large fisheries and fish farms in the southern shore of lake Malawi, as well as the massive agricultural-targeted national investment policy launched by the central government in the early 2000s, which mostly benefited the districts in the south. However, while internal migrants moved – on average – predominantly to the south-central districts and to the capital city, the areas in the North exhibited the highest growth rate of migration inflows over the period 1998–2008 (Fig. 3, right panel).

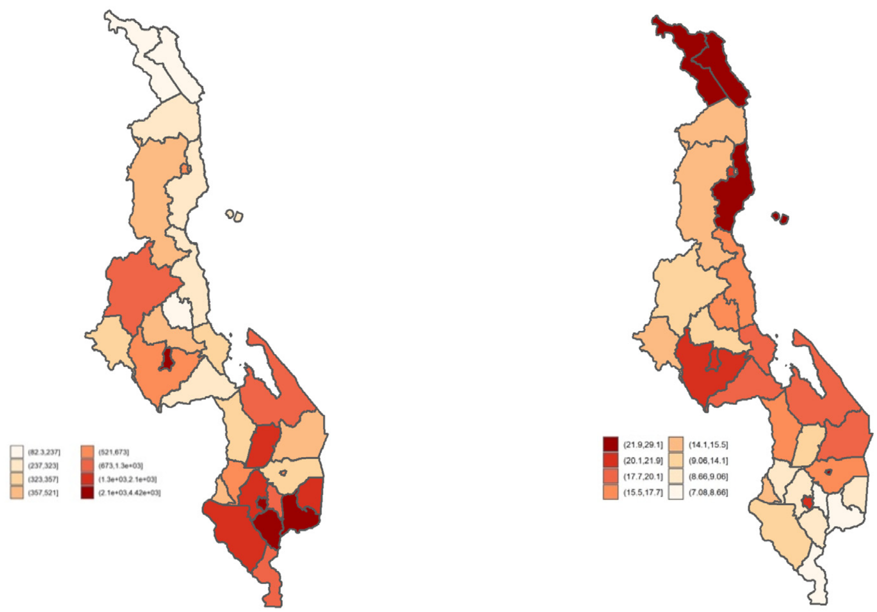
From a demographic perspective, internal migrants in Malawi tend to be slightly older than non-migrants (22.5 against 21.2 years of age respectively) and are distributed equally across genders (Table A1 in the Appendix).

2.3. Empirical specification

Our econometric specification relies on a standard gravity model of migration (e.g. Ortega & Peri, 2013) where internal bilateral migration flows from district i to district j at time t are a function of completed foreign aid projects at destination. Our baseline specification is:

$$\ln\left(Mig_{ij,t}\right) = \alpha_{ji} + \alpha_{i,t} + \beta \ln\left(AidDisbursements_{j,t-3:t-1}\right) + e_{ij,t} \qquad (1)$$

Panel a: Migrant Stock (Aggregate flow) *Panel b: Internal Immigrants (Rate of Growth)*



Notes: Shaded areas (from light to dark) denotes immigration intensity (left panel) and immigration growth (right) by district over the period 1998-2008. Source: Authors' Elaboration based on IPUMS data

Fig. 3. Migrant Stock and Migration Growth by District.

Our variable of interest $AidDisbursements_{j,t-1;t-3}$ is the 3-year average volume of aid disbursements for projects that have been concluded in district j in the three years before migration occurs. Following the existing literature, we take 3-year averages for the aid disbursements received to smooth the volatility of annual aid flows (e.g. Galiani et al., 2017; Moullan, 2013). This strategy is justified by the high volatility in the provision of foreign assistance across Malawian districts over time (see Fig. A.1 in the Appendix A). Also, we use predetermined values of aid with respect to migration inflows to alleviate potential endogeneity concerns due to reverse causality (e.g. Dreher et al., 2019; Clemens et al., 2012). While our choice on the variable of interest might bring some degree of arbitrariness, a series of robustness tests show that our results are consistent across different definitions (number of projects vs disbursements), lags and averages of foreign assistance (see Section 3.2).

The effect of ODA is first estimated without any controls, including only the set of fixed effects along the lines of Beine and Parsons (2017) and Cattaneo and Peri (2016). We deem this parsimonious model to be our preferred specification. Despite the fact that it is potentially prone to omitted variable bias, it has the advantage of not including control variables that could possibly absorb part of the overall aid effect. We subsequently add a limited number of controls to test whether our coefficient of interest is robust to their inclusion. The controls include (i) the predetermined (one-year lagged) bilateral stock of migrants from district i living in district j; (ii) the night-time light intensity, which proxies for economic activities at local level (Henderson et al., 2012); (iii) the occurrence of conflicts and the extent of climate shocks, which is measured in terms of Standardised Precipitation-Evapotranspiration Index (SPEI). Table 1 reports the descriptive statistics for the main variables included in the empirical analysis.[9]

The large set of fixed effects included in Eq. (1) significantly lowers the risk of model mis-specification and, most importantly, accounts for the so-called multilateral resistance to migration – i.e. for the fact that internal migration flows between two districts does not depend solely on their relative attractiveness, but also on the one of alternative districts (Bertoli & Fernández-Huertas Moraga, 2013). More precisely, the inclusion of origin-time dummies α_{it} controls for origin specific push factors of internal migration and leads to estimates that are consistent with the assumptions underlying the random utility model (RUM) à la Ortega and Peri (2013).[10] Furthermore, the term α_{ji} absorbs all of the (asymmetric) time-invariant dyadic determinants of internal migration, such as cultural proximity and transport costs, and generates a nest for each district-pair. This further alleviates estimation problems deriving from the potential cross-sectional dependence of the error term (Bertoli & Fernández-Huertas Moraga, 2015). Finally, all specifications are estimated with standard errors clustered at the district of destination level.

Despite the monadic nature of our variable of interest, we prefer a strategy based on a structural gravity model because our estimates of the aid coefficient β are obtained exploiting all the dyadic

[9] Among the set of controls, we did not include district's population due to data availability (at district level, it is only available for the years 1998 and 2008 – the years in which the census in Malawi was conducted). As a work-around strategy - to better account for the size at destination - we first augment the gravity model with population density at district level obtained from Harari and La Ferrara (2018), and alternatively include the (linearly) interpolated values of district's population between 1998 and 2008. As an additional robustness test, we also express the key variables in our specification (namely aid disbursements, migration and the stock of migrants) in per capita terms using the population's interpolated values. All these estimates are presented in the online Appendix (Table O4). Based on these results, we can safely conclude that district's size is not driving our results.

[10] Other than correcting for multilateral resistance, the inclusion of origin-year fixed effects on the right-hand side also captures the role of stock of residents $lnMig_{ii,t}$ as well as all migration costs specific to the district of origin (see Adovor et al., 2021 and Beine & Parsons, 2015).

Table 1
Main descriptive statistics.

Migration: Bilateral Flows and Stocks					
	N	Mean	Std. Dev.	Min	Max
Migrant Flows $_{ij,t}$	10,054	288.30	1658.03	0	65,630
Migrant Stock $_{ij,t}$	10,054	3954.55	27372.86	0	535,300
Aid (different definitions)					
	N	**Mean**	**Std. Dev.**	**Min**	**Max**
Aid Disbursements $_{j,t}$	10,054	2,026,431	4,822,938	0	32,133,124
Number of Aid Projects $_{j,t}$	10,054	1.761886	2.926194	0	17
Aid Disbursements $_{j,t-3;t-1}$	10,054	1,853,393	3,024,883	0	16,353,139
Number of Aid Projects $_{j,t-3;t-1}$	10,054	1.836301	2.534877	0	17
Control Variables					
	N	**Mean**	**Std. Dev.**	**Min**	**Max**
Nightlights $_{j,t}$	10,054	0.688	1.174	0.006	4.976
Conflict $_{j,t}$	10,054	0.327	0.381	0	1
SPEI $_{j,t}$	10,054	0.244	0.709	−1.340	1.635

Notes: **Aid Disbursements**$_{j,t-3;t-1}$ (**Number of Aid Projects**$_{j,t-3,t-1}$) refers to the 3-years average of total aid disbursements (total number of projects) concluded in the destination district *j* over the previous 3 years, expressed in constant US$. The variable **Migrant Stock** $_{ij,t}$ is constructed as the number of migrants who moved from district *i* to district *j* before year *t* (and were still resident in district *j* in 2008). See Table A2 for the complete list of sources. *Source*: Authors' Elaboration based on different datasets.

information available. In particular, estimating β with a gravity model allows us to control for all origin and bilateral specific determinants of migration, such as geographic and cultural proximity. Omitting these variables might lead to biased estimates of the aid coefficient. For instance, some districts may exhibit relatively higher immigration rates from geographically and culturally close areas of Malawi, which may wrongly be attributed to their comparably higher concentration of aid projects. In addition, the inclusion of origin-year fixed effects in a gravity model captures corrections for the so-called multilateral resistance to migration. As Bertoli and Fernandez-Huertas Moraga (2013) have shown, failing to account for multilateral resistance would lead to large distortions in the estimated coefficients. In a robustness check, we follow previous studies by Eaton and Kortum (2002) as well as Adovor et al. (2021) and estimate a monadic model with a two-step strategy (see Section 3.2.3).

In line with existing applications of the gravity model of migration (e.g. Beine & Parsons, 2015; Bertoli & Fernández-Huertas Moraga, 2015), we estimate Eq. (1) using Poisson Pseudo-Maximum Likelihood (PPML). The choice of using PPML as our preferred estimator is justified by two main considerations. First, the share of zeros in our dependent variable is approximately 22 %, which is large enough to bias the results of standard log-linear fixed effect models (see Santos-Silva and Tenreyro, 2006, 2011). Second, PPML remains consistent in presence of heteroscedasticity (see Head & Mayer, 2014; Santos-Silva & Tenreyro, 2006), and fits well with the utility-maximizing behavior of the migrants under different distributional assumptions (Schmidheiny & Brülhart, 2011).

3. Results

3.1. Baseline estimates

Table 2 reports the baseline estimates of Eq. (1). We start from a specification that only includes our variable of interest, together with the full set of fixed effects (Column 1). We then progressively add other controls, namely the stock of migrants (Column 2), *Nightlights* – as a proxy for economic attractiveness at destination (Column 3) – and the presence of *Conflicts* along with the occurrence of weather shocks (*SPEI*) at destination (Column 4).

The results suggest that foreign aid is positively associated with bilateral migration inflows. In other words, holding other factors constant, an increase in the provision of aid in a given district makes it a more attractive destination for internal migrants.

The magnitude of the aid coefficient remains remarkably stable across the different specifications (Columns 1–4). This implies that the monadic control variables in our specification do not take up part of the overall aid effect and therefore do not significantly bias the coefficient in either direction. The effect of the stock of immigrants at destination always remains statistically significant and with the expected positive sign (columns 2–4). Its elasticity is stable at around 0.3, which is in line with previous studies using gravity models applied to international migration, and confirms the role played by pre-existing migrant networks as one of the most important factors favoring migration (see Beine et al., 2016). As expected, economic activities, which are proxied by nightlight density, positively influence the attractiveness of a given district as internal migrant destination. However, there is no evidence of conflicts or climate shocks having a significant effect on within-country immigration flows.

Looking at our preferred specification, the estimated effect is not only statistically significant, but also economically relevant. Our results show that a 100 % increase in the provision of aid to a given district corresponds to a 0.8 % rise in the bilateral immigration flows within Malawi. Most of the economic relevance is, however, determined by looking at what happens when a district begins receiving foreign aid. On this respect, we propose a simple back-of-the-envelope calculation based on the following steps. First, we run our baseline specification (the one reported in column 1 of Table 2). Second, we estimate the marginal effects at different levels of the variable of interest (aid). Third, we count the difference in the predicted number of migrants from a situation of zero aid to one in which aid becomes positive. The latter corresponds more or less to a shift to the 55th percentile of the aid distribution, i.e. about 238,000 dollars. This increase will lead to 30 more migrants per dyad, which approximately corresponds to an additional 900 immigrants per district, almost 11 % of the average number of migrants per district in 2008.[11]

[11] The choice of relying on the more parsimonious specification reported in Table 2 (Column 1) as our benchmark does not affect our main results. In an additional battery of robustness tests reported in Table O2 in the online Appendix, we replicate all the main specifications discussed throughout the paper including the extended set of controls used in Column 4 of Table 2. The results are qualitatively comparable to our benchmark, leaving our conclusions substantially unaffected.

Table 2
Baseline estimates.

	(1)	(2)	(3)	(4)
Estimator	PPML	PPML	PPML	PPML
Dep. Variable	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*
Aid Disbursement$_{j,t-3;t-1}$	**0.008*****	**0.007*****	**0.006*****	**0.006*****
	(0.002)	**(0.002)**	**(0.002)**	**(0.002)**
Migrant Stock $_{ij,t-1}$		0.360**	0.323**	0.323**
		(0.162)	(0.153)	(0.154)
Nightlights $_{j,t}$			0.074***	0.075***
			(0.021)	(0.021)
Conflict $_{j,t}$				−0.004
				(0.018)
SPEI $_{j,t}$				0.014
				(0.039)
Observations	10,054	10,054	10,054	10,054
% Null	0.22	0.22	0.22	0.22
Adj. R2	0.96	0.96	0.96	0.96
Pair FE	Yes	Yes	Yes	Yes
Origin*Year FE	Yes	Yes	Yes	Yes

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ Standard errors clustered by destination in parentheses. The Table reports the results of Eq. (1) estimated with PPML with different sets of controls. The variable **Aid Disbursements$_{j,t-3;t-1}$** refers to the 3-year average of (log) total aid disbursements received by the destination district over the previous 3 years (expressed in constant US$). The additional controls include the stock of migrants from district *i* to district *j* in the previous year (in logs) as a measure of migrants' network; and three measures capturing destination specific time varying factors, average nightlight intensity, presence of Conflict, and a measure of adverse climatic conditions (SPEI). See Table A2 for a full description of the variables.

3.2. Robustness

3.2.1. Measurement issues

Foreign Aid. As discussed in Section 2.1, our estimates might be potentially biased by the lack of precise information about the financial disbursements at the location level as well as by other limitations regarding the quality of the geo-localized aid data. Another potential source of bias concerns the way we defined our variable of interest, which we construct taking the first lag of the 3-year average of aid disbursements. While this decision follows the existing literature (e.g. Galiani et al., 2017) and is essentially motivated by the high volatility of aid flows (see Fig. A1, Appendix A), it is important to rule out the possibility that this choice is not driving our results. We therefore propose a set of robustness tests in which we include several alternative definitions of our variable of interest, namely: (i) the value of projects using the non-lagged three-year average of aid disbursements (i.e. averaged across the periods t-2 to t), and the 1-year lagged value of disbursements; (ii) the number of projects, using the contemporary or lagged 3-year average as well as the 1-year lagged number.

The estimates of this exercise – reported in Table B1 in the Appendix – indicate that the positive relationship between foreign aid and internal migrant inflows holds regardless of the definition of our variable of interest. The magnitude of the *pull factor* effect rises – on average – when using the number of aid-supported projects in a given district as compared to their aggregate value in constant US$. Hence, migration decisions appear to be mostly influenced by the presence of aid projects (i.e. the extensive margin of aid effectiveness), rather than their size. In Table B1 we also test whether: (a) the effect of foreign assistance on migration decisions depends on the cumulative effects of aid projects; (b) our results hold when using commitments instead of disbursements; and (c) the results are confirmed if projects were agreed but not yet completed by the end of our sample period. While the insignificant coefficient of the not yet completed projects rules out the possibility of an "anticipation effect" of aid on internal migration, accounting for past projects in the form of the cumulative disbursement on concluded projects between 1998 and year t (Column 6, 8 – Table B1) raises the size of the aid coefficient.

Migration Flows. As reported in Section 2.2, the construction of the retrospective panel of district-to-district migration from the 2008 Population Census of Malawi could be subject to different forms of measurement error, leading to biases in the estimates of our parameter of interest.[12] To address these issues, we use information on internal migration from the Demographic and Health Survey (DHS) (rounds IV, V, and VII). The DHS provides individual level data on several socio-economic variables at both the household and the individual level, in addition to information on whether the respondent has migrated to the current location. An exploratory analysis shows that the information on migration flows from the DHS and the Population Census are highly comparable: the share of individuals interviewed in any of the DHS surveys that declared to have migrated in some point of their life is 47.3 %, which is very close to the value we obtain from the population census (45 %). Nonetheless, it must be noted that the information we gather from the DHS differs from the Population Census in (at least) three main respects. First, the place of residence of an individual in the DHS is not defined on official administrative boundaries, but at the level of survey clusters. Since each district can include several clusters (see Fig. A2 in the Appendix), DHS data allow us to match aid and migrant destinations with a higher level of precision on the basis of their geographic coordinates. Thus, aid projects could in principle be simultaneously attributed to different clusters based on their proximity to the exact location according to a given spatial buffer.[13] Second, the DHS only covers a specific segment of the population, i.e. those in the 15–49 range. Finally, respondents are only asked whether they have migrated to the location where they are inter-

[12] For a discussion on measurement concerns related to different types of data on internal migration, definitions of migration rates used in the literature, survey design and selection concerns, also see Kirchberger (2021).

[13] Each district includes on average 17, 27, and 28 clusters in 2004, 2010, and 2015 respectively. In the same order, each cluster includes on average 29, 36, and 37 persons aged 15–49. Note that the reported GPS location of each DHS cluster is randomly displaced to preserve the anonymity of all respondents (within 0–2 km for urban DHS clusters and within 0–5 km for rural clusters, with one percent of clusters randomly selected to be displaced by up to 10km). While we cannot control directly for this, our strategy to work with larger buffer zones (as detailed in the rest of the section) is in line with standard practice in the literature and should partially attenuate this problem (see Gollin et al., 2021).

viewed and when: no information is provided on where they come from. Hence, the DHS overlooks the dyadic dimension of internal migration and therefore prevents us from relying on our preferred gravity approach.

Based on this information, we run the following specification:

$$Mig_{pd,t} = \beta_0 + \beta_1 \ln(AidDisb.)_{d,t-3;t-1} + \beta_2 \ln Age_{p,t} + \beta_3 \ln Edu_{p,t} + \beta_4 Rural_{p,t} + \vartheta_j + \vartheta_t + \varepsilon_{pd,t} \quad (2)$$

where the dependent variable $Mig_{pd,t}$ measures whether the individual p living in cluster d at time t declared to have migrated the year before. $\ln Aid_{d,t-3;t-1}$, our variable of interest, measures the 3-year average aid disbursements received in the *proximity* of cluster d in the year before migration occurred.

Proximity of aid to a cluster is determined by computing the distance of each geo-localized aid project from the centroid of the cluster. In what follows, we estimate Eq. (2) by comparing different buffers, ranging from 10 to 60 km. All regressions also control for district ϑ_j and year ϑ_t fixed effects as well as for individual characteristics including age, years of education (in logs) and place of residence (urban/rural). We estimate Eq. (2) with a linear probability model (LPM), with standard errors clustered at the district level.

Results are reported in Table 3. The evidence of a positive and significant relationship between the location of aid-supported projects and the probability of migrating into a given area (i.e., evidence of aid acting as a *pull factor* for internal migration) is not sensitive to either the level of geographic granularity, nor to the different definition of migration being used.[14]

3.2.2. Alternative specifications

To test the robustness of our results, we check whether they hold against alternative estimators, different econometric specifications, and to different cuts to the sample. More specifically, we first replicate the baseline estimates with EK Tobit and a standard log linear OLS model and then re-estimate the gravity specification using different combinations of fixed effects. Further, we test whether our results survive when removing (a) the top migrant destinations; (b) the most frequent bilateral migration routes; (c) the top aid recipients and (d) all observations with zero aid inflows from the sample. Table B2 in the Appendix provides a summary of the main tests performed. As an additional check, we replicate all the results presented in this paper by restricting our sample to the period 2003–2008, which excludes the first years of our sample characterized by a relatively higher share of null migration flows and in which no aid projects among those considered were concluded.[15] Overall, this set of robustness tests indicates that our parameter of interest is rather stable across model specifications, estimators and sample selections, and remains very close in magnitude to our baseline estimate.[16]

3.2.3. Endogeneity concerns

An important econometric issue in our specification is the potential endogeneity of geo-localized aid projects, which may stem from two different sources: reverse causality and omitted variable bias. Reverse causality could be a concern if, for instance, internal migration shocks triggered by extreme events – such as natural disasters and conflicts – lead to humanitarian response by donors. To the best of our knowledge, no such disruptive event occurred in Malawi during the period analyzed in this paper, and according to the data, no sudden changes in the provision of humanitarian aid have occurred in the years 1998–2008. Omitted variables are plausibly the most relevant source of bias in the context of our analysis. For instance, the potential omission of unobserved factors, such as changes in the political landscape and/or in socio-economic conditions, might co-determine aid and migration. This is particularly compelling in our analysis as we are only able to include a limited set of district specific controls given constraints in terms of data availability in Malawi.

Potential endogeneity concerns are traditionally addressed by means of an instrumental variable (IV) approach. However, the presence of a monadic endogenous variable in a dyadic setting as in Eq. (2) makes the IV approach hardly viable in practice, as the instrument should have an *ij,t* dimension to qualify.

3.3. Two-step strategy

A possible solution is to implement an instrumental variable (IV) approach using a two-step strategy along the lines of Eaton and Kortum (2002) and Head and Mayer (2014).[17] This two-step approach departs from the standard dyadic nature of gravity models as the coefficient of the aid variable in the second step indicates how foreign aid affects the average volume of immigrants a specific district receives (relative to all other districts) in a given year. Therefore, with no dyadic terms in the second step, we can instrument our variable of interest using an IV that only varies by district and time.

In the two-step version of Eq. (1), we first estimate a gravity model with a full set of fixed effects – namely origin-year ($\alpha_{i,t}$), destination-year ($\alpha_{j,t}$) and district-pair (α_{ji}) fixed effects – using the baseline sample. The *first-step* model specification can be written as:

$$\ln(Mig_{ij,t}) = \ln\left(MigrantStock_{ij,t-1}\right) + \alpha_{ji} + \alpha_{i,t} + \alpha_{j,t} + e_{ij,t} \quad (3)$$

Second, the estimated destination-year fixed effects ($\widehat{\alpha_{j,t}}$) are then used as dependent variable in the *second-step* model and regressed on aid disbursements as follows:

$$\widehat{\alpha_{j,t}} = \beta \ln\left(AidDisbursement_{j,t-3;t-1}\right) + \alpha_j + \alpha_t + \left(\omega_{j,t} + \epsilon_{j,t}\right) \quad (4)$$

where α_j and α_t are district and time dummy variables, respectively. The composite error term in Eq. (4) encompasses the impact on internal emigration of district specific unobserved factors $\omega_{j,t}$ and the measurement error of destination-year fixed effects $\epsilon_{j,t}$. Reassuringly, when comparing the baseline with a two-step non-IV specification the aid elasticity remains substantially unchanged (column 3 of Table 4), suggesting that $\omega_{j,t}$ is a random effect and the orthogonality condition holds (see Head & Mayer, 2014 for a discussion).

3.4. Construction and validity of the instrument

Following the existing literature on aid effectiveness (Chauvet & Ehrhart, 2018; Dreher et al., 2019; Nunn & Qian, 2014), we build an instrument that exploits the exogenous variation in the supply of

[14] While we assume the error term in Eq. (2) being plausibly correlated within districts, our estimates are robust to the inclusion of robust standard errors clustered at different levels, including the *cluster d* dimension.

[15] This latter set of estimates is reported in Table O5 in the Online Appendix. Despite the large share of null flows in the pre-2003 period, the reason for which we decided to stick to the 1998–2008 time-coverage for our baseline estimates is driven by our commitment to preserve consistency across samples which compare our benchmark definition of aid (concluded aid projects) with alternative definitions employed in the different tests throughout the paper, namely incomplete/unconcluded ODA projects. AidData reports "agreed" ODA projects over a time span that goes from 1996 to 2011, which took on average around 45 months or, 3.75 years, to be completed. Therefore, only by going back to 1998 allows us to capture the potential anticipation/announcement effect on emigration decisions.

[16] The inclusion of a reduced set of fixed effects generally leads to very small changes in the size of the aid coefficient. The change in the aid elasticity becomes particularly relevant only when we exclude origin specific time dummies (Table B2, Column 3).

[17] See a recent paper by Adovor et al. (2021) for an application of this approach in a similar setting.

Table 3
Robustness tests: measurement issues/migration flows.

	(1)	(2)	(3)	(4)
Estimator	LPM	LPM	LPM	LPM
Dep. Variable	*Migrated*	*Migrated*	*Migrated*	*Migrated*
Buffer	10 km	20 km	40 km	60 km
Aid Disbursement $_{d,t-3;t-1}$	**0.0035***** (0.0005)	**0.0025***** (0.0004)	**0.002***** (0.0003)	**0.0016***** (0.0004)
Observations	75,510	75,510	75,510	75,510
Adj. R2	0.0508	0.0500	0.0495	0.0487
District f.e.	Yes	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1. Standard errors clustered at district level in parentheses. The results are obtained with a Linear Probability Model. Controls include age, years of education (both in logs) and a dummy indicating the urban/rural residence of each individual.

Table 4
IV regression.

Specification	(1)Reduced Form	(2)Baseline 1st Stage	(3) 2nd Stage no IV	(4)Baseline 2nd Stage
Estimator	**OLS**	**OLS**	**OLS**	**2SLS**
Dependent Variable	$\widehat{\alpha_{jt}}$	**Aid**	$\widehat{\alpha_{jt}}$	$\widehat{\alpha_{jt}}$
IV	**5.291***** **(1.354)**	**70.917***** **(20.675)**		
Aid Disbursement$_{j,t-3;t-1}$			**0.009***** **(0.003)**	**0.075***** **(0.023)**
Observations	341	341	341	341
Destination * Year f.e.	Yes	No	No	No
Pair f.e.	Yes	No	No	No
Destination f.e.	No	Yes	Yes	Yes
Year f.e.	No	Yes	Yes	Yes
Kleibergen–Paap F stat.	/	11.766		

Notes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.10. Robust Standard errors clustered by destination in parentheses. The reported reduced form and first stage statistics (Columns 1 and 2) refer to the IV specification of the baseline model (Column 1 in Table 2).

aid weighted by a district's probability of receiving aid. We define the probability of receiving aid from donor k as $\bar{p_{jk}} = \frac{1}{11}\sum_{t=1}^{11} p_{jk,t}$ – where $p_{jk,t}$ is a binary indicator assuming value one if district j hosts at least one agreed aid project from donor k at time t. We multiply this probability by the average net volume of aid disbursed by donor k over the previous 3-year spell to all other countries but Malawi, $ODA^{(j)}_{k,t-3;t-1}$. Finally all donor specific variables are aggregated at the district level. The resulting weighted sum is used as the IV for aid disbursement in Eq. (3). Such constructed IV is plausibly related to the volume of concluded projects in a district under the commonly adopted assumption that an exogenous shock in the total supply of aid should affect the allocation of foreign assistance proportionally. Our first stage then becomes:

$$Aid\ Disbursement_{j,t-3;t-1} = \gamma_1 \sum_k \left(ODA^{(j)}_{k,t-3;t-1} * \bar{p_{jk}} \right) + \alpha_j + \alpha_t + \epsilon_{j,t} \qquad (5)$$

The term γ_1 captures the correlation of our instrument with the endogenous variable. The first stage statistics reported in Table 4 (Column 2) indicates a positive correlation between the IV and our endogenous variable, while the Kleibergen-Paap (KP) F-statistic is above the conventional levels for most of the specifications, a fact that indicates that our estimate should not be biased by the adoption of a weak instrument. From the conceptual point of view, the exclusion restriction is expected to hold, as the total amount of aid spent by all donors outside of Malawi hardly affects within-country migration patterns.[18] However, it is not possible to formally test for the exogeneity of the instrument given that the model is exactly identified.

As work-around strategy, we *first* plot the trends displayed by aid flows and migration over time. One issue in such empirical design is in fact that the time invariant component of the IV (the probability of receiving aid) can be endogenous (Borusyak et al., 2021). As discussed in other works using a similar strategy (e.g. Dreher et al., 2019, 2021), controlling for district fixed effects reduces the identification strategy to a difference-in-differences framework. Hence, we need to prove that the exogenous variation in the provision of aid by all donors does not differentially affect migration in districts with a higher or lower probability to receive foreign projects. In Fig. B1 in the Appendix we plot the variables of interest for districts below and above the average probability of receiving aid. Panels (a) and (b) show that trends in aid and migration are parallel across the two groups, supporting our assumption on the excludability of the instrument.[19]

Second, we extend the set of instruments by coupling the shift-share instrument described above with additional IVs – namely President's Co-ethnicity and Political Switching. This allows us to

[18] Theoretically, an increase in total aid favoring economic development in third countries may also indirectly affect migration in Malawi. However, this concern more likely applies to the case of international rather than internal flows; therefore, we do not expect this channel to affect the validity of our exclusion restriction.

[19] Panel (c) of Fig. B.1 in Appendix B shows the trends in the exogenous component of the IV. The trend looks similar to those showed in Panels a and b, but the presence of a common trend is not a concern given that this is captured by the inclusion of year fixed effects in our specification.

perform a standard test for overidentifying restrictions.[20] The estimates reported in the online Appendix (Table O1) not only show that the additional IVs leave the aid coefficients substantially unaffected compared to the just-identified model (Table 4), but also substantiate the assumptions regarding the exogeneity of the instruments. In conclusion, by combining the first stage statistics with the reduced form results (reported in Table 4, Columns 1–2), we can cautiously conclude that the effect of the instrument on the dependent variable runs entirely through the endogenous variable i.e. it appears that there are no direct effects of the instruments on the dependent variable.

3.5. Discussion of the 2SLS results

Table 4 replicates the baseline estimates with the IV strategy described above. All the remaining results presented in the paper have also been re-estimated with 2SLS and are reported in the online Appendix (Table O3). Once endogeneity is accounted for, the magnitude of the aid coefficient increases substantially. This denotes potential sources of bias in the data and suggests that the baseline estimates ought to be considered as a lower bound of the "*true*" effect of foreign assistance.[21] A downward bias in the OLS coefficient is to be expected, for instance, in case of aid projects mainly directed to areas with lower levels of internal migration (e.g. rural areas). Also, measurement error in the main variable of interest—which might be imprecisely defined—could bias the OLS coefficient towards zero (attenuation bias).

4. Extensions

4.1. Sectoral aid

Several authors (e.g. Clemens et al., 2012; Qian, 2015) argue that the impact of aid is difficult to interpret as it encompasses many different types of projects and each type affects a different set of outcomes. Indeed, as reported in Section 2.1, completed aid projects in Malawi span over different sectors. This includes some of the 'early-impact' type (Clemens et al., 2012), which can foster internal migration under the promises of short-term economic opportunities, as well as other projects, whose attractiveness grounds on opportunities of access to public services that are not available in the place of origin. To investigate the heterogeneous impact of aid on migration, we group the projects according to AidData sectoral codes (see Table O6 for all the details), mostly focusing on the distinction between projects in social infrastructure/services and economic infrastructure/services.

In Table 5 we first replicate our baseline results by separately estimating the impact of aid in economic and social infrastructures – along with projects that do not fall in any of these two categories (Column 1–3). We then include all sectoral categories in the same regression (Column 4). The results show that aid projects in social and economic infrastructures are those which make districts more attractive for internal migrants in Malawi. We consider these results to be plausible, since aid-supported projects in the social and economic sectors can positively affect the quality of public services and create income and employment opportunities for the local population, respectively.

4.2. Migrants' characteristics

The results discussed so far point towards a positive role played by foreign aid on migration inflows. However, this relationship might be different when looking at rural versus urban areas, as well as across different migrant characteristics such as gender and age. Given the rapid urbanization that Malawi has been experiencing over the last 2 decades and consistently with the results reported in Table 5 (which emphasize the role played by "economic" aid as a determinant of internal migration – that is plausibly linked to the larger employment creation effect), we expect the overall aid effect to be driven by urban destinations. Additionally, Anglewicz et al. (2019) shows that there is no evidence of gender polarization in Malawi's internal mobility, with men primarily moving for work and women for marriage related reasons. Last, as showed in Table A3 in the Appendix, in Malawi the younger cohorts of the population and people in the working age are those more likely to move internally. Hence, in light of these considerations, we expect the presence of aid projects to be particularly effective in attracting men, younger cohorts and working age population from other districts.

By exploiting Census information on the age, gender and the urban/rural status of each respondent, we test these hypotheses and report the results in Table 6. More precisely, we perform a seemingly unrelated regression exercise and test for the statistical equality of aid coefficients across the different split samples. The results are fairly consistent with our priors. The effect of aid is most likely to materialize when projects target urban areas (Columns 6–7) and is significant for both genders (Columns 1–2), although with a higher coefficient for males.[22] As for the age cohorts, while the effect of aid appears to be stronger for younger cohorts of the population (including those in the working age) (Columns 3–5), we cannot reject the null hypothesis on the effect being constant across age groups (p-value = 0.16).

4.3. Accounting for the push factor effect

In this last section, we test whether the influence of aid-supported projects on internal migrant inflows is counterbalanced by a parallel *push* factor effect. We also explore the relationship between foreign assistance and *international emigration* from Malawian districts i.e. whether aid provision plays a role in influencing the local population's decision to move abroad. Theoretically speaking, the same welfare-enhancing opportunities created by development assistance – which make a district of destination relatively more attractive with respect to alternative locations – could in principle favor emigration by enabling a larger share of the population in the districts of origin to finance migration costs (*Budgetary Constraint Channel*). Under different assumptions, such opportunities could also lead to more incentives to reduce emigration through increasing opportunity costs and diminishing the net benefits of migration (*Income Channel*). Recent empirical research focusing on international emigration collected some evidence in favor of the latter hypothesis, although the findings point towards a quantitatively small impact of foreign aid (see for instance Lanati & Thiele, 2020; Clist & Restelli, 2021). In order to account for the role of aid-supported projects as a *push factor* for internal mobility we conduct a series of econometric exercises whose results are reported in Table 7. First, we follow Ariu et al.

[20] The two measures capture the share of population in a district belonging to the same ethnic group of the president (*President's Co-ethnicity)* and the political change of a district's administrative government toward the presidential party, that can result from local election or from government change (*Political Switching).* The rationale for both instruments comes from standard political economy arguments for which presidents' birthplaces and co-ethnicity with the ruling party matters for the allocation of foreign aid. See Khomba and Trew (2022) for more discussion about the relevance and the conditions for excludability of the two instruments.

[21] Note that a systematic downward bias of the OLS coefficient is reported in studies using a shift-share IV strategy similar to the one implemented in this paper (e.g. Nunn & Qian, 2014; Chauvet & Ehrhart, 2018; Dreher et al., 2019).

[22] Notice that with mostly urbanized areas we are considering the districts that have more than average respondents declaring to leave in a non-rural area.

Table 5
Disaggregated analysis: ODA by sector.

Estimator Dep. Variable	(1) PPML *Migrant Flows*	(2) PPML *Migrant Flows*	(3) PPML *Migrant Flows*	(4) PPML *Migrant Flows*
Disbursement for Social Projects $_{j,t-3;t-1}$	**0.007**** **(0.002)**			**0.006**** **(0.002)**
Disbursement for Economic Projects $_{j,t-3;t-1}$		**0.016*** **(0.003)**		**0.016**** **(0.007)**
Disbursement for Miscellaneous Projects $_{j,t-3;t-1}$			0.005 (0.005)	−0.002 (0.003)
Observations	10,054	10,054	10,054	10,054
% Null	0.22	0.22	0.22	0.22
Adj. R2	0.96	0.96	0.96	0.96
Pair f.e.	Yes	Yes	Yes	Yes
Origin * year f.e.	Yes	Yes	Yes	Yes

Notes*** p < 0.001, ** p < 0.01, * p < 0.05 Standard errors clustered by destination in parentheses. For detailed information on the sectoral classification and division across Social, Economic, and Residual projects, see Table O6 in the online Appendix.

Table 6
Disaggregated analysis: different types of migrants.

Estimator Dep. Variable	(1) PPML *Migration (Men)*	(2) PPML *Migration (Women)*	(3) PPML *Migration (Youth)*	(4) PPML *Migration (Work.)*	(5) PPML *Migration (Old)*	(6) PPML *Migration (Urban)*	(7) PPML *Migration (Rural)*
Aid Disbursements $_{j,t-3;t-1}$	**0.010***** **(0.002)**	**0.006***** **(0.002)**	**0.008***** **(0.002)**	**0.007***** **(0.002)**	**0.007** **(0.002)**	**0.006*** **(0.003)**	0.005 (0.003)
Observations	9,966	9,790	9,647	10,032	4,750	1,320	8,734
% Null	0.22	0.2	0.37	0.27	0.74	0.09	0.24
Adj. R2	0.94	0.94	0.97	0.87	0.4	0.97	0.96
Pair f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin * year f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald Test Chi2 (p-value)	12.39 (0.00)		0.16 (0.69)			3.97 (0.04)	

Notes: *** p < 0.001, ** p < 0.01, * p < 0.05. Robust Standard errors clustered by destination in parentheses. The different number of observations across columns relates to (a) the larger number of singleton observations that are dropped by the estimation routine in the gender (columns 1 and 2) and age (columns 3 to 5) decomposition exercises; and (b) the different sample size following the distinction between rural and urban districts (columns 6 and 7). See the note to Table A2 for more detailed information concerning the distinction between urban vs rural of migration flows.

(2016) and use net flows—defined as differences between inflows and outflows of migrants in a given district of destination—as dependent variable. As some *net flows* have negative values, we use OLS rather than the PPML estimator. The estimates reported in the first column of Table 7 suggest that an increase in the size of aid at district level is associated with a positive variation of migrant *net flows* (Column 1). In other words, while aid makes potential domestic destinations more attractive, we find no evidence of a counterbalancing *push factor* effect which significantly affects internal migration patterns. Similarly, when replacing our variable of interest with the ratio of ODA per capita between destination and origin districts in our baseline specification (Column 2), we again find the attraction effect of aid projects dominating its *push factor* counterpart. These conclusions are substantiated by additional gravity estimates obtained from alternative specifications in which we drop origin-year fixed effects to isolate the impact of aid projects at the origin (Column 3 and 4). The estimates reveal that while foreign aid in the district of destination is positively associated with bilateral migration inflows, there's no evidence of a significant push factor effect of ODA projects at the origin.

Finally, to explore the relationship between foreign assistance and international emigration, we make use of the information included in the additional annex of the 2008 Malawian Census regarding *whether and when any member of the household migrated abroad in the past 10 years* (no matter whether the member was still living abroad or returned home in the meanwhile). By using information on the residence of the household head at the moment of the declared emigration, we construct the flow of international emigrants in each district at year t.[23] We aggregate international emigrants by district of origin. The results for this exercise are reported in Column 5 of Table 7. Interestingly, the aid coefficient turns negative and significant at 10 % level, which seems to suggest that providing higher amounts of aid lowers the incentives for the local population to emigrate internationally.[24] Taken together, these findings imply that – other factors held constant – larger provision of aid shapes migration patterns in Malawi by rising the stock of immigrant population at district level.[25] In substantive terms, the evidence appears to be at odds with the budgetary constraint channel

[23] Similarly to our baseline, this exercise is potentially subject two potential sources of bias. First, attributing emigrants to the district where the household head declared to be resident in the year a member emigrated might be prone to the same measurement issues reported in section 2.2. Second, the reported emigration figures by definition cannot keep track of entire households migrating to another country.

[24] We are not considering emigration rates for international mobility because we have information on population at district level only for two years (1998 and 2008). As an unreported check, we linearly interpolate district population and compute the corresponding emigration rate. The coefficients of the PPML estimation point toward an even larger negative coefficient of the Aid variable when net emigration rates are considered (the coefficient varies between −0.009 to −0.015, depending on the controls included). The results also hold when estimating a gravity model without origin-year fixed effects and including conflicts, nightlights and climate shocks as origin-specific determinants (push factors) of internal migration. The coefficients of the control variables are not statistically significant at conventional levels and their inclusion does not substantially affect our parameters of interest. All these results are available upon request.

Table 7
Accounting for the Push Factor Effect.

Estimator Dep. Variable	(1) Pooled OLSNet Emig. Rate (log)	(2) PPML *Migration*	(3) PPML *Migration*	(4) PPML *Migration Rate*	(5) Pooled OLS *International Emigration Rate (log)*
Aid Disbursement $_{j,t-3;t-1}$	**0.008**** **(0.004)**		**0.006***** **(0.001)**		
Aid Disbursement $_{I,t-3;t-1}$			**0.003** **(0.002)**	**0.003** **(0.002)**	−**0.009*** **(0.005)**
Aid $_{j,t-3;t-1}$/Aid $_{I,t-3;t-1}$		**0.081**** **(0.020)**			
Model	Net Emig. Pull model	Relative Oda pull model	Aid in both districts	Origin Side Push model	International Emigration
Observations	10,230	10,054	10,054	10,054	321
% Null	0.23	0.22	0.22	0.22	0.02
Adj. R2	0.96	0.96	0.96	0.96	0.90
Pair f.e.	Yes	Yes	Yes	Yes	No
Ori * Year f.e.	Yes	Yes	No	No	No
Dest * Year f.e	No	No	No	Yes	No
Origin f.e.	No	No	No	No	Yes
Year f.e.	No	No	Yes	No	Yes

Notes: *** p < 0.01, ** p < 0.05, * p < 0.10. Robust Standard errors in parentheses clustered by destination in Column (13) and by origin in Column (4–5), respectively. The dependent variable in Column (1) is the log of the net emigration rate. In Columns (2–4) it is replaced by bilateral internal migration rate, and by international emigration rate in a given district of origin in Column (5). Columns (3) to (5) include a set of origin related controls in place of origin*year fixed effects. For the same reason, Columns (3) and (5) also include the same set of controls at destination.

and the positive role of aid in favoring net emigration by enabling a larger share of the population to finance their moving costs.

4.4. Transmission channels

In this section, we dig deeper into some of the potential channels through which foreign aid can affect internal migration decisions in Malawi. Specifically, we empirically test two potential mechanisms: (1) the capacity of aid to create local economic opportunities; and (2) its role as a source of amenities and public services at district level.

4.5. Economic opportunities

The potential role of aid as determinant of internal migration is partly grounded in its capacity to spur economic growth. While the literature on the aid-growth nexus at macro-level is inconclusive (see Arndt et al., 2010), there seems to be some consensus on a positive relationship in the recent studies based on more refined information on aid projects at sub-national level. For instance, a recent paper by Khomba and Trew (2022) shows that economic growth in Malawian districts is positively influenced by the volume of aid inflows. They argue that aid is likely to be a major instrument in leveraging economic growth in the country, as it accounts for more than 70 % of the overall development spending. Using the same data on geo-localized aid projects employed by Khomba and Trew (2022), we regress the growth rate of nighttime light (NTL) density on the volumes of aid received at district level. NTL density proxies for the intensity of economic activities at geo-localized level and is commonly used in the literature (see Henderson et al., 2012). The regression includes district and year fixed effects, with standard errors clustered at the district level. The results reported in Table 8 (Column 1) show that aid-supported projects are positively related to economic growth at district level. This finding corroborates the results of the disaggregated sectoral analysis (Table 5), and generally supports our hypothesis on the significant role of aid on internal migration decisions through the creation of greater economic opportunities.

4.6. Public services provision

Next, we look at the capacity of aid-supported projects to provide access to specific types of public services, such as health care facilities, schools and basic infrastructures, that can plausibly affect the decision to migrate internally (see Dustmann & Okatenko, 2014; Gollin et al., 2021). To test whether aid matters for the provision of public and social services to the local population, we employ individual data from rounds 3 and 4 of the Afrobarometer Survey. The survey covers a total of 2,384 individuals for Malawi, based in 68 and 69 clusters in the years 2005 and 2008.[26]

We focus on the following facilities: *Schools*, *Health Clinics*, *Electricity*, *Piped Water* and *Sewage Systems* and use information on whether a given facility is "*…present in the primary sampling unit/enumeration area, or within easy walking distance*" (Afrobarometer's survey codebook). We employ a linear probability model in which individuals' responses (0 or 1) are regressed on aid volumes received by the district where the household resided at the time of the survey (2005 and 2008).[27] All regressions control for individual characteristics (gender, age, residence in rural/urban areas) as well as district and time fixed effects. The results reported in Table 8 (Columns 2–6) suggest that the probability for an individual to live in proximity of some key facilities is generally higher in locations which receive larger volumes of foreign aid.

5. Conclusions

The policy and academic debate around the relationship between ODA and migration has almost exclusively been centered

[26] Enumeration areas for each Afrobarometer survey location are fully consistent with the AidData procedure (BenYishay et al., 2017). Afrobarometer follows a random selection process designed to generate a representative cross section of the population of voting age in each country. The sampling is based on geographic primary sampling units that form the Enumeration Areas (EA). Such units are selected with a probability proportional to their population size. A respondent is selected within a randomly selected household for each EA. Gender balance in the sample is ensured by alternating men and women in consequent interviews.

[27] Note that despite the first and second round of the survey were also overlapping with our period of interest, we could not use them due to the absence of the relevant questions in the previous questionnaires.

Table 8
Mechanisms.

	(1)	(2)	(3)	(4)	(5)	(6)
Mechanism	*Growth*	*Public services*				
Dep. Variable	*Avg. Nightlight*	*School*	*Clinic*	*Electricity*	*Pipes*	*Sewer*
Aid Disbursements$_{j,t-3;t-1}$	**0.001****	**0.027*****	**0.016***	**0.021****	**0.039*****	**0.039*****
	(0.001)	**(0.006)**	**(0.007)**	**(0.007)**	**(0.006)**	**(0.006)**
Observations	186	2,209	2,209	2,209	2,209	2,209
R-squared	0.89	0.895	0.457	0.521	0.538	0.456
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District f.e.	Yes	Yes	Yes	Yes	Yes	Yes
Year f.e	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$. Robust Standard errors clustered by district in parentheses. The specification reported in Column (1) is estimated as a pooled OLS with district and time FE and includes a set of district level controls (Population density, SPEI, conflict). The dependent variable – the Average Nightlight – has been obtained from NOAA's satellite data, while population density comes from Harari and La Ferrara (2018) and has been computed in the same way as the SPEI and the conflict related variables (see the notes to Table A2). Columns (2) to (6) refer to individual respondents to the Afrobarometer survey (Rounds 3 and 4) and are estimated using a Linear Probability Model. The dependent variable in each column takes the value of 1 if the related infrastructure (School, Health Clinic, access to Electricity Grid, Piped Water and Sewage System) is located within easy walking distance or if the respondent has easy access to them. All equations from (2) to (6) include district and time fixed effects, in addition to a set of individual characteristics (gender, age, rural/urban location), and are weighted using the sample weights (also provided by Afrobarometer).

around the potential role of foreign assistance as an instrument to manage (and curb) international emigration from developing countries. Yet, from a developing country perspective, international migration accounts for a relatively small share of total population movements, as it is characterized by substantive upfront moving costs. Especially in poor and deprived contexts, internal emigration decisions, namely *whether* and *where to* emigrate – are likely to be more sensitive to the welfare enhancing effects of foreign assistance.

In this paper, we have showed that ODA acts as a *pull factor* for internal migration in Malawi. This positive relationship between ODA and internal migrant inflows holds across a series of robustness tests – which include an instrumental variable (IV) approach. Even our most conservative estimate of the aid elasticity leads to an impact which is not only statistically significant, but also economically relevant. Taking our point estimates at face value, moving from zero to positive aid inflows leads to 900 more migrants per district, which is about 11 % of the average number of migrants per district in 2008. Conversely, we find no evidence of a counterbalancing *push factor* effect of aid-supported projects which either rises the number of international emigrants from Malawian districts or significantly affects internal migration patterns.

When investigating the potential channels at work, our analysis reveals that the positive welfare effects of foreign assistance manifest themselves not only through an increase in economic opportunities, but also via improved access to local public services in recipient districts. This result corroborates with previous research on the importance of aid-supported projects in affecting non-monetary dimensions of well-being, particularly in low-income countries.

From a policy point of view, this paper highlights a so far unexplored dimension of foreign aid i.e. its capacity to drive within-country migration by affecting the distribution of economic and income opportunities across internal areas as well as the spatial disparities in the access of public services. In a context where most of ODA funds within African countries are delivered to richer urban districts (Briggs, 2018), our findings pose additional challenges in donors' aid allocation decisions. For instance, while aid targeted to urban areas might help cities to better manage the process of rapid urbanization (Henderson & Turner, 2020), it can also favor population growth in already congested cities and magnify the existing rural/urban gaps in living standards. Hence, donors should plan the spatial distribution of aid funds also in view of the potential role of aid in shaping internal population movements. Investigating whether aid has similar pull-factor effects on internal migration in developing countries other than Malawi – particularly in contexts characterized by different levels of aid dependency – is a promising area of further research.

CRediT authorship contribution statement

Mauro Lanati: Conceptualization, Methodology, Writing. **Marco Sanfilippo:** Conceptualization, Methodology, Writing. **Filippo Santi:** Conceptualization, Data analysis, Formal analysis, Methodology, Writing.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to two anonymous reviewers, Paul Clist, Robert Lucas, Rainer Thiele, Martin Ruhs, Léa Marchal, Lennart Kaplan, Frank Borge Wietzke, Sarah Langlotz for helpful comments and suggestions. We also thank participants at the HiCN Annual Workshop on "*Conflict, Migration, and Displacement*" at the University of Goettingen, the webinar *'Development Aid & Migration: Unexplored Dimensions and Latest Research Findings'* organized by the Migration Policy Centre (EUI), the Economics Department lunch seminar series at the University of Turin, and the Junior Seminar on the Economics of Migration. Daniel Chris Khomba and Alex Trew kindly shared their data on co-ethnicity and political switching. Mauro Lanati received funding from Stiftung Mercator - Grant/Award Number: PN 14–297

Appendix A. Data and description of the variables

Table A1
Migration inflows and stocks by gender & district of destination.

	1998				2008			
	Women		Men		Women		Men	
District	Flow	Stock	Flow	Stock	Flow	Stock	Flow	Stock
Blantyre City	10,210	102,060	11,830	108,100	52,080	309,540	54,690	320,140
Lilongwe City	5490	32,680	6310	38,360	25,420	134,160	26,560	147,750
Thyolo	8680	182,460	8950	155,190	18,870	305,000	18,850	278,780
Mulanje	7600	170,290	7710	141,770	15,380	270,910	14,970	242,820
Chikwawa	7600	118,380	7430	118,650	15,070	215,910	15,180	215,100
Blantyre	5490	101,180	5340	91,900	11,100	170,670	10,740	160,360
Balaka	5460	88,710	4970	75,550	10,820	159,230	10,630	146,220
Phalombe	4910	94,030	4790	79,920	10,090	162,810	9720	147,290
Chiradzulu	4430	93,170	4700	76,320	8660	150,620	7950	134,260
Mangochi	2290	22,240	2190	20,590	8070	56,570	8150	55,250
Kasungu	2510	25,150	2970	27,690	7300	65,260	7540	69,580
Zomba City	1300	10,000	1490	10,690	7260	39,880	7510	41,560
Nsanje	3670	65,840	4240	64,810	7400	117,710	7200	115,220
Mzuzu	1710	8220	1580	8470	7300	38,310	6800	38,030
Lilongwe	1020	12,260	1070	12,900	6050	39,380	6630	40,350
Mzimba	1400	13,100	1330	13,610	4330	37,690	4680	37,890
Mwanza	1200	24,100	1450	21,600	4340	46,090	4440	43,690
Machinga	890	9890	1030	9220	3980	27,980	4160	27,560
Neno	1910	29,860	1800	26,870	3560	54,330	3530	51,370
Ntcheu	840	11,150	1040	10,840	3350	27,570	3540	27,710
Salima	820	9110	840	9920	3210	24,110	3530	26,330
Zomba	1130	10,830	1070	11,590	3470	27,350	3120	28,550
Mchinji	1080	10,940	1160	12,510	3270	27,150	2980	29,440
Dowa	1150	11,780	880	11,540	3000	29,180	3000	27,830
NkhataBay-Likoma	620	4530	650	5220	2610	16,060	3250	17,950
Dedza	660	6420	730	6470	2560	19,450	2790	20,690
NkhotaKota	820	10,110	1010	12,230	2360	23,560	2470	26,860
Karonga	380	4380	350	3900	2280	14,710	2180	13,440
Rumphi	720	5400	730	5480	2150	16,650	2190	16,910
Ntchisi	610	5980	550	6090	1730	14,470	1630	14,520
Chitipa	230	1910	290	1780	830	5780	970	5620

Source: Authors' elaboration based on Malawi 1998 and 2008 Censuses (IPUMS, 2019).

Table A2
Description of the main variables.

Domain and Source	Variable Name	Description
Official Development Assistance Data **Sources:** Malawi Aid Management Platform Geocoded Research Release, 2000–2011 OECD-DAC International Development Statistics (Used to build the IV)	**Aid Disbursements**$_{j,t-3;t-1}$	*Log of the 3-year average of total aid disbursements received by the destination district **j** and concluded over the previous 3 years (expressed in constant US$).*
	Number of Aid Projects$_{j,t-3;t-1}$	*Log of the 3-year average of the number of aid projects concluded in the destination district **j** over the previous 3 years.*
	Stock of Aid Disbursements $_{j,t}$	*Log of Cumulated Disbursement of Aid Projects concluded up to time **t** in district **d** (expressed in constant US$)*
	Disbursement for Incomplete Projects$_{j,t-3;t-1}$	*Log of the 3-year average the resources allocated to aid projects in **j**, launched in year/t^{-3}(−|−) to **t-1**, but completed after 2008 (expressed in constant US$).*
	Disbursement for Social Projects$_{j,t-3;t-1}$	*Log of the 3-year average of total aid disbursements dedicated to Social-related CRS Sectors, received by the destination district **j** and concluded over the previous 3 years (expressed in constant US$).*
	Disbursement for Economic Projects$_{j,t-3;t-1}$	*Log of the 3-year average of total aid disbursements dedicated to Economic-related CRS Sectors, received by the destination district **j** and concluded over the previous 3 years (expressed in constant US$).*
	Disbursement for Miscellaneous Projects$_{j,t-3;t-1}$	*Log of the 3-year average of total aid disbursements dedicated to non-Social, non-Economic related CRS Sectors (or whose destination is unclear), received by the destination district **j** and concluded over the previous 3 years (expressed in constant US$).*
	Instrument $_{j,t-1}$	*3-year average of the net spending by each international donor operating in district **j** everywhere but in Malawi in the previous 3 years, weighted by the probability of each donor to be involved in district **j** over the period 1998–2008.*
Internal Migration Data **Source:** IPUMS, 2019. https://doi.org/10.18128/D020.V7.2	***Migrant Stock***$_{ij,t-1}$	*Stock of Migrants born in district **i** and living in district **j** as in year/t (in logs).*
	Migrant Flow$_{ij,t}$	*Total number of people that moved from district **i** to district **j** at time **t** (Dependent Variable)**
	Migrant Flow (Men)$_{d,t}$	*Total number of Men that moved from district **i** to district **j** at time **t** (Dependent Variable)*
	Migrant Flow (Women) $_{j,t}$	*Total number of Women that moved from district **i** to district **j** at time **t** (Dependent Variable)*
	Migrant Flow	*Total number of 0–14 y.o. children that moved from district **i** to district **j** at*

Table A2 (*continued*)

Domain and Source	Variable Name	Description
	(Youth) $_{j,t}$	*time **t** (Dependent Variable)*
	Migrant Flow (Work) $_{j,t}$	*Total number of 15–64 y.o. working age migrants that moved from district **i** to district **j** at time **t** (Dependent Variable)*
	Migrant Flow (Old) $_{j,t}$	*Total number of 65 + y.o. elders that moved from district **i** to district **j** at time **t** (Dependent Variable)*
	Migrant Flow (Rural vs Urban) $_{j,t}$	*Total number of people that moved from district **i** to district **j** at time **t** (Dependent Variable). Districts are classified as "predominantly urban" and "predominantly rural"***
Additional controls:	***Nightlights*** $_{j,t}$	*Average Night-stime light Luminescence in district **j***
Sources:NOAA-DMSPHarari and La Ferrara (RESTAT 2018)	***Conflict*** $_{j,t}$	*Presence of any form of conflict in district **j** (dummy)*
	SPEI $_{j,t}$	*Crop affecting environmental variable in destination district **j***

Notes: Subscripts – o indicates the district of origin; d refers to the district of destination (when referring to internal migration); t refers to time. All variables taken from Harari and La Ferrara (2018) were originally available at cell level, and have been processed and rescaled to match the boundaries of each district. In the robustness tests, other definitions of aid have also been used and they are described in the text. The census was run by the Malawi National Statistical Office and is distributed by the Integrated Public Use Microdata Series (IPUMS, 2019) at the University of Minnesota. From IPUMS, the data are available as a systematic sample of every 10th household with a random start, which was drawn by the Minnesota Population Center to preserve the anonymity of respondents while preserving the representativeness of the data. For additional details on the sampling, see: https://international.ipums.org/international-action/sample_details/country/mw#tab_mw2008a.

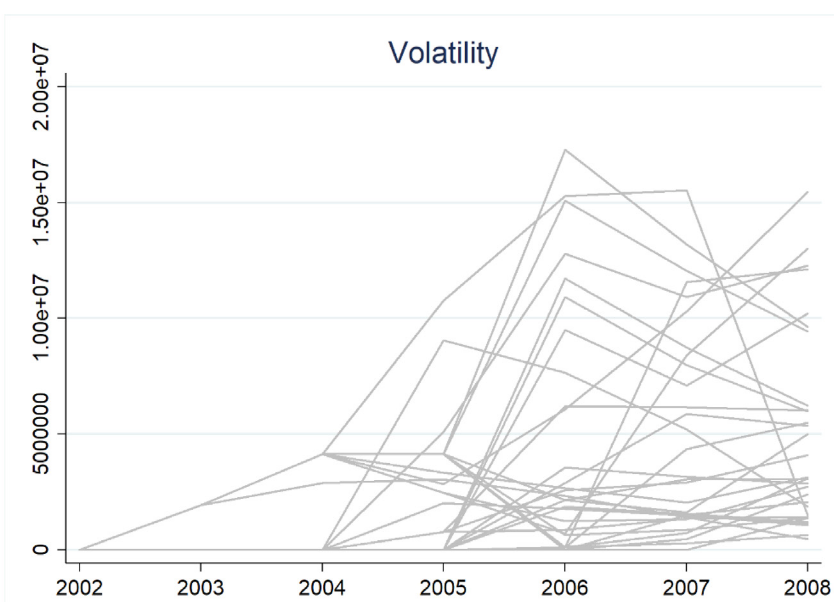
*Neither the census-based data we use in the main analysis, nor the DHS data we employ in one of our robustness tests distinguish between temporary vs permanent mobility (or seasonal migration).

** The distinction between **rural vs urban** migration flows is based on the information provided by the census, which reports the rural/urban status of each respondent. We grouped districts according to the share of individuals living in urban areas in 2008. Using the sample mean (around 10 %), we classified as "urban" the districts reporting higher shares and "rural" all the others. The group of urban districts corresponds to those hosting the major towns of the country, e.g. Karonga, Rumphi, Mzuzu, Lilongwe City, Mwanza, Zomba City, Blantyre City. The data are fairly consistent with national level data from the World Development Indicators, which report that the rural population accounted for around 85 % of the total in 2008 (see: https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=MW, accessed on November 16th, 2020).

Table A3
Migration flows by age groups.

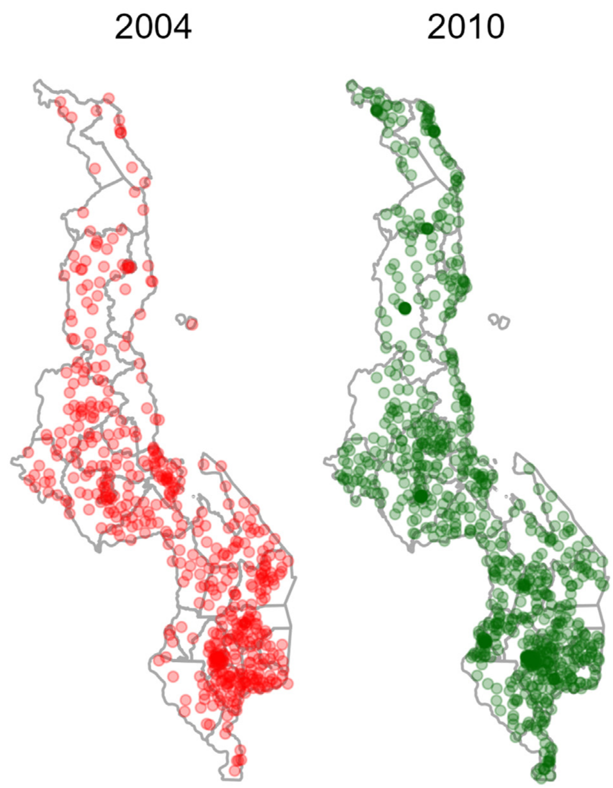
Year	Children	Working age	Elderly	Total
1998	105,100	70,160	2,050	177,310
1999	93,770	44,480	1,320	139,570
2000	125,880	70,250	1,750	197,880
2001	129,580	62,220	1,700	193,500
2002	125,750	64,230	1,670	191,650
2003	156,730	91,480	2,200	250,410
2004	174,000	96,750	2,010	272,760
2005	192,710	111,450	1,720	305,880
2006	201,360	125,800	1,840	329,000
2007	204,930	114,470	1,690	321,090
2008	297,130	218,790	3,560	519,480

Notes: Migrants flows distribution by age group and year of migration. Children refers to migrants less than 15 years old. Working age include people between 18 and 64. Elderly includes all migrants aged 65 and more. *Source*: Authors' elaboration based on Malawi 2008 Census (IPUMS, 2019).



Notes: Time series of the yearly volume of concluded aid projects in each district (in constant US$).

Fig. A1. Volatility of Aid Disbursements by District.



Notes: Distribution of DHS Clusters based on cluster centroids. The clusters' density is representative of the overall population at regional (for the year 2004) and district (for the years 2010 and 2015) level. Source: Authors' elaboration of DHS data

Fig. A2. Distribution of DHS clusters.

Appendix B. Robustness Tests

Table B1
Robustness tests: measurement issues/aid variable.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimator	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML
Dep. Variable	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*	*Migrant Flows*
Aid Disbursements$_{j,t-2;T}$	**0.008**** **(0.003)**							
Number of Aid Projects$_{j,t-3;t-1}$		**0.130***** **(0.019)**						
Number of Aid Projects$_{j,t-2;t}$			**0.134***** **(0.023)**					
Aid Disbursements$_{j,t-1}$				**0.003+** **(0.002)**				
Number of Aid Projects$_{j,t-1}$					**0.058**** **(0.025)**			
Stock of Aid Disbursements $_{j,t}$						**0.102***** **(0.017)**		
Commitment for Aid Projects$_{j,t-3;t-1}$							**0.007***** **(0.002)**	
Disbursement Incomplete Proj.$_{j,t-3;t-1}$								0.000

Table B1 (*continued*)

Estimator Dep. Variable	(1) PPML *Migrant Flows*	(2) PPML *Migrant Flows*	(3) PPML *Migrant Flows*	(4) PPML *Migrant Flows*	(5) PPML *Migrant Flows*	(6) PPML *Migrant Flows*	(7) PPML *Migrant Flows*	(8) PPML *Migrant Flows*
								(0.003)
Observations	10,054	10,054	10,054	10,054	10,054	10,054	10,054	10,054
% Null	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22
Adj. R2	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Standard errors clustered by destination in parentheses. All columns estimates the impact of different, alternative definitions of foreign aid on bilateral internal migration, using same econometric specification reported in Column (1) of Table 2 **Aid Disbursements**$_{j,t-2;t}$ refers to the 3-year average of total aid disbursements received by the destination district over the current and the previous 2 years (expressed in constant US$). **Number of Aid Projects**$_{j,t-3;t-1}$ and **Number of Aid Projects**$_{j,t-2;t}$ report the effect of aid as considered in the baseline and in Column (1), but computed on the number of projects' locations of concluded projects rather than their value. **Aid Disbursements**$_{j,t-1}$ and **Number of Aid Projects**$_{j,t-1}$ refers to the single-year volume (number) of aid projects, concluded in the destination district *j* in the previous year. **Stock of Aid Disbursements** $_{j,t}$ refers to the cumulate disbursements for concluded projects (in constant US$). **Commitment for Aid Projects**$_{j,t-3;t-1}$ and **Disbursement for Incomplete Proj.**$_{j,t-3;t-1}$ also refers to a lagged 3-years average, but reports the volume of commitments and of incomplete projects respectively (both expressed in constant US$). Consistently with PPML functional form, All aid variables are expressed in logs.

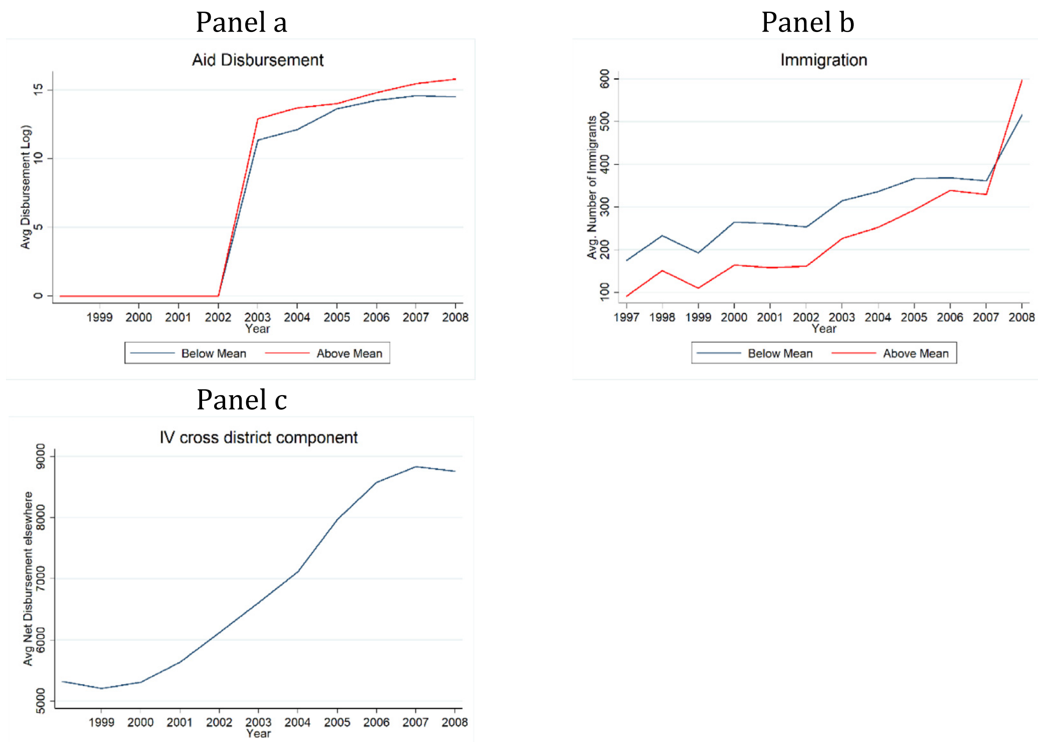
Table B2
Robustness tests: alternative specifications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Type of Robustness Test	***Alternative Estimators***		***Alternative Sets of Fixed Effects***				***Robustness to sample selection***			
	Pooled OLS	EK Tobit	Pair Only	Pair + T	Pair + O + T	O + T + D	No Top Destination	No Top Migr. Corridors	No Top Recipients	No Zero Aid Flows
Aid Disbursements$_{j,t-3;t-1}$	**0.008**** **(0.003)**	**0.007***** **(0.002)**	**0.040***** **(0.005)**	**0.009***** **(0.002)**	**0.009***** **(0.002)**	**0.009***** **(0.002)**	**0.008***** **(0.001)**	**0.007***** **(0.002)**	**0.008***** **(0.002)**	**0.011***** **(0.004)**
Observations	10,230	10,230	10,054	10,054	10,054	10,230	9,064	9,999	9,075	5,436
% Null	–	–	0.22	0.22	0.22	0.23	0.23	0.22	0.23	0.16
Adj. R2	0.85	0.85	0.94	0.95	0.95	0.35	0.95	0.96	0.96	0.96
Origin*Year FE	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes
District Pair FE	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	No	No	No	No
Origin FE	No	No	No	No	Yes	Yes	No	No	No	No
Destination FE	No	No	No	No	No	Yes	No	No	No	No

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Robust Standard errors clustered by destination in parentheses. With the exclusion of Columns (1) and (2), the dependent variable is the flow of migrants at time t. The dependent variable of the OLS model in Column (1) is log $(1 + N)$, where N is the flow of migrants from district *i* to district *j*. The dependent variable in Column (2) is also represented in logs. The data are then considered as left-censored, where the value of the censoring is set to the lower non-null value of bilateral migration recorded for each pair of districts. Columns (3) to (6) report the coefficients from PPML, fitted including different sets of fixed effects (Pair = Origin * Destination; O = Origin; D = Destination; T = Time). The models in Columns (7), (8) and (9) are based on the specification of Column (1) in Table 2 and test the robustness of the baseline estimates to the exclusion of the three top migrant destinations (Blantyre City, Thyolo, Lilongwe City, and Mulanje), the five major bilateral corridors (Thyolo to Balaka, Blantyre to Mulanje, Chikwawa to Blantyre City, Blantyre (District) to Zomba City, and Phalombe to Chikwawa), and the top three aid recipients district respectively (Karonga, Lilongwe (District), and Mangochi). Finally, Column (10) replicates the specification of Column (1) from Table 2, but limiting the sample to the period 2003–2008 (that is, removing the years in which no aid projects among those considered was concluded).

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.worlddev.2022.106134.



Notes: Panel a and b compare the trend of both our variable of interest (log of the 3-years average of concluded disbursements) and the dependent variable (immigration in recipient districts) in districts receiving less than the average aid with respect to those receiving more than the average aid disbursement in a given year t. Panel c reports the trend over time of the average of the time varying component of our instrument.

Fig. B1. Shift-Share Instrument.

References

Adovor, E., Czaika, M., Docquier, F., & Moullan, Y. (2021). Medical brain drain: How many, where and why? *Journal of Health Economics, 76*, 102409.

Anglewicz, P., Kidman, R., & Madhavan, S. (2019). Internal migration and child health in Malawi. *Social Science & Medicine, 235*, 112389.

Ardington, C., Case, A., & Hosegood, V. (2009). Labor supply responses to large social transfers: Longitudinal evidence from South Africa. *American Economic Journal: Applied economics, 1*(1), 22–48.

Ariu, A., Docquier, F., & Squicciarini, M. P. (2016). Governance quality and net migration flows. *Regional Science and Urban Economics, 60*, 238–248.

Arndt, C., Jones, S., & Tarp, F. (2010). Aid, growth, and development: Have we come full circle? *Journal of Globalization and Development, 1*(2).

Becerra-Valbuena, L. G., & Millock, K. (2021). Gendered migration responses to drought in Malawi. *Journal of Demographic Economics*, 1–41.

Beine, M., & Parsons, C. (2015). Climatic factors as determinants of international migration. *The Scandinavian Journal of Economics, 117*(2), 723–767.

Beine, M., & Parsons, C. R. (2017). Climatic factors as determinants of international migration: Redux. *CESifo Economic Studies, 63*(4), 386–402.

Beine, M., Bertoli, S., & Fernández-Huertas Moraga, J. (2016). A practitioners' guide to gravity models of international migration. *The World Economy, 39*(4), 496–512.

BenYishay, A., Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L., and Runfola, D. (2017). Geocoding afrobarometer rounds 1-6: Methodology & data quality. *AidData (Available online at http://geo.aiddata.org.)*.

Berthélemy, J. C., Beuran, M., & Maurel, M. (2009). Aid and migration: Substitutes or complements? *World Development, 37*(10), 1589–1599.

Bertoli, S., & Fernández-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics, 102*, 79–100.

Bertoli, S., & Fernández-Huertas Moraga, J. (2015). The size of the cliff at the border. *Regional Science and Urban Economics, 51*, 1–6.

Borusyak, K., Hull, P., & Jaravel, X. (2021). Quasi-experimental shift-share research designs. *The Review of Economic Studies*. https://doi.org/10.1093/restud/rdab030.

Briggs, R. C. (2018). Poor targeting: A gridded spatial analysis of the degree to which aid reaches the poor in Africa. *World Development, 103*, 133–148.

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica, 82*(5), 1671–1748.

Cai, S. (2020). Migration under liquidity constraints: Evidence from randomized credit access in China. *Journal of Development Economics, 142*, 102247.

Cattaneo, C., & Peri, G. (2016). The migration response to increasing temperatures. *Journal of Development Economics, 122*, 127–146.

Chauvet, L., & Ehrhart, H. (2018). Aid and growth: Evidence from firm-level data. *Journal of Development Economics, 135*, 461–477.

Clark, D. E., Herrin, W. E., Knapp, T. A., & White, N. E. (2003). Migration and implicit amenity markets: Does incomplete compensation matter? *Journal of Economic Geography, 3*(3), 289–307.

Clemens, M. A., Radelet, S., Bhavnani, R. R., & Bazzi, S. (2012). Counting chickens when they hatch: Timing and the effects of aid on growth. *The Economic Journal, 122*(561), 590–617.

CHAI - Clinton Health Access Initiative Malawi (2015). Resource Mapping in Malawi, Case Study. http://www.clintonhealthaccess.org/content/uploads/2015/08/Case-Study_HF-Resource-Mapping.pdf.

Clist, P., & Restelli, G. (2021). Development aid and international migration to Italy: Does aid reduce irregular flows? *The World Economy, 44*(5), 1281–1311.

De, R., & Becker, C. (2015). The foreign aid effectiveness debate: Evidence from Malawi. *AidData Working Paper #6.* Williamsburg, VA: AidData at William & Mary.

Dolan, C. B. (2018). Health aid projects have both expanded and constrained the capacity of health facilities to deliver malaria services to under-five children in Malawi. *BMJ Global Health, 3*(6), e001051.

Dreher, A., & Lohmann, S. (2015). Aid and growth at the regional level. *Oxford Review of Economic Policy, 31*(3–4), 420–446.
Dreher, A., Fuchs, A., & Langlotz, S. (2019). The effects of foreign aid on refugee flows. *European Economic Review, 112*, 127–147.
Dreher, A., Fuchs, A., Parks, B., Strange, A. M., & Tierney, M. J. (2021). Aid, China, and growth: Evidence from a new global development finance dataset. *American Economic Journal: Economic Policy, 13*(2), 135–174.
Dustmann, C., & Okatenko, A. (2014). Out-migration, wealth constraints, and the quality of local amenities. *Journal of Development Economics, 110*, 52–63.
Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica, 70*(5), 1741–1779.
FAO IFAD IOM WFP (2018). The Linkages between Migration, Agriculture, Food Security and Rural Development. Technical Report (available at http://www.fao.org/3/CA0922EN/CA0922EN.pdf). Licence: CC BY-NC-SA 3.0 IGO.
Galiani, S., Knack, S., Xu, L. C., & Zou, B. (2017). The effect of aid on growth: Evidence from a Quasi-experiment. *Journal of Economic Growth, 22*, 1–33.
Gollin, D., Kirchberger, M., & Lagakos, D. (2021). Do urban wage premia reflect lower amenities? Evidence from Africa. *Journal of Urban Economics, 121*, 103301.
Harari, M., & La Ferrara, E. (2018). Conflict, climate, and cells: A disaggregated analysis. *Review of Economics and Statistics, 100*(4), 594–608.
Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review, 60*(1), 126–142.
Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics* (pp. 131–195).
Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review, 102*(2), 994–1028.
Henderson, J. V., & Turner, M. A. (2020) Urbanization in the Developing World: too early or too slow? NBER Working Paper N. 27201
IPUMS. (2019). Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minnesota Population Center, Minneapolis (MN) https://doi.org/10.18128/D020.V7.2.
Khomba, D. C., & Trew, A. (2022). Aid and local growth in Malawi. *The Journal of Development Studies*. https://doi.org/10.1080/00220388.2022.2032668.
Kirchberger (2021). Measuring internal migration. *Regional Science and Urban Economics*, 103714. https://doi.org/10.1016/j.regsciurbeco.2021.103714. ISSN 0166-0462.
Kotsadam, A., Østby, G., Rustad, S. A., Tollefsen, A. F., & Urdal, H. (2018). Development aid and infant mortality. Micro-level evidence from Nigeria. *World Development, 105*, 59–69.
Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives, 34*(3), 174–192.
Lanati, M., & Thiele, R. (2018). The impact of foreign aid on migration revisited. *World Development, 111*, 59–74.
Lanati, M., & Thiele, R. (2020). Foreign assistance and emigration: Accounting for the role of non-transferred aid. *The World Economy, 43*, 1951–1976.
Lucas, R. E. B. (2015). Internal migration in developing economies: An overview, KNOMAD Working Paper.
Martorano, B., Metzger, L., & Sanfilippo, M. (2020). Chinese development assistance and household welfare in sub-Saharan Africa. *World Development, 129*, 104909.
Miller Runfola, D., & Napier, A. (2016). Migration, climate, and international aid: Examining evidence of satellite, aid, and micro-census data. *Migration and Development, 5*(2), 275–292.
Mishra, P., & Newhouse, D. (2009). Does health aid matter? *Journal of Health Economics, 28*(4), 855–872.
Moullan, Y. (2013). Can foreign health assistance reduce the medical brain drain? *The Journal of Development Studies, 49*(10), 1436–1452.
Nunn, N., & Qian, N. (2014). US food aid and civil conflict. *American Economic Review, 104*(6), 1630–1666.
Odokonyero, T., Ijjo, A. T., Marty, R., Muhumuza, T., & Moses, G. O. (2018). The impact of aid on health outcomes in Uganda. *Health Economics, 4*, 733–745.
Ortega, F., & Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies, 1*(1), 47–74.
Pickbourn, L., Caraher, R., & Ndikumana, L. (2022). Does project-level foreign aid increase access to improved water sources? Evidence from household panel data in Uganda. *World Development, 158*, 105994.
Qian, N. (2015). Making progress on foreign aid. *Annual Review of Economics, 7*(1), 277–308.
Schmidheiny, K., & Brülhart, M. (2011). On the equivalence of location choice models: Conditional logit, nested logit and Poisson. *Journal of Urban Economics, 69*(2), 214–222.
Santos-Silva, J., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics, 88*(4), 641–658.
Santos-Silva, J., & Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters, 112*(2), 220–222.
Tierney, M. J., Nielson, D. L., Hawkins, D. G., Roberts, J. T., Findley, M. G., Powers, R. M., & Hicks, R. L. (2011). More dollars than sense: Refining our knowledge of development finance using AidData. *World Development*, *39*(11), 1891-1906.
UNDP (2009). Human development report: Mobility and migration. United Nations Development Program.
World Bank (2016). *Malawi urbanization review: Leveraging urbanization for national growth and development*. The World Bank.
Young, A. (2013). Inequality, the urban-rural gap, and migration. *The Quarterly Journal of Economics, 128*(4), 1727–1785.