Università degli Studi Firenze

DISEI
Dipartimento di Scienze per l'Economia e l'Impresa

Working Papers - Economics

The Impact of Regional Growth on Internal Migration: A District-Level Analysis for Malawi

Mauro Lanati and Rainer Thiele

Working Paper N. 15/2025

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

The Impact of Regional Growth on Internal Migration: A District-Level Analysis for Malawi

by

Mauro Lanati* and Rainer Thiele**

Abstract: Research on the nexus between development and migration has mainly focused on cross-border flows. How income changes affect migration within developing countries is much less well researched even though addressing this topic might provide essential information about the process of structural transformation needed for economic development. In this paper, we provide new evidence on the link between income growth and internal migration for Malawi, one of the poorest countries worldwide where migration is predominantly internal. Employing a gravity approach and performing an instrumental variable regression based on a shift-share instrument, we robustly find that, on average, rising incomes – proxied by changes in nightlight intensity – are associated with higher emigration rates. This effect is mainly driven by people emigrating from comparably richer urban areas. In the poorer rural districts, by contrast, migration tends to fall with increasing economic activity, which is in accordance with the notion that poverty may force people to leave their home in response to negative shocks. Our results also suggest a specific sorting pattern by education levels: While in urban areas rising incomes mainly facilitate the emigration of lower-skilled people to non-urban destinations, in rural areas it is higher-skilled people who most likely leave their home in response to falling incomes.

JEL Codes: O55; R23; O01
Keywords: Economic Development; Internal Migration; Malawi; Sub-Saharan Africa

* Department of Economics and Management, University of Florence (Italy). Email: mauro.lanati@unifi.it
** Kiel Institute for the World Economy, Kiel University, Kiel (Germany). Email: Rainer.Thiele@ifw-kiel.de
We thank Luca Tiberti, Sarah Langlotz, Lennart Kaplan and Filippo Santi for helpful comments and suggestions. We also acknowledge helpful comments from participants at the German Development Economics Conference (2025) in Frankfurt, the Workshop on Globalization and Development in Göttingen and seminar presentations in Göttingen and Florence.

1. Introduction

Research on the nexus between development and migration dates at least back to the seminal paper by Harris and Todaro (1970), who proposed that (expected) income differentials between (urban) destinations and (rural) origins are the key determinant of migration decisions within countries. The topic has recently regained momentum at the international level, partly in response to large numbers of arrivals of irregular migrants from developing countries which prompted policy makers in potential OECD destinations to stress the importance of tackling the root causes of international migration through raising living standards at the origin. However, policy makers' presumption that migration can be reduced by supporting local economic development has been challenged in the academic literature that accompanied the recent debate. The predominant view is that a hump-shaped relationship exists between home-country incomes and emigration pressure (e.g. Clemens, 2014). This would imply that higher incomes in developing countries, which tend to be located on the rising part of the inverted U-curve, lead to more rather than less emigration. The basic economic mechanism behind the migration hump is that at low initial levels of per capita GDP, additional income facilitates emigration for liquidity-constrained individuals in countries of origin, thus raising the number of people who leave. At some point, the liquidity constraint becomes less binding, so further increases in real incomes cause the emigration rate to fall from its peak through incentive effects a la Harris and Todaro (1970).

Evidence supporting the migration-hump hypothesis mainly comes from cross-country studies (e.g. Dao et al. 2018; Djajic et al. 2016). The cross-country estimates are best interpreted as capturing the long-term association between economic development and emigration. By contrast, most panel data studies of the development-migration nexus, which tend to capture short- to medium-term effects, point to a universal negative income-migration relationship once cross-country heterogeneity is accounted for through the inclusion of an appropriate set of fixed effects (e.g. Benček and Schneiderheinze 2024; Böhme et al. 2020; Lanati and Thiele 2024). An exception is Clemens (2020), who finds that increasing GDP per capita is on average associated with more emigration in poor countries and that the effect reverses only after GDP per capita exceeds about $10,000.

The relationship between income changes and migration within developing countries, to which the original Harris and Todaro (1970) paper referred, is much less well researched (Lucas 2016). This is even though in low-income settings internal migration often predominates and understanding how regional income gaps are linked to internal migration may provide essential information about the process of structural transformation needed for economic development, as this process is typically associated with massive population movements from rural to urban areas, and from agriculture to non-agricultural activities (Lagakos 2020). If financial constraints or other barriers prevent a significant number of people from migrating despite large existing income gaps, this might contribute to a slowdown of structural transformation. As illustrated by Bryan and Morten (2019) for the case of Indonesia, encouraging internal migration in developing countries can give rise to substantial productivity gains.

Empirical studies for two middle-income countries suggest that the migration hump hypothesis might also be relevant in the context of internal migration. Employing census data from Indonesia, Bazzi (2017) estimates that positive income shocks in poor rural areas increase emigration, whereas the opposite effect occurs for the most developed regions within the country. Kleemans (2023) devises a model that distinguishes two different motivations for migration: It can either be used as an ex-post risk-coping strategy in response to negative income shocks or serve as an investment, in which case positive income shocks may increase migration by overcoming liquidity constraints that prevent households from paying up-front migration costs. Also looking at individual migration decisions in Indonesia, she provides evidence that is in accordance with her model predictions. Additional migration after negative shocks often involves temporary moves to nearby rural areas, whereas migration as an investment strategy is more likely to be directed to urban destinations and take place over longer durations. For Columbia, Acosta and Gu (2022) find a hump shape when they restrict their analysis to out-migration from rural areas. In the full sample, however, the effect turns around to form a u-shape. Hirvonen (2016) uses panel data from the overwhelmingly rural Kagera region in Tanzania and examines the impact of weather-induced income shocks on migration; his findings indicate that the low rates of internal migration in the country can, at least partly, be explained by liquidity constraints.

In this paper, we provide new evidence on the link between income growth and internal migration for Malawi, one of the poorest countries worldwide, which in the period under consideration stands out as exhibiting the highest rate of internal migration in Sub-Saharan Africa (Morten 2015). Our empirical analysis is based on a gravity model of migration. We proxy regional economic activity by the intensity of nightlights and derive district-to-district migration data from the most recent population census based on which it is possible to construct a district-to-district panel dataset for internal migration flows.[1] To address endogeneity concerns emanating particularly from measurement error in the nightlight data, we perform an instrumental variable (IV) regression along the lines of Gehring et al (2022). Specifically, we exploit the availability of geo-coded aid data for Malawi and construct a shift-share instrument that interacts an exogenous component based on total aid commitments by development banks and an endogenous component based on district-specific aid disbursements.

Our contribution to the literature is threefold. First, in substantive terms, this study is the first that investigates the relationship between income growth and internal migration for a low-income country, where liquidity constraints are expected to be particularly binding. At the same time, a sizeable share of the poorest population might be forced to emigrate if, for instance, falling agricultural yields push their income below thresholds needed for survival. Second, our analysis comes closer to establishing a causal interpretation of the growth-migration link than previous papers, most notably those dealing with international migration. At the sub-national level, on which we focus here, the problem of unobserved cross-sectional heterogeneity tends to be less severe than in multi-country settings with strongly varying national context conditions. In addition, our IV approach mitigates potential endogeneity concerns. Third, we disaggregate migration by education level, which enables us to provide new evidence for a single low-income economy on the issue of migrant sorting based on differences in human capital, complementing Young's (2013) analysis that covers a broad cross-section of developing countries.

We find that, on average, rising incomes are associated with higher emigration rates. This effect is mainly driven by people emigrating from comparably richer urban areas. In the poorer non-urban districts, by contrast, migration tends to fall with increasing economic activity, which is in accordance with the notion that extreme poverty may force people to leave their home. Our results also suggest a specific sorting pattern by education levels: While in urban areas rising incomes mainly facilitate the emigration of lower-skilled people to non-urban destinations by loosening

[1] More recent Census data for Malawi from the 2018 wave only include information on migration status and residency one year prior to the survey which doesn't allow us to construct a panel for district-to-district internal migration flows.

budgetary constraints, in rural areas it is higher-skilled people who most likely leave their home in response to falling incomes, in many cases moving to urban areas.

The remainder of the paper is structured as follows. In section 2, we introduce our econometric approach. Section 3 describes the data used in the empirical analysis, putting a particular emphasis on discussing the nighttime-light (NTL) intensity variable that serves as a proxy for regional economic activity. Section 4 presents the estimation results. The paper closes with some concluding remarks.

2. Empirical Specification

Our baseline specification relies on a gravity model of migration (e.g. Beine and Parsons 2015, Beine and Parsons 2017), where internal bilateral migrant flows from district *i* to district *j* at time *t* are a function of the average intensity of per capita nighttime lights at the origin. Specifically, we obtain:

$$\frac{\mathrm{M}_{ij,t}}{\mathrm{M}_{ii,t}} = exp\left[S_i + S_{j,t} + \ ln\left(\mathrm{MigStocks}_{ij,t}\right) * \delta + ln(NTLpc_{i,t-1}) * \gamma + X'_{i,t} * \partial + \tau_{ij,t}\right] \quad (1)$$

$\frac{\mathrm{M}_{ij,t}}{\mathrm{M}_{ii,t}}$ is the bilateral district-to-district emigration rate along the lines of Beine and Parsons (2015). $\mathrm{M}_{\mathrm{ii,t}}$ is recovered from population data by subtracting the total number of migrant flows in district *i*. *MigStocks* capture the standard network effect which is proxied by the stock of emigrants from district *i* residing in district *j* at time *t*. S_i and $S_{j,t}$ are origin and destination-year fixed effects, respectively. $X'_{i,t}$ is a vector of time-varying origin-specific control variables, and *NTLpc* denotes per-capita nightlights, our main variable of interest. *NTLpc* enters the model with a lag to partially mitigate potential reverse causality, i.e. the possibility that emigration affects income levels at the origin. The controls include proxies for standard push factors in the district of origin – conflict and climate shocks – as well as the district's population density. In one specification, we also add district-pair fixed effects to capture all the main dyadic factors - such as cultural and geographical distance - that influence emigration decisions and thereby attenuate a possible omitted-variable bias.

However, equation (1) does not fully account for multilateral resistance to migration (Beine et al. 2016), given that the inclusion of origin-year dummies would completely absorb the effect of our origin-specific variable of interest. In addition, relying solely on the dyadic setup would not allow us to address endogeneity issues in a methodologically sound way by performing an IV analysis. If the multilateral resistance term is not properly accounted for and enters the error component, this renders any relevant instrument invalid by definition as it creates a correlation between the instrumental variable and the residuals.[2]

Different sources of endogeneity, most notably measurement error, are likely to remain even if we employ lagged NTL intensity.[3] Hence, we additionally rely on a monadic setup to carry out an IV analysis, which reduces to:

$$\frac{\sum_{j=1}^{J} \mathrm{M}_{ij,t}}{\mathrm{M}_{ii,t}} = \beta_i + \beta_t + \ln\left(NTLpc_{i,t-1}\right) * \varphi + X'_{i,t} * \partial + \epsilon_{i,t} \tag{2}$$

where β_i and β_t are district and year fixed effects, respectively. In equation (2), we instrument NTL per capita using a shift-share type of instrument along the lines of Dreher et al. (2019, 2021) and Gehring et al (2022). The first stage regression at the district-time level becomes:

$$\ln\left(NTLpc_{i,t-1}\right) = \beta_i + \beta_t + \left(AvgAid_i * BCR_t\right) * \alpha + X'_{i,t} * \partial + \tau_{i,t} \tag{3}$$

where $X'_{i,t}$, β_i and β_i are the same set of control variables, origin and year fixed effects as in Equation (2). Our preferred instrument is the interaction between a time-varying exogenous component BCR_t and a district-specific and time-invariant endogenous term $AvgAid_i$. The first term is defined as the Regional Development Banks' (RDBs) and World Bank's total commitments to all recipients but Malawi in year t divided by the total volume of DAC Donors' Capital Subscriptions in the same year. Capital subscriptions are financial contributions made by OECD-DAC members and other donors to multilateral development banks like the World Bank and regional development banks, which constitute a key resource for these institutions to mobilize

[2] It must be noted that origin-time dummies account for multilateral resistance to migration only when Ortega and Peri's (2013) distributional assumption on the stochastic component holds, i.e. that the multilateral resistance term does not vary across destinations (see Beine et al. 2016).

[3] Gibson (2021) cautioned against using NTL data taken from DMSP as their inaccuracies can cause mean-reverting errors, which implies that spatial inequality will be understated if these data are used to measure patterns of economic activity. However, as Pischke (2007) pointed out, we can still get consistent estimates of the true beta in an IV setup as long as the instruments are only correlated with true Xs but not with any of the measurement errors. We believe this is the case with our instruments, as the decision of the development banks to allocate aid, which we use as the basis of our IV, does not depend either on the capacity of NTL to capture economic activity in a given district, or on the inaccuracies coming from the DMSP satellite images. Rather, aid allocation most likely depends on the actual socio-economic characteristics of a given ADM1 district.

funds and provide development assistance to eligible recipients. The ratio between RDBs' and the World Bank's aid commitments and the volume of capital subscriptions is often greater than 1, as these institutions can spend more than the volume of funds they receive from donor countries, due to their ability to leverage their capital base through borrowing in international markets and mobilizing resources from other sources. This variable can be regarded as exogenous, given that the extent to which the development banks go to the market to reap more funds compared to what they receive from the DAC group is due to factors that have no relation to the socio-economic situation in Malawian districts. The term BCR_t is multiplied by the average per-capita World Bank disbursements $AvgAid_i$ received by district i in Malawi over the period 1997-2007, thus yielding an instrumental variable with an i,t dimension. We predict a sizable correlation between aid disbursements and local economic conditions. Shon et al. (2024) reveal a consistent pattern across a broad cross-section of 38 Sub-Saharan African countries: more impoverished regions receive significantly less aid as compared to highly urbanized and richer regions. Along similar lines, Briggs (2017 and 2021) focusses on multilateral aid flows and finds that foreign assistance from the World Bank and the African Development Bank (AfDB) have been allocated to areas with higher population, greater per-capita income, and better infrastructure. This positive spatial correlation between aid and economic opportunities also applies to the specific case of Malawi. Figure 1 compares the allocation of the average value of World Bank aid projects across Malawi's districts with the average NTL intensity levels; it shows a strong concentration of projects in the capital Lilongwe and other major cities where NTL intensity is high.

The intuition behind the validity of the shift-share IV approach follows the logic of a difference-in-difference method (see for instance Dreher et al. 2019). The identifying assumption is that emigration from districts with differing probabilities of receiving high volumes of aid will not be affected differently by changes in the capacity of the World Bank and RDBs to finance aid projects beyond the volume of capital subscriptions - other than via the impact of NTL. Figure 2 shows that trends in NTL intensity and emigration are parallel across the two groups, supporting our assumption on the excludability of the instrument.

In a robustness check, we employ two alternative IVs. The first alternative is also a shift-share instrument but constructed with a different endogenous component, the probability of receiving aid defined as $\bar{P}_{i,w} = \frac{1}{11}\sum_{y}^{11} P_{i,w,y}$, where $P_{i,w,y}$ is a binary indicator variable that is one when recipient *i* received a positive amount of bilateral aid from the World Bank in year *y* (see Dreher et al. 2019 and 2021). While both shift-share instruments pass the usual diagnostic tests assessing their validity and strength, we prefer the one that uses actual disbursements because F-test results suggest that it is slightly stronger. We also propose the per-capita value of World Bank future committed aid projects (t+1, t+2) in the economic and production sectors as alternative IVs. The rationale behind the choice of these IVs is that future aid commitments targeted at the economic and production sectors of a specific district depend on previous economic conditions, while current emigration decisions are unlikely to be sensitive to future allocations of committed aid projects by the World Bank. There might be time-varying factors that potentially influence both future aid allocations and current emigration decisions, but these are more likely to be country-specific (e.g. macro shocks), or they are district-specific but do not exhibit significant variation over time and their effects are therefore absorbed by the set of fixed effects (e.g. ethnic fractionalization). The validity of the proposed alternative instruments is substantiated by exogeneity tests and first-stage statistics (Appendix Tables B7 and B8).

3. Data

We employ the 2008 Population Census of Malawi to construct a retrospective panel of district-to-district migration over the period 1998–2008. We define as migrants all those individuals who moved to the current district of residence at the ADM1 level from any other district in a specific year. Along the lines of Lanati et al (2023), starting from the year 2008 and going backwards until 1998, we aggregate over origin and destination districts to build a dyadic matrix of annual migrant flows.[4]

We also exploit Census information on the skill level of internal migrants. Migrants are classified according to four different levels of education: less than primary, primary, secondary and tertiary.[5]

[4] It has to be noted that the IPUMS anonymization procedure allows to track the previous residence of migrants only at the district level. This limitation prevents us from being more granular at the geo-localized level as well as from considering within-district movements. While within-district movements are not negligible in numbers, they mostly involve short-distance moves for marriage-related purposes, family reasons or relocations to nearby villages for farmland, which are likely to be temporary movements (see Becerra-Valbuena & Millock, 2021).

[5] There is a discrepancy between bilateral flows included in the baseline and the correspondent aggregate internal movements for which information on education are available. This is because information on education is missing in the Census or simply the respondents sometimes fall outside the defined population or target group for that variable and are therefore classified as NIU (not in universe). In Table B1, we compare our baseline results with the corresponding results using aggregate flows for which the information on education levels is available and find no majors differences.

As recently argued by Clemens and Mendola (2024), migrant selection depends – among other factors - on rural-urban location and is strongly positive on education. Along similar lines, yet with a more specific focus on within-country mobility in developing countries, Young (2013) reveals a clear sorting on education by region and type of emigration. More specifically, by relying on Demographic and Health Surveys from a diverse and large set of developing countries, he finds that rural-to-urban emigrants have education levels that are significantly higher compared to the ones who decide to stay or move to other rural areas. On the contrary, urban-to-rural migrants have substantially lower education levels in comparison to those who remain at their origin or move to other urban areas.[6] Both flows are in accordance with a model where populations sort themselves geographically on the basis of skills, for which there is relatively higher demand in urban areas (Young 2013).[7]

To capture the rural-urban dimension of migration, we classify as urban the four major metropolitan areas of Blantyre, Lilongwe, Mzuzu and Zomba which host most non-agricultural economic activities, while the remaining districts – including secondary cities, e.g., townships and district centers – where more than 80% of the total population live are predominantly rural and agrarian (McBride and Moucheraud 2022). Not surprisingly, the four districts identified as urban areas are the ones that exhibit the highest levels of per capita nightlight intensity in our period under consideration.

We also rely on information from alternative data sources – namely the Malawi's LSMS Second Integrated Household Survey 2004-2005 and the Third Integrated Household Survey 2010-2011- to corroborate our main econometric results and provide descriptive evidence on the channels through which the income effect manifests itself.

The descriptive statistics reported in Figure 3 and Appendix Table A1 show that internal migrants in Malawi are predominantly low-skilled without any school education and move mainly from rural to other rural districts. The share of those who move to urban areas increases with the level of education. In urban areas, where education is more valuable and income opportunities are greater, most internal emigrants are low-skilled and move to rural districts. These results confirm

[6] According to Young (2013), both rural-to-urban and urban-to-rural migration is sizeable, with 20 to 25 percent of individuals, on average, moving in each direction as young adults.
[7]Note that while skills are unobservable, they are correlated with education and therefore can be approximated empirically.

the empirical findings on within-country migration that Young (2013) obtained for a large sample of developing countries.

Our measure of economic activity is based on satellite data displaying the per-capita average intensity of NTLs. The main advantage of night-time light intensity is its availability at the regional level, which is particularly useful in the African context where regional GDP estimates are typically poor or unavailable. As concerns reliability, Gibson et al. (2021) pointed to low predictive performance of NTL data for lower-level spatial units, lower-density areas, and smaller areas. However, the predictive power of NTLs is shown to be higher for larger spatial units such as the ADM1 regional districts considered in our empirical analysis. At the country level, Henderson et al. (2012) find a high correlation between changes in NTL intensity and GDP, while Baskaran et al (2024) document a similarly strong relationship between per capita nightlights and per capita GDP. Hodler and Raschky (2014) corroborate the high correlation for subnational administrative regions covering a broad set of over 120 countries, and Asher et al (2021) show that the time-variation of nightlights in India appears to be a good proxy for economic opportunities such as manufacturing and services employment. For the case of Malawi, the NTL statistics reported in Figure 2 (bottom panel) accurately capture the economic downturn that occurred in the country around 2002-2003.[8]

The NTL data most frequently used in economics comes from the Defense Meteorological Satellite Program (DMSP), which has well-documented limitations in predicting GDP (Gibson 2021, Gibson et al. 2021). Newer data with better predictive performance from the Visible Infrared Imaging Radiometer Suite (VIIRS) exists but cannot be matched with the information on Malawi's internal migration because VIIRS data have only become available after 2012. Instead, we employ the dataset introduced by Li et al. (2020). While raw DMSP NTL time series data are not comparable across years due to the lack of on-board calibration, varied atmospheric conditions, satellite shift, and sensor degradation, Li et al. (2020) generate a globally harmonized and temporally consistent NTL data set from 1992 to 2018 through the integration of DMSP and VIIRS

[8] The downturn was caused mainly by the aftermaths of drought and food shortages, macro imbalances, rising inflation and the temporary suspension of budget support from several donors, including the IMF, EU and other bilateral donors such as UK, US, Denmark and others, that occurred roughly in the same period. See report by the Swedish International Development Cooperation Strategy, available at https://cdn.sida.se/publications/files/-evaluation-of-general-budget-support-malawi-country-report.pdf

data using a stepwise calibration approach which makes the yearly cross-sections comparable over time.

The NTL dataset includes a small fraction of spatial units with zero average nightlight values, which occurs either because the area is extremely sparsely populated or because the satellite sensors cannot capture dimly lit areas. In the baseline model, we drop the zeros but include them in two robustness checks employing the scaled-log transformation (e.g., Hodler and Raschky 2014) and the inverse hyperbolic sine (IHS) transformation (see, for example, Amare et al. 2020) of NTLs, respectively. These transformations leave our conclusions unaffected (see Tables B2 and B3).

Based on data provided by Harari and La Ferrara (2018), we proxy for climate shocks using the Standardised Precipitation-Evapotranspiration Index (SPEI) and employ a dummy for the presence of any form of conflict in the district of origin according to the Armed Conflict Location and Event (ACLED) database. Districts' population values are only available for the two years – 2008 and 1998 – in which the census was conducted. For the remaining years, they are obtained through linear interpolation. Information on World Bank aid disbursements by district that we need for our instruments comes from the Geocoded Official Development Assistance Dataset (GODAD) (Bomprezzi et al., 2025), while OECD-DAC data are used to construct the exogenous component of the IV.

Variable definitions and summary statistics of the variables for each of the 30 Malawian districts are shown in Tables A2 and A3.

4. Results

The notion of a hump-shaped relationship between income and international migration receives empirical support mainly from cross-sectional studies of south-north movements (e.g. Clemens 2014; Dao et al. 2018). To provide first indicative evidence on whether the migration hump is also confirmed in the setting of a low-income country such as Malawi, Figure 4 shows the cross-sectional correlation between its district-level NTL intensity and internal migration. The curve that emerges is broadly in accordance with the upward-sloping part of the migration hump; only at the lower tail of the income distribution, there is some indication of a negative income-migration link.

This pattern holds for the whole country as well as when we look at emigration from rural and urban districts separately. The absence of a turning point at higher income levels might reflect that even wealthier Malawians are too poor to get over the threshold where liquidity constraints no longer dominate incentive effects, while the negative slope for the poorest might point to desperation-led emigration that becomes less likely with rising incomes.[9] Figure 4 also highlights the existence of a large rural-urban gap in district-level NTL intensity.

Turning to a more rigorous regression analysis, we first report our baseline estimations of the gravity equation (1) with different sets of fixed effects (see Table 1). We use Poisson Pseudo Maximum Likelihood (PPML) as our preferred estimator to account for the sizeable share of zeros in the dependent variable (23.4 %), which is large enough to potentially bias the results of standard log-linear fixed effects models (Santos-Silva and Tenreyro 2006). Standard errors are clustered by district of origin. Our findings are robust to alternative ways of clustering standard errors – e.g. by district pairs (Table B4).

Across alternative specifications, higher per-capita nightlight intensity at the district of origin is robustly associated with a rise in internal migration, which contrasts with recent panel studies on international migration where the estimated income-migration relationship tends to switch from positive to negative when accounting for cross-sectional heterogeneity (e.g. Benček and Schneiderheinze, 2024; Lanati and Thiele 2024). In the Malawian context, improvements in economic conditions therefore on average facilitate internal emigration more strongly than they provide an incentive to stay in the district of origin. This result is not only statistically significant, but also economically relevant: a simple back-of-the-envelope calculation based on our baseline estimates shows that passing from the 1st percentile (extremely poor) to the 25th percentile of per capita NTL intensity leads to an average increase in the district-to-district emigration rate of 51.3%.

As concerns the control variables, it appears that only migrant stocks are shown to retain their significance when we include district-time and bilateral fixed effects (column 5), corroborating previous cross-country studies that identified existing migrant networks as a main driver of

[9] Figure 4 depicts the pooled correlation between per capita nightlight intensity and emigration rates, where multiple cross-sections are aggregated for the whole period 1998-2008. The yearly correlations reported in Figure A1 are largely in line with the results of the pooled analysis.

international migration (e.g. Lanati and Thiele 2018). Following Cattaneo and Peri (2016), we also estimated a parsimonious model that includes only the set of fixed effects with no additional controls. This specification has the advantage that it does not include variables that possibly could take up part of the overall income effect. According to Table B5, the size and significance of the coefficients of NTL intensity are hardly affected by the exclusion of the control variables, which suggests that a potential omitted-variable bias is not likely to be a major concern in our empirical analysis.

The positive relationship between per capita NTL intensity and emigration rates is confirmed by the monadic model estimates (Equation 2) reported in Table 2. The standard PPML estimates with and without fixed effects (columns 1-2) exhibit somewhat higher effect sizes as compared to the gravity model, and the IV analysis (columns 3-6) leads to coefficients that are still larger in magnitude when including the fixed effects, which points to a downward bias in the baseline estimates that might for example reflect measurement error.[10] The first-stage results of the IV estimation and the reduced form regression provide support for the validity of the selected instruments. A negative and statistically significant spatial correlation emerges between the shift-share instrument and NTL intensity (Table 2, column 8). Our interpretation of the first-stage relationship follows Gehring et al. (2022). In accordance with the negative α coefficient in Equation (3) we estimate, they suggest that districts which on average receive less development assistance from the World Bank tend to benefit more from the capacity of international development banks to spend resources in excess of donors' contributions. The Kleibergen-Paap F-statistic is above the rule-of-thumb critical value of 10. Yielding a coefficient of the IV that is similar to the one of the first stage regression, the reduced form regression (Table 2, column 7) suggests that the impact of the proposed shift-share instrument on emigration rates operates *only* through its relationship with economic conditions, which in turn influence emigration decisions. F-tests and reduced form regressions equally confirm the validity and strength of the two alternative IVs (Tables B7 and B8). For the instrument based on future aid commitments in the

[10] In a robustness check, we apply the same IV strategy as in the baseline estimation, but based on a two-step approach along the lines of Eaton and Kortum (2002). As shown in Table B6, the results remain qualitatively unchanged but point to a somewhat lower magnitude of the effect of NTL intensity on migration which is, however, still above the magnitude obtained in the baseline regression.

economic and production sector the Hansen-J test additionally confirms the validity of the over-identifying restrictions.

To explore further whether there are non-linearities in the income-migration relationship as indicated by Figure 3 above, we split the sample according to the yearly level of per-capita nightlights at the origin.[11] In this disaggregated analysis, we have to rely on standard PPML estimates because there isn't enough variation in the aid projects allocated to non-urbanized districts, making our IV unsuitable for the smaller sub-samples. The results reported in Table 3 reveal that the positive effect of NTL on internal emigration found in the baseline specification is driven by emigrants from comparably richer and urbanized districts (column 2), while for rural districts the coefficient is negative but insignificant (column 1). This finding is corroborated in a monadic regression along the lines of equation (2) with migration data taken from the Malawian Living Standards Measurement Survey (LSMS) (Table B10). When focusing on emigrants leaving poorer areas, the per-capita NTL coefficient switches sign and becomes negative and statistically significant, supporting the proposition that falling incomes due to negative shocks force people to leave their district of origin.[12] The LSMS 2004/2005 suggests that in our period under consideration emigration was indeed a non-negligible response to shocks, most notably among non-agricultural households in rural areas (see Table A4).

Table 4 shows that the significantly positive link between NTL intensity and emigration from urban areas, which points to liquidity constraints that are mitigated by higher incomes, only holds for migrants with lower education levels, i.e. predominantly those who move from urban to non-urban districts. Again, using LSMS migration data leaves the main conclusion unaffected (Table B10). In rural areas, emigration falls with rising incomes for migrants who have at least completed primary schooling, i.e. those who are more likely than not to move to urban destinations.[13] The insignificant effect for those with less than primary education might reflect that this group includes many of the very poorest people who do not have the means to leave their home in response to negative income shocks. Our results are in line with a study for a neighboring low-income country,

[11]If we alternatively employ a quadratic specification to test for non-linearities, both the linear and the squared NTL term turn out to be positively and significantly associated with emigration (Table B9), suggesting that the positive impact grows with rising NTL intensity.
[12]The proposition that negative shocks, in particular related to climate change, lead to more migration is dominant in the literature even though the empirical evidence on the relationship is mixed so far (e.g. Bertoli et al. 2022).
[13] With LSMS migration data, the coefficient remains negative but insignificant also for higher education levels (Table B10).

Tanzania, which shows that weather shocks in rural areas increase the probability of migration only for households in the middle of the wealth distribution (Kubik and Maurel 2016).

5. Concluding Remarks

In this paper, we have analyzed the link between income growth and internal migration for Malawi. Our empirical results suggest that, on average, rising incomes – proxied by changes in nightlight intensity – are associated with higher urban emigration rates, whereas in rural areas emigration is falling with increasing economic activity except for the least educated.

These findings are in accordance with Kleemans' (2023) theoretical model where migration can either be used as an ex-post risk-coping strategy in response to negative income shocks – this might on average apply to the poorer rural districts in Malawi – or serve as an investment, with positive income shocks enabling households to pay up-front migration costs – this might on average apply to urban districts in Malawi. They deviate from Kleemans (2023) in that her estimates indicate temporary migration after negative shocks while our estimates for rural Malawian districts point to permanent emigration when incomes fall. This discrepancy may in part be due to the fact that we do not capture within-district movements which are more likely to be temporary than movements across districts.

Our results also suggest a specific sorting pattern for internal migrants based on differences in skill levels that is in line with previous findings by Young (2013) for a broad cross-section of African countries. While in urban areas rising incomes mainly facilitate the emigration of lower-skilled people to non-urban destinations, in rural areas it is higher-skilled people who are most likely to leave their home in response to falling incomes, often moving to urban destinations.

From a policy perspective, the paper's findings suggest that poor Malawians who reside predominantly in rural areas would need targeted support to cope with recurrent income shocks. This could either involve assistance in districts of origin – such as the provision of draught-resistant seeds or temporary cash transfers – that obviates the need to emigrate, or it could facilitate emigration and integration at destination, for example through providing access to land for housing. People who intend to emigrate from urban areas, but cannot realize their plans due to liquidity constraints, would also benefit from policies that reduce mobility costs.

References

Acosta, K., & Gu, H. (2022). Locked up? The development and internal migration nexus in Colombia. Regional and Urban Economics Working Paper 304, Columbian Central Bank, January 2022.

Amare, M., Arndt, C., Abay, K.M., & Benson, T. (2020). Urbanization and child nutrition outcomes. The World Bank Economic Review 34(1), 63-74.

Asher, S., Lunt, T., Matsuura, R., & Novosad, P. (2021). Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in India using the shrug open data platform. The World Bank Economic Review, 35(4), 845-871.

Baskaran, T., Bhalotra, S., Min, B., & Uppal, Y. (2024). Women legislators and economic performance. Journal of Economic Growth, 29(2), 151-214.

Bazzi, S. (2017). Wealth heterogeneity and the income elasticity of migration. American Economic Journal: Applied Economics, 9(2), 219–255.

Becerra-Valbuena, L. G., & Millock, K. (2021). Gendered migration responses to drought in Malawi. Journal of Demographic Economics, 87(3), 437-477.

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

Beine, M., & Parsons, C. (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.

Benček, D., & Schneiderheinze, D. (2024). Higher economic growth in poor countries, lower migration flows to the OECD: Revisiting the migration hump with panel data. World Development, 182, 106655.

Bertoli, S., Docquier, F., Rapoport, H., & Ruysen, I. (2022). Weather shocks and migration intentions in Western Africa: insights from a multilevel analysis. Journal of Economic Geography, 22(2), 289-323.

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

Bomprezzi, P., Dreher, A., Fuchs, A., Hailer, T., Kammerlander, A., Kaplan, L., Marchesi, S., Masi, T., Robert, C., & Unfried, K. (2025). Wedded to Prosperity? Informal Influence and Regional Favoritism. CEPR Discussion Paper 18878.

Briggs, R. C. (2017). Does foreign aid target the poorest? International Organization, 71(1), 187-206.

Briggs, R. C. (2021). Why Does Aid Not Target the Poorest? International Studies Quarterly 65(3), 739–752.

Bryan, G., & Morten, M. (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. Journal of Political Economy, 127(5), 2229-2268.

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

Clemens, M. A. (2014). Does development reduce migration? In International Handbook on Migration and Economic Development (Vol. 8592, pp. 152–185). Edward Elgar Publishing.

Clemens, M.A. (2020). The emigration life cycle: How development shapes emigration from poor countries. Center for Global Development Working Paper 540, August 2020.

Clemens, M.A., & Mendola, M. (2024). Migration from developing countries: Selection, income elasticity, and Simpson's paradox. Journal of Development Economics, 171, 103359.

Dao, T. H., Docquier, F., Parsons, C., & Peri, G. (2018). Migration and development: Dissecting the anatomy of the mobility transition. Journal of Development Economics, 132, 88–101.

Djajic, S., Kirdar, M. G., & Vinogradova, A. (2016). Source country earnings and emigration. Journal of International Economics, 99, 46–67.

Dreher, A., Fuchs, A., Hodler, R., Parks, B. C., Raschky, P. A., & Tierney, M. J. (2019). African leaders and the geography of China's foreign assistance. Journal of Development Economics, 140, 44-71.

Dreher, A., Fuchs, A., Parks, B., Strange, A., & 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.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. Econometrica, 70(5), 1741-1779.

Gehring, K., Kaplan, L. C., & Wong, M. H. (2022). China and the World Bank—How contrasting development approaches affect the stability of African states. Journal of Development Economics, 158, 102902.

Gibson, J. (2021). Better night lights data, for longer. Oxford Bulletin of Economics and Statistics, 83(3), 770-791.

Gibson, J., Olivia, S., Boe-Gibson, G., & Li, C. (2021). Which night lights data should we use in economics, and where? Journal of Development Economics, 149, 102602.

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. American Economic Review, 60 (1), 126–142.

Henderson, J.V, Storeygard, A., & Weil, D.N. (2012). Measuring economic growth from outer space. American Economic Review, 102(2), 994-1028.

Hirvonen, K. (2016). Temperature changes, household consumption, and internal migration: Evidence from Tanzania. American Journal of Agricultural Economics, 98(4), 1230-1249.

Hodler, R., & Raschky, P. A. (2014). Regional favoritism. The Quarterly Journal of Economics, 129(2), 995-1033.

Kubik, Z., & Maurel, M. (2016). Weather Shocks, Agricultural Production and Migration: Evidence from Tanzania. Journal of Development Studies, 52(5), 665-680.

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., Sanfilippo, M., & Santi, F. (2023). Aid and internal migration in Malawi. World Development, 162, 106-134.

Lanati, M., & Thiele, R. (2024). The link between economic growth and emigration from developing countries: Does migrants' skill composition matter? Kyklos, 77(1), 213-248.

Li, X., Zhou, Y., Zhao, M. *et al.* (2020). A harmonized global nighttime light dataset 1992–2018. Nature Scientific Data, **7**, 168.

Lucas, R.E.B. (2016). Internal migration in developing countries: An overview of recent evidence: Geopolitics, History, and International Relations 8(2), 159-191.

Kleemans, M. (2023). Migration choice under risk and liquidity constraints. Working Paper, University of Illinois, Urbana-Champaign.

McBride, K., & Moucheraud, C. (2022). Rural–Urban Differences: Using Finer Geographic Classifications to Reevaluate Distance and Choice of Health Services in Malawi. *Health Systems & Reform*, *8*(1), e2051229.

Morten, M. (2015). Migration and Labor Mobility Internal Migration in Sub-Saharan Africa. Paper presented at CSAE annual conference.

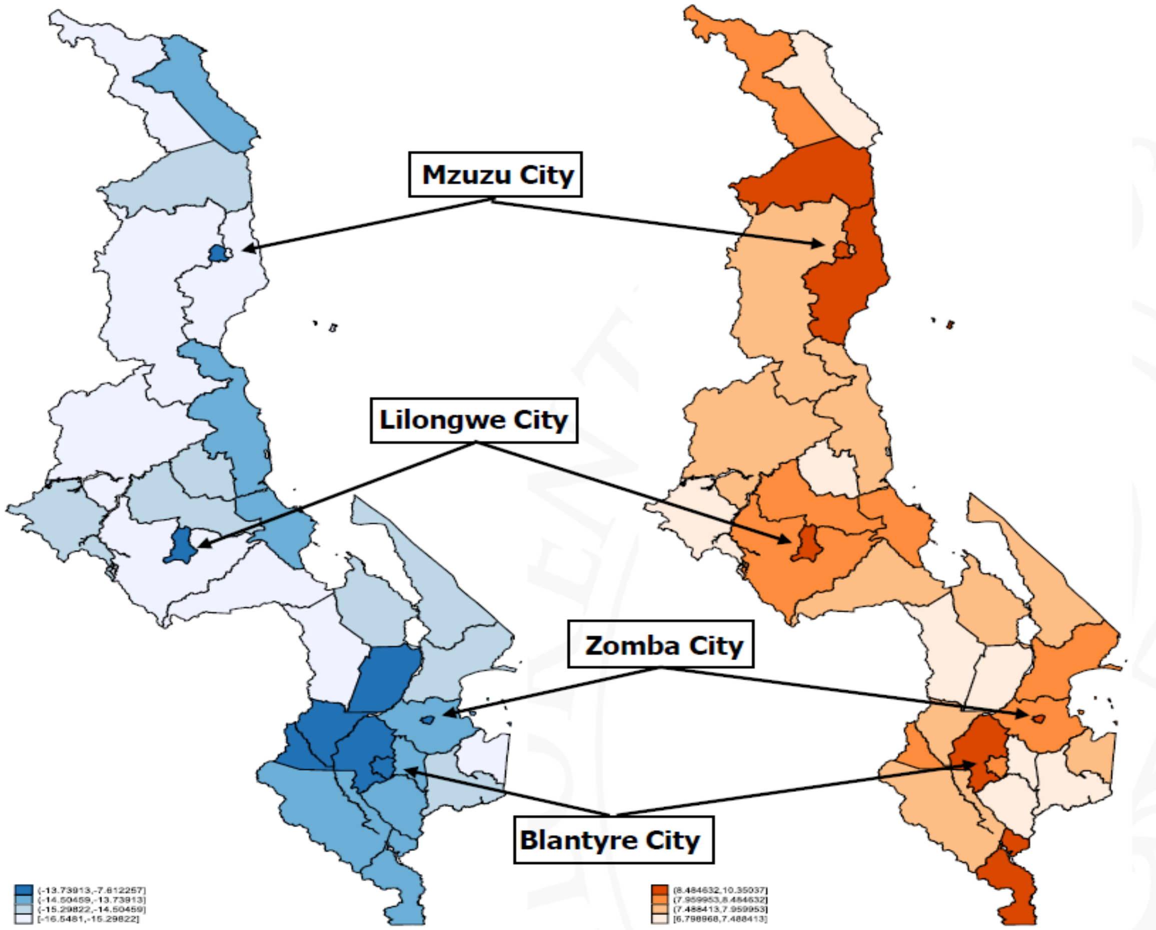
Pischke, S. (2007). Lecture notes on measurement error. London School of Economics, London.

Silva, J. S., & Tenreyro, S. (2006). The log of gravity. The Review of Economics and Statistics, 641-658.

Shon, H., Lee, H., & Kim, B. (2024). Spatial pattern of aid allocation in the early 21st century: Evidence from 38 sub-Saharan African countries. Papers in Regional Science, 103(3), 100026.

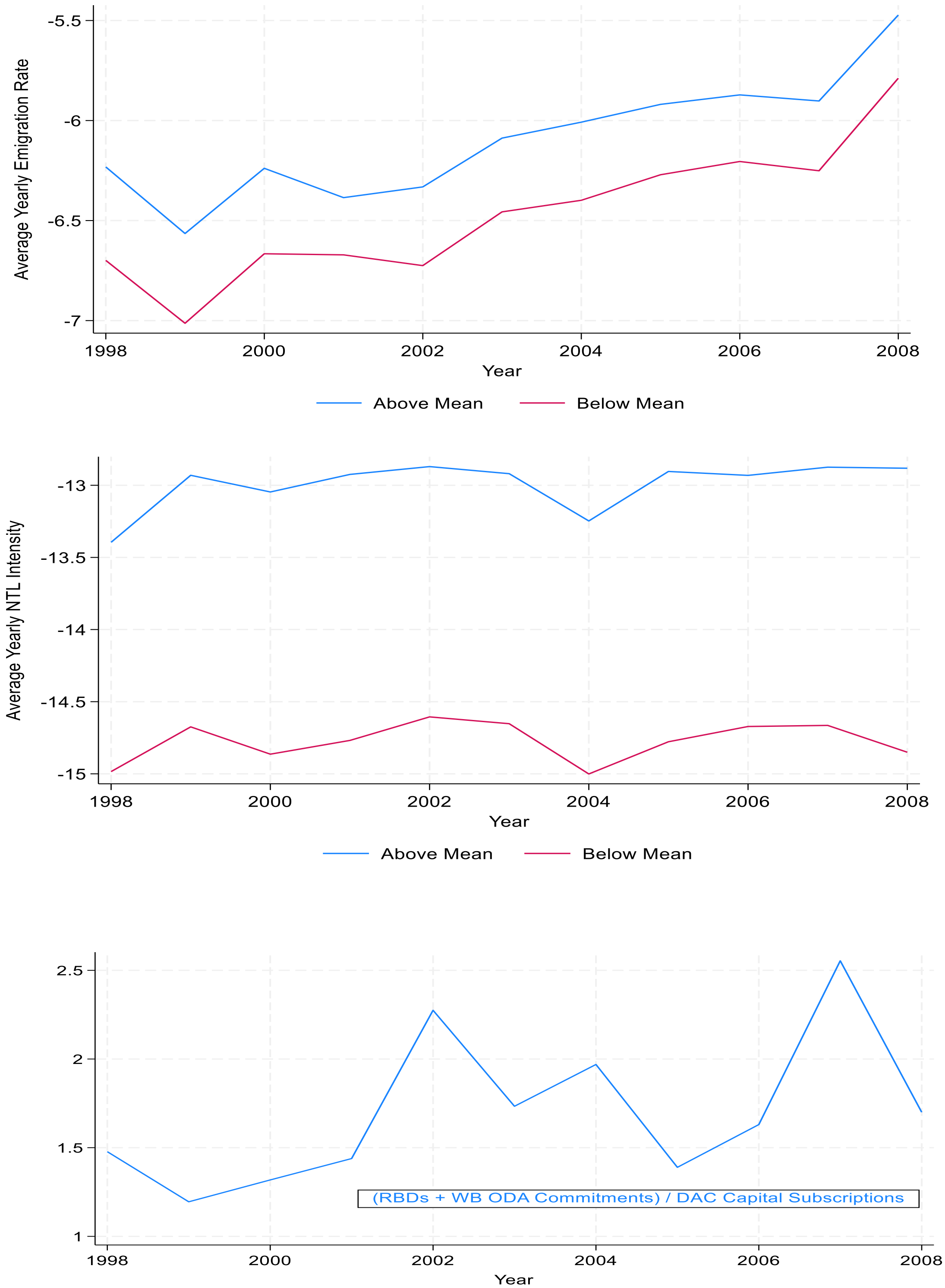
Young, A. (2013). Inequality, the Urban-Rural Gap, and Migration. The Quarterly Journal of Economics, 128 (4), 1727–85.

Figure 1: Average per Capita Nightlights vs Average Per Capita Aid Disbursements



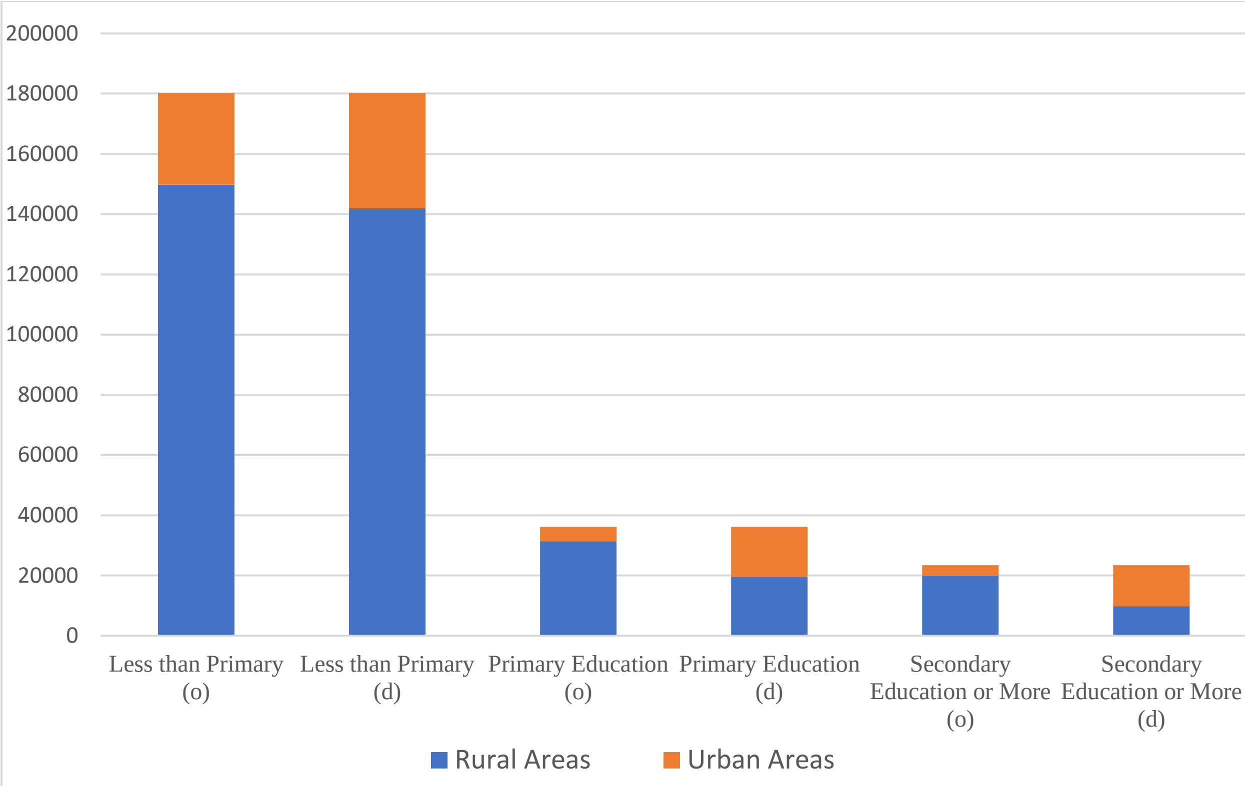
Notes: The Figure reports the values of average per-capita nightlights (left) vis-à-vis the average per capita aid disbursements (right). Values are expressed in natural logarithms. The four cities identified in the maps were the four major urban areas in Malawi prior to 2009. For the construction of the averages, we use observations from the baseline sample.

Figure 2: Parallel-trends assumption



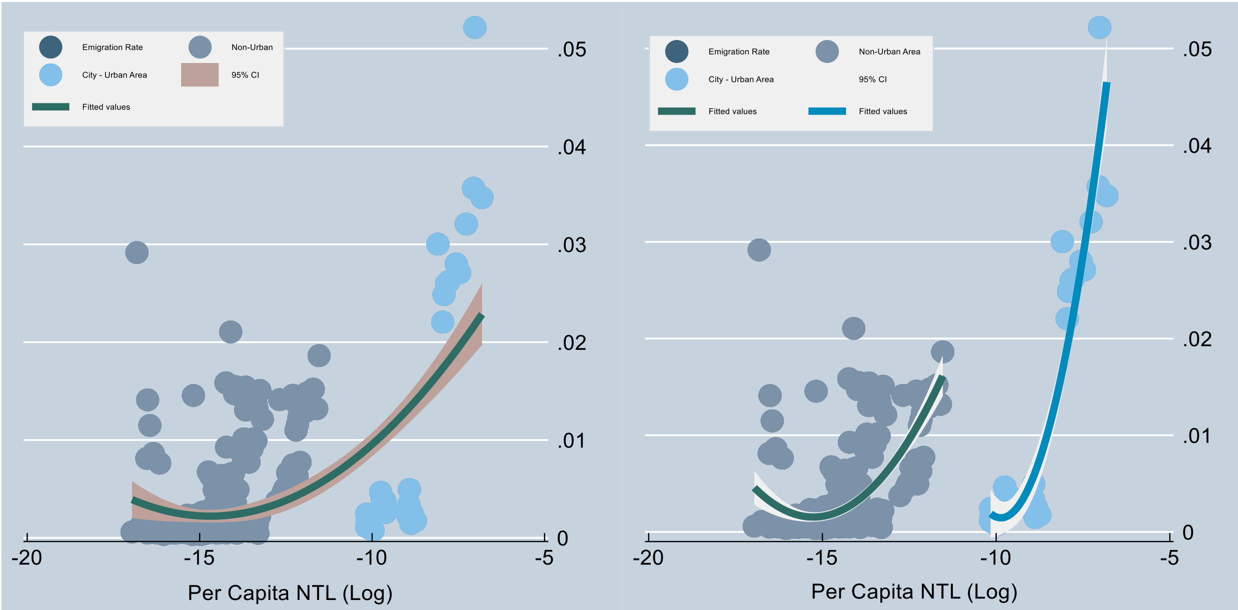
Notes: The Upper Panel shows the average (log) emigration rate over time within the group that is below the median of the average volume of aid received and the group that is above the median. The Middle Panel shows WB's yearly average aid disbursements (log, t-1) in Malawi over time within these two groups. The Bottom Panel reports the share of WB disbursements over DAC's Capital Subscriptions over time. For the construction of the averages, we use observations from the baseline sample.

Figure 3: Migrants by Skill Level and Area of Origin (o) & Destination (d) *(Rural vs Urban)*



Notes: The graph shows the total number of internal emigrants disaggregated by skill level – *Less than primary, Primary and Secondary or more* – and by district of origin and destination (rural vs urban). Urban districts are Lilongwe City, Blantyre city, Mzuzu city and Zomba city.

Figure 4: Correlation between District-level Nightlight Intensity and Internal Migration



Notes: The left-hand graph shows the quadratic trend from the relationship between yearly emigration rate and per capita nightlights at district level. The right-hand graph shows the same relationship by dividing between urban and non-urban districts. Urban districts are: Lilongwe City, Blantyre city, Mzuzu city and Zomba city

Table 1 – Gravity Model Estimates

	(1)	(2)	(3)	(4)	(5)
Estimator	PPML	PPML	PPML	PPML	PPML
Depvar	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$
Log Mig Stock $_{i,t}$	0.894^{***}	0.893^{***}	0.943^{***}	0.911^{***}	1.761^{***}
	(29.94)	(30.46)	(41.48)	(182.11)	(17.83)
Log NTL pc $_{i,t-1}$	**0.248^{***}**	**0.247^{***}**	**0.221^{***}**	**0.170^{***}**	**0.177^{***}**
	(7.34)	**(7.50)**	**(11.10)**	**(5.71)**	**(5.61)**
SPEI $_{i,t}$	-0.125^{***}	-0.275^{***}	-0.394^{***}	-0.0186	-0.0611
	(-3.31)	(-2.94)	(-5.14)	(-0.47)	(-1.52)
Conflict $_{i,t}$	0.167	0.201	0.152^{**}	0.0134	-0.00211
	(1.33)	(1.33)	(2.53)	(0.51)	(-0.08)
Pop Density $_{i,t}$	-0.00384^{***}	-0.00387^{***}	-0.00368^{***}	0.00168^{*}	0.00133
	(-5.63)	(-5.76)	(-11.84)	(1.84)	(1.30)
N	10080	10080	10080	10080	9905
Year Fes		X			
Dest-Year Fes			X	X	X
Origin FEs				X	
Cty-Pair FEs					X
% Zeros	23.6%	23.6%	23.6%	23.6%	22%

t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by district of origin.
The Table reports the PPML estimates of Equation (1) with different sets of fixed effects.

Table 2: Monadic Analysis

Estimator Model	(1) PPML Monadic	(2) PPML Monadic	(3) IV-PPML Monadic	(4) IV-PPML Monadic	(5) IV-2SLS Monadic	(6) IV-2SLS Monadic	**(7)** PPML Monadic	**(8)** IV-PPML Monadic
Depvar	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$ln\left(\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}\right)$	$ln\left(\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}\right)$		
							Reduced Form	First Stage
Log NTL pc $_{i,t-1}$	**0.306***** **(3.54)**	**0.292**** **(2.08)**	**0.300***** **(3.67)**	**0.692***** **(2.97)**	**0.313***** **(2.88)**	**0.661**** **(2.71)**		
SPEI $_{i,t}$	-0.370*** (-7.70)	0.0326 (0.44)	-0.557*** (-7.22)	-0.0513 (-0.58)	-0.403*** (-7.47)	-0.0349 (-0.40)	0.236** (2.49)	0.138* (1.78)
Conflict $_{i,t}$	0.592** (2.30)	0.0419 (1.04)	0.875*** (3.59)	0.00456 (0.08)	0.693*** (2.62)	0.0161 (0.29)	0.0768 (1.57)	0.0612 (1.01)
Pop Density $_{i,t}$	-0.00215* (-1.75)	0.00370** (2.42)	-0.00180* (-1.84)	0.000830 (0.79)	-0.00171 (-1.58)	0.000702 (0.65)	-0.00141 (-1.03)	0.00443** (2.00)
Instrumental Variable								
IV $_{i,t}$							-0.000018*** (-3.57)	-0.000016*** (-13.38)
N	336	336	336	336	336	336	336	336
KP-F Stat					138.108	11.883		
Origin Fes		X		X		X	X	X
Year Fes		X		X		X	X	X

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by district of origin.
The table reports the PPML and IV results of model 2 (columns 1-6). Columns (7-8) show the reduced form and first stage statistics of the IV-PPML model. IV-PPML statistics are obtained with the control function approach.

Table 3: Disaggregation across Area and Per-Capita NTL Quintiles

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimator	PPML	PPML	PPML	PPML	PPML	PPML	PPML
Subsample	Rural	Urban	NTL pc	NTL pc	NTL pc	NTL pc	NTL pc
			<=20th	<=40th	<=60th	<=80th	All
Log NTL pc $_{i,t-1}$	-0.0217	0.523***	-0.129***	-0.0681**	-0.0655**	-0.0342	0.170**
	(-0.80)	(3.52)	(-3.83)	(-2.03)	(-2.48)	(-1.45)	(1.99)
N	8760	1320	2178	4170	6150	8130	10080
Origin FEs	X	X	X	X	X	X	X
Dest-Year FEs	X	X	X	X	X	X	X

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the results of model (1) estimated on urban vs non-urban districts of origin and on subsamples based on per capita NTL quintiles.

Table 4: Disaggregation across Area and Skill Levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Estimator	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML
District of Origin	Rural	Urban	Rural	Rural	Rural	Rural	Rural	Urban	Urban	Urban	Urban	Urban
Skill	All	All	Low	L-M	M-H	High	(3+4)	Low	L-M	M-H	High	(3+4)
Log NTL pc $_{i,t-1}$	-0.0217	0.523***	0.0196	-0.0391	-0.0577	-0.273***	-0.0691**	0.692***	0.523***	-0.198	0.0908	-0.155
	(-0.80)	(3.52)	(0.57)	(-1.18)	(-1.61)	(-2.66)	(-2.03)	(13.83)	(5.06)	(-0.57)	(0.40)	(-0.54)
N	8760	1320	8760	8760	8760	5124	8760	1320	1289	1208	397	1212
Dest-Year Fes	X	X	X	X	X	X	X	X	X	X	X	X
Origin FEs	X	X	X	X	X	X	X	X	X	X	X	X

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors are clustered by district of origin.
Low=less than primary; L-M=primary; M-H=secondary; High=university; (3+4) = secondary + university. Urban districts are: Lilongwe City, Blantyre city, Mzuzu city and Zomba city. Coefficients of the control variables are not reported. The table reports the results of model (1).

Appendix

Appendix A includes descriptive information complementing the evidence shown in the main text. Figure A1 shows the correlation between NTL intensity and emigration rates for each yearly cross section, Figure A2 the same correlation across different skill levels. Table A1 reports the number of migrants by area (rural vs urban) and skill level. Tables A2 and A3 define the variables employed in the empirical analysis and report the main summary statistics for each district, respectively. Table A4

Figure A1 – Yearly Correlation between District-level Nightlight Intensity and Internal Migration

Figure A2 - Correlation between District-level NTL Intensity and Internal Migration across Skill Levels

Table A1 – Internal Migration / Descriptive Statistics by Skill Level

Table A2 – Description of the Main Variables

Table A3 – Summary statistics by origin (district of previous residence)

Table A4 – Migration as a Response to Shocks, LSMS 04-05

Appendix B reports all the tests in support of the robustness of the results presented in the main analysis. This includes a comparison between the baseline specification and the same model which aggregates all migrants for which skill levels are reported (Table B1), regressions using different transformations of nightlight data (Tables B2 and B3), estimates with alternative clustering of standard errors (Table B4), estimates of parsimonious versions of Equations (1) and (2) without controls (Table B5), the IV estimation with a two-step strategy (Table B6), the IV estimation performed with alernative instruments (Table B7) and the respective first-stage results (Table B8), a regression including a squared NTL to test for non-linearities (Table B9), and the disaggregated analysis estimated with information on internal movements extracted from the LSMS Third Integrated Household Survey 2010-2011 (Table B10).

Table B1 – All Skill Levels vs Baseline Estimates

Table B2 – Scaled-Log Transformation of NTL (1+NTL)

Table B3 – NTL in Inverse Hyperbolic Sine (IHS) Transformation

Table B4 – Alternative Clustering of Standard Errors

Table B5 – Estimates without Additional Controls

Table B6 – IV Estimates using Two-Step Strategy

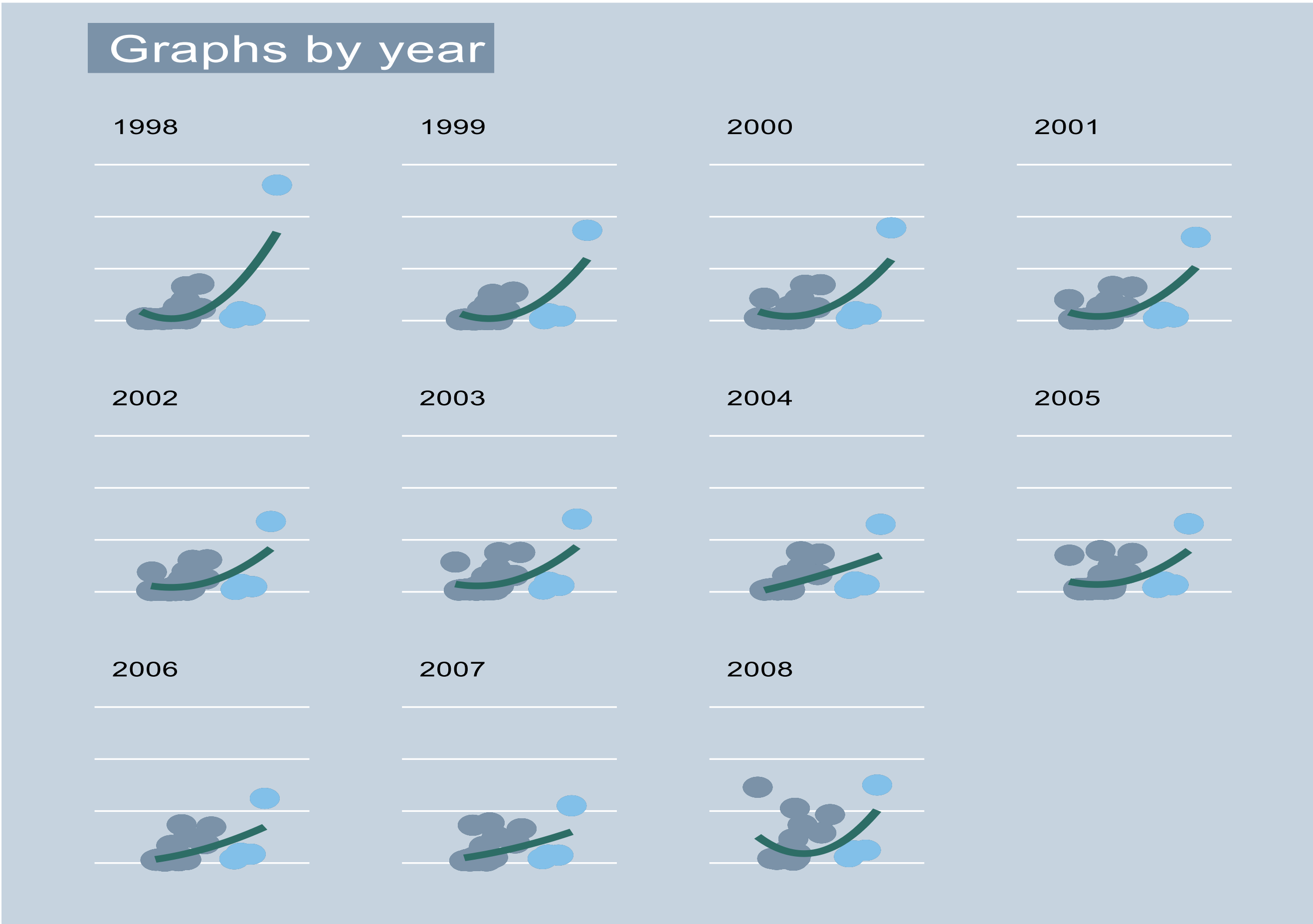
Table B7 – Monadic Estimation with Alternative Instruments

Table B8 – Alternative Instruments, First-Stage Results

Table B9 – Including a Squared Term

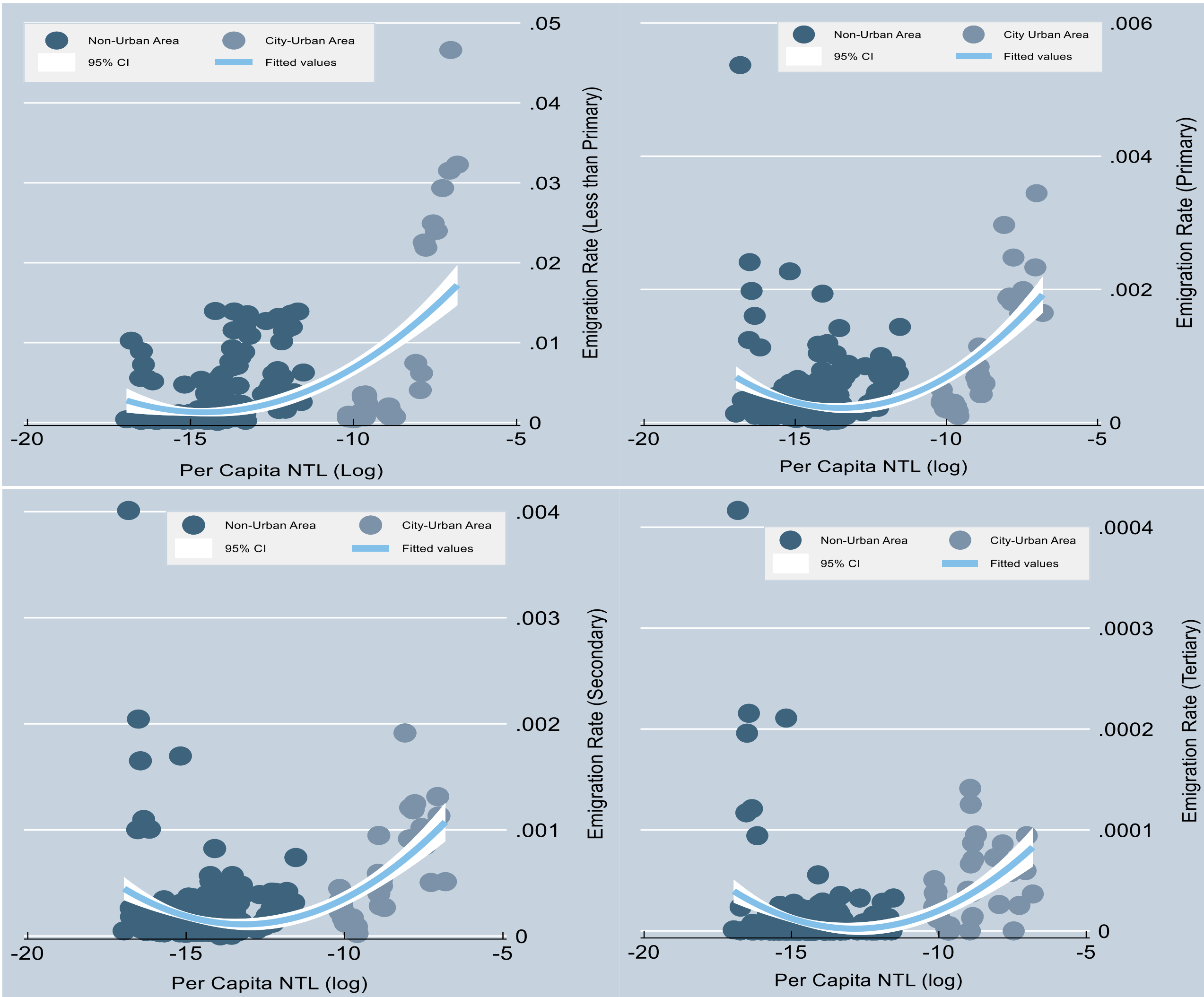
Table B10 – Monadic Estimates with LSMS Migration Data

Figure A1: Yearly Correlation between District-level Nightlight Intensity and Internal Migration



Notes: The graphs show the quadratic trend from the relationship between yearly emigration rate and per capita nightlights (log) at district level. The figures divide between urban (blue) and non-urban (grey) districts. Urban districts are: Lilongwe City, Blantyre city, Mzuzu city and Zomba city

Figure A2: Correlation between District-level NTL Intensity and Internal Migration across Skill Levels



Notes: The graphs show the quadratic trend from the relationship between yearly emigration rate and per capita nightlights at district level across skill levels: less than primary (upper left), primary (upper right), secondary (bottom left) and university (bottom right). Emigration rates are calculated using total population at district level as denominator. Urban districts are: Lilongwe City, Blantyre city, Mzuzu city and Zomba city.

Table A1 – Internal Migration by Skill Level

Internal Emigration	(1) Less than primary	(2) Primary	(3) Secondary	(4) Tertiary	(5) Secondary or more (3) + (4)
From Rural Districts					
Rural to urban	36237	15323	11313	908	12221
From rural to rural	113588	16089	7469	354	7823
Total from rural	149825	31412	18782	1262	20044
Share rural to urban	0.24	0.49	0.60	0.72	0.61
Share rural to rural	0.76	0.51	0.40	0.28	0.39
From Urban Districts					
Urban to urban	2177	1251	1225	173	1398
Urban to rural	28295	3507	1913	99	2012
Total from urban	30472	4758	3138	272	3410
Share urban to urban	0.07	0.26	0.39	0.64	0.41
Share urban to rural	0.93	0.74	0.61	0.36	0.59

Table A2 – Description of the Main Variables

Domain and Source	Variable Name	Description
Official Development Assistance Data **Source:** GODAD Project based on the OECD's Creditor Reporting System (CRS) (Used to build the **IVs**)	***Econ Aid Projects*** $_{i,t+1\text{-}t+2}$	*World Bank per-capita committed Eco sector total flows (in levels, constant 2014 US$) in the period t+1, t+2*
	Prod Aid Projects $_{i,t+1\text{-}t+2}$	*World Bank per-capita committed Eco sector total flows (in levels, constant 2014 US$) in the period t+1, t+2*
	IV1 $_{i,t}$	***Shift-Share Instrument Constructed as follows:*** *Regional Development Banks and World Bank's total commitments to All recipients but Malawi in year t divided by the total volume of DAC Donors' Capital Subscriptions. This term is multiplied by the Per-Capita Average World Bank's Volume of Aid Disbursed to Malawian District i.over the period 1998-2008.*
	IV2 $_{i,t}$	***Shift-Share Instrument Constructed as follows:*** *World Bank's total commitments to All recipients but Malawi in year t divided by the total volume of DAC Donors' Capital Subscriptions. This term is multiplied by Malawian District i's probability of receiving aid over the period 1998-2008.*
Internal Migration Data **Source:** Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D020.V7.2	***Migrant Stocks / Network*** $_{ij,t-1}$	*Stock of Migrants born in district **i** and living in district **j** as in year **t-1** (in logs).*
	Migrant Flow $_{ij,t}$	*Total number of people that moved from district **i** to district **j** at time **t** (Dependent Variable)*
	International Migrant Flow $_{i,t}$	*Total number of people that moved from district **i** to a foreign country at time **t** (Dependent Variable)*
	Migrant Flow (Rural) $_{i,t}$	*Total number of people that moved from district **i** to district **j** at time **t** from predominantly rural districts other than Lilongwe City, Blantyre city, Mzuzu city and Zomba city*
	Migrant Flow (Urban) $_{i,t}$	*Total number of people that moved from district **i** to district **j** at time **t** from urban districts, namely Lilongwe City, Blantyre city, Mzuzu city and Zomba city*
	Migrant Flow (Less than Primary) $_{j,t}$	*Total number of people that moved from district **i** to district **j** at time **t** who did not complete primary school*
	Migrant Flow (Primary) $_{j,t}$	*Total number of people that moved from district **i** to district **j** at time **t** who completed primary school*
	Migrant Flow (Secondary) $_{j,t}$	*Total number of people that moved from district **i** to district **j** at time **t** who completed secondary school*
	Migrant Flow (Tertiary) $_{j,t}$	*Total number of people that moved from district **i** to district **j** at time **t** who hold a university degree*
Additional controls: **Sources:** NOAA-DMSP Harari and La Ferrara (RESTAT 2018)	***Conflict*** $_{j,t}$	*Presence of any form of conflict in the district of origin **i** (dummy) at time **t***
	SPEI $_{j,t}$	*Crop affecting environmental variable in the district of origin **i** at time **t***
	Population Density $_{i,t}$	*Population Density in the district of origin **i** at time **t***

Notes: Subscripts – "*i"* indicates the district of origin; "*j*" refers to the district of destination (when referring to internal migration); t refers to time (year). 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.

Table A3: Summary Statistics by Origin (District of Previous Residence)

	Mean	Max	Min	SD
Balaka				
Emigration Rate $_{i,t-1}$	.009	.015	0.007	.002
Log NTL pc $_{i,t-1}$	-13.607	-13.37	-14.018	.181
Log Average WB Aid Disb (US$)	19.76196			
Blantyre				
Emigration Rate $_{i,t-1}$	.006	.012	0.004	.002
Log NTL pc $_{i,t-1}$	-12.367	-12.088	-12.766	.203
Log Average WB Aid Disb (US$)	21.53441			
Blantyre city				
Emigration Rate $_{i,t-1}$	.003	.005	0.003	.001
Log NTL pc $_{i,t-1}$	-9.663	-9.604	-9.752	.043
Log Average WB Aid Disb (US$)	21.42621			
Chikwawa				
Emigration Rate $_{i,t-1}$	.003	.005	0.002	.001
Log NTL pc $_{i,t-1}$	-14.324	-14.159	-14.576	.138
Log Average WB Aid Disb (US$)	20.45442			
Chiradzulu				
Emigration Rate $_{i,t-1}$	.002	.004	0.001	.001
Log NTL pc $_{i,t-1}$	-13.948	-13.266	-14.738	.482
Log Average WB Aid Disb (US$)	19.63637			
Chitipa				
Emigration Rate $_{i,t-1}$	.001	.003	0.001	.001
Log NTL pc $_{i,t-1}$	-15.571	-14.781	-16.736	.713
Log Average WB Aid Disb (US$)	20.36343			
Dedza				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-15.321	-14.796	-15.704	.269
Log Average WB Aid Disb (US$)	20.69592			
Dowa				
Emigration Rate $_{i,t-1}$	.001	.002	0.001	0
Log NTL pc $_{i,t-1}$	-14.96	-14.381	-15.773	.468
Log Average WB Aid Disb (US$)	21.29887			
Karonga				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-14.145	-13.311	-14.903	.483
Log Average WB Aid Disb (US$)	19.79247			
Kasunga				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-15.822	-15.395	-16.438	.282
Log Average WB Aid Disb (US$)	20.96893			
Lilongwe				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-15.817	-15.386	-16.961	.425
Log Average WB Aid Disb (US$)	22.33529			
Lilongwe City				
Emigration Rate $_{i,t-1}$	.001	.002	0.001	.001
Log NTL pc $_{i,t-1}$	-10.079	-9.974	-10.156	.063
Log Average WB Aid Disb (US$)	22.41962			
Machinga				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0

Log NTL pc $_{i,t-1}$	-15.048	-14.647	-15.452	.272
Log Average WB Aid Disb (US$)	21.14565			
Mangochi				
Emigration Rate $_{i,t-1}$	.001	.001	0.000	0
Log NTL pc $_{i,t-1}$	-15.177	-14.88	-15.423	.205
Log Average WB Aid Disb (US$)	21.32816			
Mchinji				
Emigration Rate $_{i,t-1}$	.001	.001	0.000	0
Log NTL pc $_{i,t-1}$	-15.368	-14.859	-16.325	.544
Log Average WB Aid Disb (US$)	19.63637			
Mulanje				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-14.551	-14.154	-15.000	.272
Log Average WB Aid Disb (US$)	20.31819			
Mwanza				
Emigration Rate $_{i,t-1}$	.014	.019	0.011	.002
Log NTL pc $_{i,t-1}$	-12	-11.537	-12.682	.341
Log Average WB Aid Disb (US$)	19.64962			
Mzimba				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-16.081	-15.717	-16.516	.265
Log Average WB Aid Disb (US$)	21.04587			
Mzuzu City				
Emigration Rate $_{i,t-1}$	.003	.005	0.001	.001
Log NTL pc $_{i,t-1}$	-8.887	-8.741	-9.012	.072
Log Average WB Aid Disb (US$)	21.35715			
Neno				
Emigration Rate $_{i,t-1}$	.015	.021	0.010	.003
Log NTL pc $_{i,t-1}$	-13.678	-13.183	-14.238	.36
Log Average WB Aid Disb (US$)	19.18127			
Nkhata Bay and Likoma				
Emigration Rate $_{i,t-1}$	.001	.002	0.001	.001
Log NTL pc $_{i,t-1}$	-15.377	-14.954	-16.353	.404
Log Average WB Aid Disb (US$)	20.95288			
Nkhota kota				
Emigration Rate $_{i,t-1}$	.001	.002	0.001	0
Log NTL pc $_{i,t-1}$	-14.514	-14.26	-14.780	.145
Log Average WB Aid Disb (US$)	20.01774			
Nsanje				
Emigration Rate $_{i,t-1}$	.006	.009	0.004	.001
Log NTL pc $_{i,t-1}$	-14.326	-13.892	-14.747	.251
Log Average WB Aid Disb (US$)	20.78201			
Ntcheu				
Emigration Rate $_{i,t-1}$	.001	.002	0.001	0
Log NTL pc $_{i,t-1}$	-15.475	-15.269	-15.822	.183
Log Average WB Aid Disb (US$)	20.12159			
Nthisi				
Emigration Rate $_{i,t-1}$	.001	.002	0.001	0
Log NTL pc $_{i,t-1}$	-14.652	-13.568	-15.355	.601
Log Average WB Aid Disb (US$)	19.23517			
Phalombe				
Emigration Rate $_{i,t-1}$	.012	.029	0.006	.007
Log NTL pc $_{i,t-1}$	-16.283	-15.18	-16.818	.526
Log Average WB Aid Disb (US$)	20.3463			

Rumphi				
Emigration Rate $_{i,t-1}$	.002	.004	0.001	.001
Log NTL pc $_{i,t-1}$	-14.831	-14.066	-15.350	.443
Log Average WB Aid Disb (US$)	21.07434			
Salima				
Emigration Rate $_{i,t-1}$	.001	.002	0.000	0
Log NTL pc $_{i,t-1}$	-14.161	-13.881	-14.803	.283
Log Average WB Aid Disb (US$)	20.76648			
Thyolo				
Emigration Rate $_{i,t-1}$	.001	.001	0.000	0
Log NTL pc $_{i,t-1}$	-13.937	-13.565	-14.457	.295
Log Average WB Aid Disb (US$)	20.42923			
Zomba				
Emigration Rate $_{i,t-1}$	.001	.003	0.001	.001
Log NTL pc $_{i,t-1}$	-14.287	-13.884	-14.896	.317
Log Average WB Aid Disb (US$)	21.24459			
Zomba City				
Emigration Rate $_{i,t-1}$	.031	.052	0.022	.008
Log NTL pc $_{i,t-1}$	-7.524	-6.815	-8.093	.432
Log Average WB Aid Disb (US$)	21.093			

Notes: summary statistics at the level of district of origin. In **bold** the urban areas

Table A4: Migration as a Response to Shocks, LSMS 04-05

Household Head	(1) All	(2) Agricultural	(3) Non-Agricultural
*Districts: **Non-Urban***			
Type of Shock	Response: Emigration (%)	Response: Emigration (%)	Response: Emigration (%)
All Shocks	1.79%	1.67%	4.43%
Income	1.64%	1.96%	3.13%
Asset	1.73%	1.50%	9.86%
Both	2.26%	2.09%	8.04%
*Districts: **Urban***			
Type of Shock	Response: Emigration (%)	Response: Emigration (%)	Response: Emigration (%)
All Shocks	1.25%	0.95%	2.62%
Income	1.27%	0.92%	2.60%
Asset	1,01%	0.70%	3.67%
Both	1.28%	1.08%	2.51%

Notes: Data are from Malawi's LSMS Second Integrated Household Survey 2004-2005 – module "Recent shocks to household welfare". Urban districts are: Lilongwe City, Blantyre city, Mzuzu city and Zomba city. Response "Went elsewhere to find work for more than 1 month" as the main reaction to Shock to Regain Former Welfare Level (1). Shocks have been classified to those that caused a reduction in income, assets or both. The statistics distinguish between agricultural and non-agricultural households.

Table B1: All Skill Levels vs Baseline Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Model	Gravity	Monadic	Monadic	Gravity	Monadic	Monadic
Estimator	PPML	PPML	IV-PPML	PPML	PPML	IV-PPML
Depvar	$M_{ij,t}/M_{ii,t}$	$\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}$	$\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}$	$\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}$
Log Mig Stock $_{ij,t}$	1.761^{***}			2.834^{***}		
	(17.83)			(20.42)		
NTL pc (IHS) $_{i,t-1}$	**0.177^{***}**	**0.292^{**}**	**0.692^{***}**	**0.208^{***}**	**0.403^{**}**	**0.850^{***}**
	(5.61)	**(2.08)**	**(2.97)**	**(5.19)**	**(2.55)**	**(2.90)**
SPEI $_{i,t}$	-0.0611	0.0326	-0.0513	-0.107^{***}	0.0760	-0.0265
	(-1.52)	(0.44)	(-0.58)	(-2.70)	(0.68)	(-0.23)
Conflict $_{i,t}$	-0.00211	0.0419	0.00456	-0.0346	0.0287	0.0910
	(-0.08)	(1.04)	(0.08)	(-1.11)	(0.38)	(1.14)
Pop Density $_{i,t}$	0.00133	0.00370^{**}	0.000830	0.000495	0.00391^{**}	0.000506
	(1.30)	(2.42)	(0.79)	(0.45)	(2.37)	(0.30)
N	9905	336	336	9905	336	336
Year Fes		X	X		X	X
Origin Fes		X	X		X	X
Dest-Year FEs	X			X		
Cty-Pair FEs	X			X		

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table compares the estimates of the baseline model (Tables 1 and 2) with the corresponding results obtained using aggregate flows for which the information on education levels are available.

Table B2: Scaled-Log Transformation of NTL (1+NTL)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model	Gravity	Gravity	Gravity	Gravity	Gravity	Monadic	Monadic
Estimator	PPML	PPML	PPML	PPML	PPML	PPML	IV-2SLS
Depvar	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}$	$ln\left(\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}\right)$
Log Mig Stock $_{ij,t}$	0.900***	0.900***	0.936***	0.910***	1.763***		
	(45.36)	(47.61)	(65.52)	(149.08)	(17.33)		
Log (NTL+1) pc $_{i,t-1}$	0.374***	0.372***	0.368***	0.835***	0.830***	0.902***	1.163***
	(10.43)	(10.42)	(13.48)	(23.05)	(30.74)	(5.99)	(5.71)
SPEI $_{i,t}$	-0.153***	-0.211**	-0.314***	-0.0514	-0.0877	0.0143	0.0970*
	(-5.39)	(-2.03)	(-2.59)	(-0.90)	(-1.40)	(0.20)	(1.78)
Conflict $_{i,t}$	0.212*	0.233	0.177*	-0.0442	-0.0569*	-0.0221	0.0461
	(1.80)	(1.53)	(1.71)	(-1.51)	(-1.65)	(-0.40)	(1.29)
Pop Density $_{i,t}$	-0.00330***	-0.00333***	-0.00321***	0.000256	-0.0000830	0.00263	-0.00112
	(-4.47)	(-4.60)	(-5.90)	(0.18)	(-0.05)	(1.61)	(-1.11)
N	10230	10230	10230	10230	10054	341	341
Year FEs		X				X	X
Origin FEs				X		X	X
Dest-Year FEs			X	X	X		
Cty-Pair FEs					X		

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Table reports the baseline results with NTL expressed in Scaled-Log. Standard errors are clustered by district of origin.

Table B3: NTL in Inverse Hyperbolic Sine (IHS) Transformation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model	Gravity	Gravity	Gravity	Gravity	Gravity	Monadic	Monadic
Estimator	PPML	PPML	PPML	PPML	PPML	PPML	IV-2SLS
Depvar	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$M_{ij,t}/M_{ii,t}$	$\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}$	$ln\left(\sum_{j=1}^{J} M_{ij,t}/M_{ii,t}\right)$
Log Mig Stock $_{ij,t}$	0.903^{***}	0.902^{***}	0.938^{***}	0.910^{***}	1.767^{***}		
	(43.49)	(45.72)	(60.37)	(150.17)	(17.43)		
NTL pc (IHS) $_{i,t-1}$	**0.330^{***}**	**0.328^{***}**	**0.322^{***}**	**0.782^{***}**	**0.781^{***}**	**0.815^{***}**	**1.205^{***}**
	(9.71)	**(9.73)**	**(12.48)**	**(22.35)**	**(21.01)**	**(4.38)**	**(4.69)**
SPEI $_{i,t}$	-0.151^{***}	-0.230^{**}	-0.341^{***}	-0.0501	-0.0868	0.0241	0.0987
	(-4.86)	(-2.17)	(-2.70)	(-0.85)	(-1.35)	(0.32)	(1.69)
Conflict $_{i,t}$	0.218^{*}	0.245	0.184^{*}	-0.0472	-0.0602^{*}	-0.0224	0.0410
	(1.73)	(1.52)	(1.68)	(-1.58)	(-1.75)	(-0.40)	(1.07)
Pop Density $_{i,t}$	-0.00354^{***}	-0.00356^{***}	-0.00341^{***}	0.000458	0.000115	0.00297^{*}	-0.00116
	(-4.50)	(-4.62)	(-5.86)	(0.32)	(0.07)	(1.76)	(-1.13)
N	10230	10230	10230	10230	10054	341	341
Year FEs		X				X	X
Origin FEs				X		X	X
Dest-Year FEs			X	X	X		
Cty-Pair FEs					X		

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Table reports the baseline results with NTL expressed in Inverse Hyperbolic Sine (IHS) transformation. Standard errors are clustered by district of origin.

Table B4: Alternative Clustering of Standard Errors

	(1)	(2)	(3)	(4)	(5)
Model	Gravity	Gravity	Gravity	Monadic	Monadic
Estimator	PPML	PPML	PPML	PPML	IV-PPML
Clustering SE	Origin-Year	District-Pair	VCE Robust	Origin-Year	Origin-Year
Log NTL pc $_{i,t-1}$	**0.170*****	**0.170*****	**0.170*****	**0.292*****	**0.692*****
	(3.40)	**(5.71)**	**(6.62)**	**(4.35)**	**(3.56)**
SPEI $_{i,t}$	-0.0186	-0.0186	-0.0186	0.0326	-0.0513
	(-0.28)	(-0.47)	(-0.41)	(0.34)	(-0.56)
Conflict $_{i,t}$	0.0134	0.0134	0.0134	0.0419	0.00456
	(0.35)	(0.51)	(0.49)	(0.73)	(0.09)
Pop Density $_{i,t}$	0.00168	0.00168*	0.00168**	0.00370***	0.000830
	(1.12)	(1.84)	(2.19)	(2.61)	(1.11)
N	10080	10080	10080	336	336
Year FEs				X	X
Origin FEs	X	X	X	X	X
Dest-Year FEs	X	X	X		

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B5: Estimates without Additional Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimator	PPML	PPML	IV- PPML	PPML	PPML	PPML	PPML
Depvar	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$\sum_{j=1}^{J} \mathrm{M}_{ij,t} / \mathrm{M}_{ii,t}$	$\mathrm{M}_{ij,t}/\mathrm{M}_{ii,t}$	$\mathrm{M}_{ij,t}/\mathrm{M}_{ii,t}$	$\mathrm{M}_{ij,t}/\mathrm{M}_{ii,t}$	$\mathrm{M}_{ij,t}/\mathrm{M}_{ii,t}$
	Monadic	Monadic	Monadic	Gravity	Gravity	Gravity	Gravity
Log NTL pc $_{i,t-1}$	**0.275*****	**0.323****	**0.702*****	**0.275*****	**0.279*****	**0.310****	**0.169***
	(3.54)	**(1.99)**	**(3.18)**	**(3.54)**	**(3.72)**	**(2.02)**	**(1.84)**
IV1 $_{i,t}$			-0.0000179***				
			(-3.01)				
N	336	336	336	10140	10140	10140	9965
Year Fes		X	X				
Dest-Year Fes					X	X	X
Origin FEs		X	X		X	X	X

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by district of origin.
The table reports the PPML results of model 2 and 1 with fixed effects but without additional controls.

Table B6: IV Estimation with Two-Step Strategy

	(1)	(2)	(3)
Estimator	PPML	IV-PPML	IV-PPML
Model	Gravity	Monadic	Monadic
Depvar	$\text{M}_{ij,t}/\text{M}_{ii,t}$	$\hat{S}_{i,t}$	$\hat{S}_{i,t}$
Log Mig Stock $_{ij,t}$	1.880^{***}		
	(22.41)		
Log NTL pc $_{i,t-1}$		0.793^{***}	0.779^{**}
		(2.79)	(2.74)
SPEI $_{i,t}$		-0.205^{**}	-0.203^{**}
		(-2.21)	(-2.17)
Conflict $_{i,t}$		-0.0609	-0.0616
		(-1.05)	(-1.03)
Pop Density $_{i,t}$		0.00117	0.00122
		(0.63)	(0.64)
Instrument – First Stage			
IV1 $_{i,t}$		-0.0000177^{***}	-0.0000177^{***}
		(-3.57)	(-3.57)
N	10054	336	336
Year FEs		X	X
Origin FEs		X	X
Dest-Year FEs	X		
Ori-Year FEs	X		
Cty-Pair FEs	X		

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The first column reports the estimates of the following structural gravity model: $\frac{\text{M}_{ij,t}}{\text{M}_{ii,t}} = exp[S_{ij} + S_{j,t} + S_{i,t} + \ ln(\text{MigStocks}_{ij,t}) * \delta + \tau_{ij,t}]$. The estimated origin-year fixed effect coefficients $\hat{S}_{i,t}$ are used as dependent variable in the second step IV-PPML model (Columns 2-4) obtained using the same instrument described in Table A2.

Table B7: Monadic Estimation with Alternative Instruments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimator	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	OLS	OLS	OLS	OLS
Model	Monadic	Monadic	Monadic	Monadic	Monadic	Monadic	Monadic	Monadic
Depvar	Ln(EmigRate)	Ln(EmigRate)	Ln(EmigRate)	Ln(EmigRate)	Ln(EmigRate)	Ln(EmigRate)	Ln(EmigRate)	Ln(EmigRate)
					Reduced Form	Reduced Form	Reduced Form	Reduced Form
Log NTL pc $_{i,t-1}$	**0.889*****	**0.765****	**0.691*****	**0.716*****				
	(3.13)	**(2.71)**	**(2.86)**	**(3.02)**				
SPEI $_{i,t}$	-0.0881	-0.0593	-0.0419	-0.0478	0.116**	0.125**	0.128**	0.128**
	(-0.74)	(-0.54)	(-0.39)	(-0.45)	(2.26)	(2.25)	(2.27)	(2.29)
Conflict $_{i,t}$	-0.00308	0.00729	0.0135	0.0114	0.0707*	0.0636*	0.0712*	0.0666*
	(-0.05)	(0.13)	(0.27)	(0.22)	(1.94)	(1.94)	(2.04)	(2.02)
Pop Density $_{i,t}$	0.00102	0.000847	0.000744	0.000779	-0.000261	-0.000528	-0.000362	-0.000493
	(0.78)	(0.82)	(0.66)	(0.72)	(-0.23)	(-0.55)	(-0.37)	(-0.51)
IVs								
AID Econ $_{i,t+1, t+2}$						0.00267**		0.00158
						(2.25)		(1.57)
AID Prod $_{i,t+1, t+2}$							0.00216*	0.00138
							(2.03)	(1.62)
IV2 $_{i,t}$					-27437.7***			
					(-5.92)			
N	336	336	336	336	336	336	336	336
Origin Fes	X	X	X	X	X	X	X	X
Year Fes	X	X	X	X	X	X	X	X
AID Econ $_{i,t+1, t+2}$		X		X		X		X
AID Prod $_{i,t+1, t+2}$			X	X			X	X
IV2 $_{i,t}$	X				X			

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by country of origin.

Table B8: Alternative Instruments, First-Stage Results

	(1)	(2)	(3)	(4)
Depvar	$Log\ NTL\ pc_{i,t}$	$Log\ NTL\ pc_{i,t}$	$Log\ NTL\ pc_{i,t}$	$Log\ NTL\ pc_{i,t}$
Model	Monadic	Monadic	Monadic	Monadic
Instrumental Variables				
AID Econ $_{i,t+1,\ t+2}$		0.00348^{***}		0.00167
		(325)		(1.59)
AID Prod $_{i,t+1,\ t+2}$			0.00312^{***}	0.00230^{***}
			(3.76)	(3.15)
IV2 $_{i,t}$	-30870.74^{***}			
	(-3.25)			
N	336	336	336	336
Origin Fes	X	X	X	X
Year Fes	X	X	X	X
Kleibergen-Paap - F statistic	10.536	10.558	14.169	10.264
Hansen J-Stat / Chi-sq(1) P-val				0.7083

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by country of origin.

Table B9: Including a Squared Term

	(1)	(2)	(3)	(4)	(5)	(6)
Model	Gravity	Gravity	Gravity	Gravity	Gravity	Gravity
Estimator	PPML	PPML	PPML	PPML	PPML	PPML
Migration	Internal	Internal	Internal	Internal	Internal	Internal
Skill	All	Low	L-M	M-H	High	(4+5)
Log Mig Stock $_{ij,t}$	0.911^{***}	0.939^{***}	0.653^{***}	0.591^{***}	0.520^{***}	0.588^{***}
	(181.75)	(160.43)	(41.11)	(42.63)	(17.62)	(43.09)
Log NTL pc $_{i,t-1}$	**1.372^{***}**	**1.454^{***}**	**1.388^{***}**	**0.646^{**}**	**0.885**	**0.656^{**}**
	(13.80)	**(6.25)**	**(8.55)**	**(2.44)**	**(1.38)**	**(2.52)**
Log NTL pc (Sq) $_{i,t-1}$	**0.0455^{***}**	**0.0468^{***}**	**0.0468^{***}**	**0.0231^{***}**	**0.0377^{*}**	**0.0238^{***}**
	(13.43)	**(5.95)**	**(8.22)**	**(2.61)**	**(1.84)**	**(2.77)**
N	10080	10080	10080	10080	6463	10080
Origin FEs	X	X	X	X	X	X
Dest-Year FEs	X	X	X	X	X	X

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors are clustered by district of origin.
The models include the control variables, whose coefficients are not reported. Low=less than primary; L-M=primary; M-H=secondary; High=university; (4+5) = secondary + university.

Table B10: Monadic Estimates with LSMS Migration Data

	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)
Estimator	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML
District of Origin	Rural	Urban	Rural	Rural	Rural	Urban	Urban	Urban
Skill	All	All	0	0+1	2	0	0+1	2
Log NTL pc $_{i,t-1}$	-0.232	2.550***	-0.295	-0.339	-0.102	2.526***	3.673***	0.298
	(-1.23)	(6.65)	(-1.09)	(-1.57)	(-0.38)	(2.65)	(3.46)	(0.69)
N	292	44	292	292	281	40	40	40
Year FEs	X	X	X	X	X	X	X	X
Origin FEs	X	X	X	X	X	X	X	X

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors are clustered by district of origin.
The table reports the results of Equation (2) without control variables. Education Levels: 0 = Both parents with no education; 1 = At least one parent with primary education; 2 = At least one parent with secondary education. Migrants whose parents have at least a tertiary education degree haven't been considered because it would lead to a significant reduction of the sample size. Urban districts are: Lilongwe City, Blantyre city, Mzuzu city and Zomba city.