Empirical Economics (2023) 65:1823–1854
https://doi.org/10.1007/s00181-023-02393-y

Does access to credit alter migration intentions?

Marina-Selini Katsaiti[1,2,3] · **Maroula Khraiche**[4]

Received: 3 September 2021 / Accepted: 16 February 2023 / Published online: 5 March 2023

Abstract
This study investigates whether access to credit affects migration intentions. Using Gallup survey data for the years 2009 and 2010 for 17 African countries, we document a negative link between the ability to borrow and the desire to migrate. Being able to borrow reduces the likelihood of reporting wanting to migrate, especially for those with some education, those with lower income, for individuals with a bank account, and for those who feel their assets are safe. To deal with endogeneity, we assume that migration desire is driven by borrowing which in turn is determined by access to financial services. Therefore, we estimate a two-equation system for migration and borrowing, using variables describing access to financial services as instruments. We verify our findings using identification through heteroscedasticity and using a geographical instrument. Our results indicate that efforts to increase credit access in developing economies can cement residents' attachment to their home country.

Keywords Migration · Credit access · Migration costs · Survey data · Brain drain

JEL Classification F22 · J61 · O15 · G51

We thank Thanasis Stengos for helpful comments on the paper. All errors are our own.

✉ Maroula Khraiche
maroula.khraiche@utrgv.edu

Marina-Selini Katsaiti
mskatsaiti@aua.gr

1 Department of Regional and Economic Development, Agricultural University of Athens, Amfissa, Greece

2 Department of Innovation in Government and Society, College of Business and Economics, United Arab Emirates University, Al Ain, UAE

3 Hellenic Open University, Patra, Greece

4 Department of Economics, Robert C. Vackar College of Business and Entrepreneurship, University of Texas Rio Grande Valley, Edinburg, TX, USA

1 Introduction

In 2015, the United Nations identified financial inclusion as a key target in its development agenda, with many countries having since focused on increasing access to financial services and credit. Simultaneously, researchers have documented the impact of increased financial inclusion on various economic outcomes, such as economic growth and inequality in various geographic areas (Barajas et al. 2015; Dabla-Norris et al. 2015; Hajilee et al. 2017; Buera et al. 2021). In this work, we focus on the impact of credit access on international migration for a set of African economies. Theoretically, as workers gain access to credit, the borrowed funds generate a positive income effect, which loosens budget constraints and aid workers in covering fixed moving costs, required to migrate to foreign labor markets with higher wages (Rapoport 2002; Marchal and Naiditch 2020). On the other hand, access to credit can help workers undertake investment at home, which increases the opportunity cost of migrating abroad (Bazzi 2017).[1] Alternatively, if a worker's decision to migrate to a different labor market is part of an insurance contract to mitigate family risk of domestic income shocks (such as droughts), then access to credit capable of smoothing consumption when shocks arise may reduce the incentive to migrate (Stark and Lucas 1988). Additional considerations that shape the relationship between credit and migration are income and education (Bazzi 2017). For example, individuals with high income levels may not be constrained when it comes to financing moving costs; therefore, access to borrowing would not increase the likelihood of migration. Ultimately, the effect of borrowing on migration is not definitive a priori and warrants empirical investigation. Moreover, understanding the link between migration and credit is crucial to evaluating the effect of programs that increase credit access on labor markets. We thus aim to provide new evidence on the relationship between credit access and migration.

The existing empirical literature on the effect of individual borrowing on migration is limited and sometimes mixed evidence emerges. A negative effect of increased borrowing on domestic migration is identified for Thailand (Poggi 2019), a positive one is found for China (Cai 2020) and India (Singh 2018), while no effect is found in the case of Mexico (Angelucci 2015).[2] In the case of international migration, a positive relationship is found for the poor and low-skilled, from Mexico to the USA (Angelucci 2015; Kaestner and Malamud 2014); lack of finance is associated with lower propensity to emigrate from Indonesia (Mahendra 2014), while no association is found for the case of migration from Comoros to the richer neighboring French island of Mayotte (Gazeaud et al. 2023).

Our study contributes to the literature on the relationship between migration and credit by expanding the geographical scope and investigating 17 Sub-Saharan economies. Based on the Gallup World Poll (GWP), our study examines emigration desires instead of actual migration and identifies a negative result. Migration desires have been extensively used in the literature in place of actual migration (Dustmann

[1] Mesnard (2004) also shows that credit imperfections at home affect migration length and occupational choice of return migrants.

[2] Rubalcava and Teruel (2005) study the effect using similar data over a shorter time horizon but finds that migration decreases.

and Okatenko 2014; Van Dalen et al. 2005a; Liebig and Sousa-Poza 2004). Evidence shows that migration intentions and aspirations are good predictors of future actual migration (Van Dalen and Henkens 2008, 2013; Creighton 2013). More generally, according to psychology theories of reasoned action and planned behavior, individual's intention predicts actual decisions and behaviors (Fishbein et al. 1975; Ajzen 1991; Hale et al. 2002; Ajzen and Fishbein 2005). In fact, a 2012 report by GWP shows that 28.9 percent of the 386 million aspiring migrants ended up migrating (Docquier et al. 2015). Additionally, using migration intentions is less sensitive to policies of receiving countries and country-specific migration costs.[3]

Our results build on the work of Van Dalen et al. (2005a) and Liebig and Sousa-Poza (2004), who use survey data to study the desire to migrate, by going a step further and testing whether having borrowed with interest has an effect on the migration incentive. Our work also builds on the work by Dustmann and Okatenko (2014), who study the impact of various amenities on migration propensity by proposing that access to borrowing is also a factor that keeps workers at home.[4]

Our baseline empirical model assumes that respondents prefer to migrate if their utility from migration exceeds that of staying. If utility, defined over consumption, is linear and consumption is constrained by income and borrowed funds, then a preference for migration will also be determined by income and borrowed funds. However, borrowing itself is determined by access to financial services and associated costs. Our model implicitly assumes that financial services themselves do not provide utility directly to the respondents, except in their ability to provide necessary liquidity through borrowed funds. In terms of data, we proxy access to financial services with bank account ownership and variables documenting the costs of banking services such as fees and cumbersome paperwork. The recent increase in bank account ownership by adults, that reached 34% in Sub-Saharan African countries, is a result of improvements in access to banking services through mobile phones. This, in turn, is mainly driven by advances in mobile technology (World Bank, 2020), a fact arguably exogenous to the desire to move. Our implicit exclusion restriction, that access to financial services should only affect access to credit rather than migration intentions, passes a battery of tests for validity including the cutoff F-statistic suggested by Stock and Yogo (2005) and tests for overidentifying restrictions. Nonetheless, there is the possibility that the uptake of bank accounts is not purely exogenous, violating our exclusion restriction. Therefore, we employ identification through heteroscedasticity, a method developed by Lewbel (2012), as a second strategy to deal with endogeneity. The aforementioned strategy exploits non-spherical disturbances in the residuals of the first-stage regression to construct a valid instrument when exogenous instruments may not be available. As a third identification strategy, we employ a geographical instrument, which calculates the average number of respondents in the country who report having borrowed, excluding the region of the respondent, then use the aforementioned average as an

[3] And given that we look at migration intentions in large samples across individuals, regions, and countries, as available from the GWP, we can mitigate a potential concern that the patterns in our results stem from a systematic measurement error.

[4] We also contribute to the literature documenting the relationship between migration and income (Dao et al. 2018; Clemens 2014; Abramitzky et al. 2013).

external instrument for credit access. All our empirical models verify the negative association between borrowing and migration desire.

We calculate the marginal effect of having borrowed on wanting to migrate from a bivariate probit model for each country in our sample. Additionally, we show that the negative effect of borrowing on migration is larger for those with primary, secondary or tertiary education, indicating that perhaps access to credit can cement skilled workers' attachment to the home county. We also show that those with higher income are less likely to report a desire to migrate, especially when they have borrowed. Finally, we find that conditional on being able to borrow, an increase in bank account ownership reduces the desire to move, more so when respondents also report that their assets are safe.

In Sect. 2, we discuss the dependent variable, main explanatory variable, and other controls and present descriptive statistics. In Sect. 3, we discuss the empirical strategies and we present our results. In Sect. 4, we discuss marginal effects. Finally, in Sect. 5 we conclude.

2 Data and variable definitions

The Gallup World Poll, now an annual survey, conducted in about 160 countries worldwide, is representative of more than 99% of the world's population. The survey polls at least 1000 respondents in each country (2000 for the larger countries) and asks more than 100 questions. Gallup asks the same questions, every year, in the same way in the main questionnaire. All samples are probability-based and nationally representative, allowing for direct country comparisons. However, certain questionnaire waves often carry additional questions. Since we are interested in the relationship between desired migration and access to credit in developing economies, the sample we use is limited to Sub-Saharan countries, for which the survey waves conducted in 2009 and 2010 included questions about the desire to move and whether respondents borrowed. Our sample is dictated by the simultaneous availability of both variables (information on the desire to migrate and history of borrowing), along with the necessary controls, for the same years and countries. In the following subsections, we detail how we construct the variables.

2.1 Migration intentions—dependent variable

The survey asks the respondents, "Ideally, if you had the opportunity, would you like to move permanently to another country, or would you prefer to continue living in this country?". The answers available are, "Like to move to another country, Like to continue living in this country, and Do not know." We create a dummy variable we label as "desire to move" based on the aforementioned question such that the variable takes the value 1 if the respondent answers that they would like to move and zero otherwise. We drop observations of respondents who refuse to answer.

2.2 Borrowing in Sub-Saharan Africa—main explanatory variable

The survey asks the respondents, "Over the past year, have you borrowed money from any financial institution, i.e., bank, community savings group/savings club, microfinance institution, moneylender?". We create a dummy variable, we label as "Borrowed" based on the above survey question, taking the value 1 if the respondent answers "yes" and zero otherwise.

Although we cannot determine which institution respondents received funds from, we can investigate general trends about Sub-Saharan Africans' lending and borrowing practices. Sub-Saharan Africa's financial system remains largely underdeveloped and leaves a big bulk of the population outside any formal borrowing system. Therefore, most resort to informal borrowing from friends and family. According to the World Bank, in 2013, only 5% of households were able to secure credit from commercial banks across the continent. Some countries, such as Kenya, Ethiopia, Ghana, and Uganda, rise above the continent average, with more than 10% of adults having borrowed from formal sources. Alternatively, borrowers employ semi-formal lending channels, such as savings clubs or rotation savings and credit associations, which require memberships to receive loans as a mechanism to screen applications (Aryeetey 2005). Data on savings groups are scarce; however, some estimates by FinScope on adult participation in such groups in 2016 and later range around 37% in Uganda, 13% in Kenya, 12% in Nigeria, and 16% in Tanzania (Hoop et al. 2020), while GWP finds that in 2008 the regional median of respondents reporting that they depend "a little" on community savings programs is 20%, while a mere 5% reports to depend on such programs "a lot" (English 2008).[5] Although such clubs are mainly used to acquire funds for consumption, some farmers utilize the funds to finance working capital (Aryeetey 2005). Occasionally, some borrow from moneylenders, who typically charge high interest rates and require quick repayments, thus representing a solution of last resort. In terms of borrowing from microfinance firms, the industry remains relatively small in the continent (Van Rooyen et al. 2012). The aforementioned facts about formal, semi-formal, and informal lending channels in Sub-Saharan Africa are reflected in GWP survey responses for wave years of 2009 and 2010 (same time period as our sample). For example, when asked an open-ended question about how they would borrow if they needed money to start a business, 42% of Sub-Saharan African respondents reported that they were most likely to resort to their families (Marlar 2010b). 15% reported that they would resort to banks, 10% stated that they would resort to savings club, 4% answered that they would utilize microfinance,[6] while 2% reported that they would resort to money lenders. It is evident that family and friends remain an important source of borrowing for respondents, compared to formal and semi-formal sources. But a different trend emerges for bank account owners: 50% state that they would go to a bank if they needed to borrow, versus 24% who would resort to family instead. In contrast, for those individuals without a bank account, only 8% would go

[5] It is worth noting that at least 4 out of 10 respondents in Rwanda, Central African Republic, Cameroon,and Kenya say they depend at least a little on community savings groups to make a living, the highest participation in Sub-Saharan Africa.

[6] 43% of Sub-Saharan Africans reported that they have heard of microfinance institutions but they are not available in their community (Marlar 2010a).

to a bank, if they needed money, and 46% would ask their family members. This is not surprising; in general, bank loans require bank accounts as a prerequisite (Marlar 2010b). But it is worth noting that for the survey time period, only 1 in 5 Sub-Saharan Africans reports having a banking account. Barriers to open a bank account include multiple forms of identification or documents, such as pay slips (which are hard to come by for those in the informal sector), in addition to often prohibitively high associated costs. For example, in Sierra Leone the cost of maintaining a bank account was the equivalent of that country's GDP per capita in annual fees (Beck et al. 2008).

The importance of bank accounts for formal loan generation was documented by Allen et al. (2021) in the case of Kenya's Equity Bank, which significantly expanded its customer base both in Kenya[7] and in the neighboring countries. To grow its client base, the bank opened branches in low-income and underserved areas and expanded its operations in mobile banking. The bank also reduced minimum balance requirements and maintenance fees for deposit accounts. To encourage new clients to open bank accounts, the bank offered products tailored to their needs, such as small loans, the application for which simply required a National ID card, and accepted a wider range of options for collateral. Allen et al. (2021) find that the bank's branching strategy increased the probability of residents having both bank accounts and loans, especially for those living in rural and arid areas.

This trend of increasing bank account uptake (in Kenya and elsewhere in Africa) is expected to continue as the rise of Fintech alters how traditional financial services are offered by incumbent banks. The region stands as a global leader in mobile money adoption and usage and advances in Fintech will take it beyond its main use as an e-payment and transfer system (Sy et al. 2019). New services such as micro lending, crowdfunding, and peer-to-peer lending are taking shape in the region. For now, the banking sector in Sub-Saharan Africa remains inefficient and charges extremely high interest rates to compensate for risk.

2.3 Control variables

We control for the demographic characteristics of each respondent. First, we control for marital status and whether the respondent is self-employed. We also control for household income per capita in international dollars, which is calculated by dividing each respondent's reported household income by the total number of occupants in their household.[8] Other controls are gender, age and its quadratic term and household size. We also control for education through dummy variables for primary and secondary and tertiary education. Following (Dustmann and Okatenko 2014), we use correspondence analysis to create a wealth index as an additional control, based on questions regarding whether respondents own a TV and a cellphone and whether they had enough money for housing and food over the previous year. We list the questions in detail in Appendix A.

[7] From 5% of adult population in 2006 to 36% in 2015.

[8] A single international dollar buys an amount of goods and services in a country similar to the amount bought by a US dollar in the USA.

Since the desire to migrate is highly correlated with respondents' satisfaction with general amenities and domestic conditions (Dustmann and Okatenko 2014), we also control for the degree of satisfaction respondents express toward their home country using three indices. The first index we use is the National Institutions Index, which measures the confidence in national institutions, in particular the military, judicial system, the national government, and the honesty of elections. Secondly, we include a Corruption Index, which reflects perceptions about corruption in both business and government. Finally, we use a Law-and-Order Index that measures the level of security respondents perceive for themselves and their families, based on confidence in local police and past experiences with crime. We also control for respondents' perception regarding the safety of their assets using a question that asks "If someone wants to start a business in this country, can they trust their assets and property to be safe at all times?".

The previous literature has shown that receiving remittances can affect migration plans (Piracha and Saraogi 2017; Van Dalen et al. 2005a, b) and spur on development by reducing investment constraints (Stark and Lucas 1988). We thus control for whether respondents received remittances using the following two GWP survey questions: i) "Considering all the activities you engage in to make a living, how much do you depend on receiving money from family members working in other countries?" and ii) "Considering all the activities you engage in to make a living, how much do you depend on receiving money from family members working elsewhere in the country?". Note that we also control for remittances indirectly by including income in our regressions. When recording income for the survey, respondents are "instructed to include all income from all wages and salaries in the household, remittances from family members living elsewhere, and all other sources."

Having migrant relatives abroad can also increase migration desire even in the absence of remittances, since it can signal the existence of networks abroad. Therefore, we also control for whether respondents report having a family member abroad by including a dummy corresponding to their answer to the question: "Do you have relatives or friends who are living in another country whom you can count on to help you when you need them, or not?".

2.4 Descriptive analysis

In this section, we discuss general trends we observe and the characteristics of households in the sample for which we have observations on migration intentions, borrowing, and the control variables listed in Sect. 2.3.[9]

For each country, the share of individuals who report i) a desire to migrate and ii) having borrowed from a financial institution is shown in Table 1. We plot the relation between the two aforementioned variables in Fig. 1 and include a linear trend. A cursory examination of the aforementioned linear trend indicates a positive correlation between the two variables.[10] Our data point to differences across countries.

[9] Full descriptive statistics of all variables can be found in Appendix D, Table 11.

[10] However, this correlation is statistically insignificant. The slope coefficient of the linear trend is .094 with a standard error of .107 and a p value of 0.395.

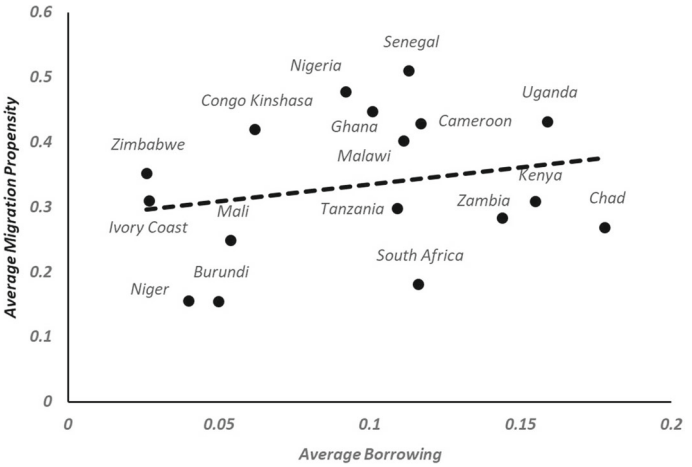

Fig. 1 Country shares: migration desire and reported borrowing

Table 1 Country shares of dependent variable and main explanatory variable

Country	No. of obs.	Desire to move		Borrowed from fin. inst.	
		Mean	SD	Mean	SD
Burundi	986	0.155	0.362	0.05	0.217
Cameroon	975	0.429	0.495	0.117	0.321
Chad	985	0.269	0.444	0.178	0.382
Congo Kinshasa	966	0.42	0.494	0.062	0.241
Ghana	884	0.447	0.497	0.101	0.301
Ivory Coast	991	0.31	0.463	0.027	0.163
Kenya	987	0.309	0.462	0.155	0.362
Malawi	999	0.402	0.491	0.111	0.314
Mali	966	0.249	0.433	0.054	0.226
Niger	983	0.156	0.363	0.04	0.195
Nigeria	762	0.478	0.5	0.092	0.289
Senegal	951	0.51	0.5	0.113	0.316
South Africa	917	0.181	0.385	0.116	0.32
Tanzania	988	0.298	0.457	0.109	0.312
Uganda	958	0.432	0.496	0.159	0.366
Zambia	938	0.283	0.45	0.144	0.351
Zimbabwe	928	0.352	0.478	0.026	0.159

Source: Gallup World Data, combined 2009 and 2010 waves

Examining Table 2, we find that individuals in the highest two quintiles of income, with secondary education or higher, and those employed, represent a higher fraction of the total number of people reporting to have borrowed and willing to migrate.

Table 2 Household characteristics

Percentage of respondents reporting they would like to move	
By income quintiles	
Poorest 20%	13.45
Second 20%	16.32
Middle 20%	17.89
Fourth 20%	21.34
Richest 20%	30.99
By education	
No formal education	9.65
Primary	24.66
Secondary	47.92
Tertiary	17.77
By employment status	
Employed	42.53
Self employed	22.43
Unemployed or out of the labor force	35.05
Percentage of respondents reporting they borrowed *By income quintiles*	
Poorest 20%	9.23
Second 20%	12.22
Middle 20%	16.87
Fourth 20%	23.55
Richest 20%	38.13
By education	
No formal education	8.06
Primary	22.58
Secondary	41.14
Tertiary	28.21
By employment status	
Employed	60.42
Self employed	25.63
Unemployed or out of the labor force	13.95

Source: Gallup World Data, combined 2009 and 2010 waves for sample used

3 Estimation

In this section, we outline our strategy for estimating the effect of borrowing on migration intentions. We discuss possible biases that can arise and then outline our identification strategies.

We assume that respondents have a linear utility function defined over consumption. Respondents will prefer to migrate if their utility of migrating exceeds the utility

of staying in the home country, given their relevant budget constraint. We define the perceived net gain in utility of migration for respondent i in country j as the latent variable $\triangle m^*_{ij}$. Since utility is linear, it will equal consumption, which in turn is determined by whether the respondent borrowed (b_{ij}) in addition to observable variables such as income, wealth, and employment status. Utility can also be determined by demographic factors such as age and marital status and by the degree of satisfaction with the home country conditions. The following equations govern the latent net benefit of migration and the binary intention to migrate (m_{ij}) reported by the respondent, respectively:

$$\begin{aligned} \triangle m^*_{ij} &= x'_{ij}\beta_1 + b_{ij}\alpha + c'_j\delta_1 + u_{ij} \\ \text{prob}[m_{ij} = 1] &= \text{prob}[x'_{ij}\beta_1 + b_{ij}\alpha + c'_j\delta_1 + u_{ij} > 0], \end{aligned} \tag{1}$$

where β_1 is the vector of coefficients associated with the vector of control variables x_{ij}, including income, wealth, employment, demographic characteristics, and variables measuring a respondent's satisfaction with the home country. α is the coefficient associated with borrowing b_{ij}. c_j are country dummies and δ_1 is the vector of the associated coefficients. Finally, u_{ij} is the random error.

Estimating Eq. 1 for migration intentions using ordinary least squares (OLS) is likely to suffer from omitted variables bias and reverse causality issues, causing endogeneity. Respondents' ability and attitude toward risk (which cannot be measured from our survey data) can affect both borrowing and migration decisions. If an omitted variable, such as ability, is positively correlated with both migration and borrowing, then an OLS estimate of α would be biased upward since $Cov(b_{ij}, u_{ij}) > 0$. For example, a respondent with high ability can fill out migration applications and loan applications in a more competent way than a low-ability respondent with identically measured characteristics (such as age and education level), making them both more likely to migrate and more likely to have received a loan. Having the aforementioned example in mind, estimating α using OLS will lead to an inflated coefficient. On the other hand, an omitted variable can be negatively correlated with borrowing and positively correlated with migration.[11] In this case, $Cov(b_{ij}, u_{ij}) < 0$ and the bias in an OLS estimate of α may be downward. Therefore, the net direction of bias is unknown.

Reverse causality may arise if the intention to migrate affects borrowing decisions. For example, those who wish to migrate might borrow to cover migration costs. Alternatively, those who wish to migrate may avoid borrowing if the cost and obligation of a loan hinder their relocation plans. In such cases, migration intentions, m_{ij}, and borrowing, b_{ij}, explain each other and thus: $m_{ij} = x'_{ij}\beta_1 + b_{ij}\alpha + u_{ij}$ and $b_{ij} = \lambda m_{ij} + x'_{ij}\beta_2 + z'_{ij}\gamma + v_{ij}$. The aforementioned reverse causality leads the OLS estimator to be biased. As Wooldridge (2015) shows, the sign of the asymptotic bias in the OLS estimator would be the same as the sign of $\frac{\lambda}{(1-\lambda\alpha)}$, if we drop x for simplicity. Assuming that $\lambda\alpha < 1$, then if $\lambda > 0$, implying that migration intentions lead to more borrowing, we will find the OLS estimator of α biased upward. Therefore, if we expect

[11] For example, a risk-taking individual may be more likely to use their own money or assets in order to invest, rather than borrow, and is more likely to engage in the risky endeavor of migration, compared to a respondent with identically measured characteristics.

that $\alpha < 0$, an upward bias would lead us to underestimate the effect of borrowing on migration intentions (OLS estimate of α would be smaller in absolute terms). On the other hand, if $\lambda < 0$, implying migration intentions reduce borrowing, then the OLS estimator would be biased downward, leading us to overestimate the effect of borrowing on migration when $\alpha < 0$ (estimate of α would be larger in absolute terms). In the special case of $\alpha = 0$, implying no effect of borrowing on migration desire, the OLS estimate would yield a positive or negative effect of borrowing on migration, depending on the sign of λ.

Due to the possible biases detailed above, we test whether the variable "having borrowed" is exogenous. We first apply the Durbin and the Wu–Hausman tests, both with the null hypothesis that the variable "having borrowed" is exogenous. Given the significance of both tests at the 5% level, we reject the null of exogeneity of the borrowing decision.[12] We thus treat the aforementioned variable as endogenous. We tackle endogeneity through the use of several identification strategies, outlined in detail in Sects. 3.1, 3.2, and 3.3.

3.1 Baseline instrumental variable estimation

In this section, we detail our baseline IV approach. Our ability to identify the causal effect of borrowing on migration intention relies on a set of variables we label z_{ij}, that explain an individual's ability to borrow without affecting migration intentions. Therefore, we specify a corresponding latent model of borrowing behavior. We assume that the latent variable b^*_{ij} is the amount a person can borrow which is again determined by the observable characteristics we labeled as x_{ij}. However, borrowing is also determined by access to financial services, which we label as z_{ij}.

We define the latent borrowing variable and the binary variable that is generated from the latent equivalent using the following equations:

$$\begin{aligned} b^*_{ij} &= x'_{ij}\beta_2 + z'_{ij}\gamma + c'_j\delta_2 \ + v_{ij} \\ \text{prob}[b_{ij} = 1] &= \text{prob}[x'_{ij}\beta_2 + z'_{ij}\gamma + c'_j\delta_2 \ + v_{ij} > 0], \end{aligned} \tag{2}$$

again, β_2 are the associated coefficients for x_{ij}, the vector of control variables. c_j are country dummies for each country j in our sample and δ_2 are the associated vectors of coefficients and v_{ij} is the random error. γ is the vector of the associated coefficients of z_{ij}.

In our discussion of the state of the financial system in Sub-Saharan Africa in Sect. 2.2, we detailed examples of how having a bank account is required in order to apply for small loans and how having low fees and less paperwork can increase access to credit. Therefore, we construct a dummy for those who report having a bank account. Then, we use a survey question asking respondents who reported having no bank account, "Why don't you have a bank account?" to create a cost of banking dummy variable. The dummy tracks whether respondents report that "bank charges and

[12] With the null hypothesis of exogeneity, the Durbin χ^2 test statistic is 4.48773, with a p value of 0.0341 and the Wu–Hausman F statistic is 4.47981 with a p value of 0.0343.

commissions are too high" or whether respondents report "cumbersome paperwork" is required by banks.

Intuitively, the premise of our identification strategy is that individuals do not derive utility directly from financial services, and therefore, access to financial services should not determine whether someone intends to migrate, except through these services' capacity to provide some needed liquidity to smooth consumption or finance investment. In other words, we assume that access to financial services per se is not a driver of migration intentions, but the credit and liquidity provided by financial services may affect migration intentions. Therefore, variables that reflect the availability of financial services could theoretically constitute valid instruments. Estimating Eqs. 1 and 2 using two-stage least squares (2SLS) regression with the IV's described above, we show that in addition to having a significant p value, our F statistic from the first-stage regressions exceeds 10, a threshold suggested by Stock and Yogo (2005) for the reliability of instruments. Additionally, the Sargan and the Basmann tests for overidentifying restrictions alongside Wooldridge's test are all not statistically significant. Wooldridge's score test is robust when the errors are not i.i.d., which is the likely case for our data. The results of the 2SLS estimation are presented in Table 4, and the first-stage results are presented in Table 5. Additionally, We run two more tests suggested by Stock et al. (2002) to check whether the IVs are weak, which we discuss in Appendix in B and report the results in Table 9. Nonetheless, overidentification tests do not guarantee the validity of instruments, particularly since our two instruments are highly correlated, with only the bank account ownership variable being highly significant. Therefore, in Sects. 3.2 and 3.3 we conduct the analysis with alternative identification strategies.

Our findings suggest that being able to borrow in the home country reduces the desire to leave the home country and migrate. A direct comparison between the 2SLS estimates in Table 4 and the OLS estimates in Table 3 suggests that the OLS results are biased upward and underestimate the extent to which borrowing reduces migration intentions. Therefore, the 2SLS may have corrected for reverse causality and omitted variables that are positively correlated with both borrowing and migration, as per our discussion in Sect. 3. Given that both our dependent variable and main explanatory variable are binary, a nonlinear model such as a bivariate probit (biprobit) may be more appropriate to fit the data. In Appendix C, we discuss the details of the biprobit model and present the results in Table 10. Our biprobit results confirm the negative association between borrowing and migration intentions identified above.[13]

Economic theory offers many explanations for the negative relationship between borrowing and attachment to the home country. Most likely, borrowing at home reduces the intensity of the "push factors," which the literature defines as factors relating to the conditions at home that propels workers to leave (in contrast, "pull factors" are related to better conditions abroad, typically higher income). If borrowed funds are used for entrepreneurial activities, then these funds present an opportunity for higher income in the future at home and perhaps more satisfactory working conditions, raising the opportunity costs of migration and reducing its appeal. Similarly, if the borrowing was used to purchase assets domestically, then the attachment to the home country

[13] The marginal effects are presented in Table 8 and are discussed in Sect. 4.

may increase (or perhaps the obligation to pay back debt may increase the likelihood of reporting wanting to stay home). Alternatively, if the borrowed funds were used to smooth consumption, then families no longer need to send a migrant abroad as a strategy to diversify risk and insure against negative economic outcomes.

Even though our IVs pass the standard tests for validity, threats to our identification strategy still exist. First, those who want to migrate may be more likely to open a bank account in the home country in order to send remittances back home. However, remittances are more commonly sent through mobile transfers, that can reach rural areas and typically don't require further identification mechanisms, such as proof of permanent address. Alternatively, a threat to identification could arise if those who wish to migrate might avoid costly bank account ownership, while those with no migration intention open bank accounts to facilitate their transactions at home. A second threat to identification is that the presence of banks in certain areas may be correlated with other amenities or commercial activities in those areas. In other words, both bank presence (and bank account ownership) and migration intentions can be driven by some other economic factors. This problem may be mitigated by adding country fixed effects and the various indices measuring the country's quality of institutions, corruption, and law and order. Third, the possibility that having a bank account is correlated with the respondents' entrepreneurial activity, which in turn can reflect a respondent's intent to migrate. One way we dealt with this concern is by controlling for whether the respondent owns a business in the country of origin or is fully self-employed. If any of these threats materialize in our data, then our estimates may not be valid and may be biased in either direction. In order to verify our findings, we employ two other identification strategies detailed in the next two sections.

3.2 Identification through heteroscedasticity

In order to verify the robustness of our results in Sect. 3, we propose the use of a different instrumentation methodology, à la Lewbel (2012). To illustrate this method briefly, we define m as the outcome variable, b as the endogenous variable, and x as the set of exogenous variables. Next, we write the empirical model using the following two equations: $m = x'\beta_1 + b\alpha + u$ and $b = x'\beta_2 + v$, where the errors u and v may be correlated and α and β_1 are coefficients we wish to estimate. Standard IV techniques, like the one used in the previous section, are based on the assumption that at least one element in x belongs in the b equation, but not in the m equation. Such an exclusion restriction raises the possibility of making the wrong assumption (that perhaps all the elements of x appear in the m equation). Lewbel (2012) suggested a new method, which offers identification through a simple linear 2SLS estimator for the coefficients of interest (α and β_1). This method is based on exploiting information embedded in the heteroscedasticity of v in order to construct valid IVs. The two steps of the Lewbel (2012) method are[14]:

1. Use OLS to estimate the first-stage residuals given by $\hat{v} = b - x'\hat{\beta}_2$.

[14] The discussion in this section is based on Baum and Lewbel (2019).

Table 3 Ordinary least squares

Variable	(1)	(2)	(3)
Borrowed	−0.0385**	−0.0386**	−0.0374**
	(0.0157)	(0.0157)	(0.0156)
Secondary edu or higher	0.0874***	0.0888***	0.0889***
	(0.0128)	(0.0127)	(0.0127)
Primary edu	0.0282**	0.0283**	0.0281**
	(0.0136)	(0.0138)	(0.0136)
National Institution Index	−0.000629***	−0.000627***	−0.000628***
	(0.0002)	(0.0002)	(0.0002)
Law-and-Order Index	−0.0014***	−0.0014***	−0.0014***
	(0.0002)	(0.0002)	(0.0002)
Corruption Index	0.0005***	0.0005***	0.0005***
	(0.0002)	(0.0002)	(0.0002)
Asset safe	−0.0242*	−0.0242*	−0.0243*
	(0.0134)	(0.0134)	(0.0134)
Gender	0.0587***	0.0579***	0.0580***
	(0.0107)	(0.0108)	(0.0108)
Married	−0.0609***	−0.0618***	−0.0590***
	(0.0113)	(0.0111)	(0.0108)
Age	−0.00407***	−0.00423***	−0.00547***
	(0.0014)	(0.0014)	(0.0005)
Age^2	−0.00002	−0.00002	
	(−0.00002)	(0.00002)	
Depend on remittances	0.0350*	0.0351*	0.0352*
	(0.0198)	(0.0194)	(0.0198)
Relative abroad	0.163***	0.164***	0.164***
	(0.0175)	(0.0174)	(0.0175)
ln(Income)	−0.0119**	−0.0120**	−0.0122**
	(0.0051)	(0.0051)	(0.0054)
Self employed	−0.0134		
	(0.0167)		
Wealth		−0.0104	
		(0.0324)	
HH size			−0.0003
			(0.0013)
Constant	0.553***	0.557***	0.581***
	(0.0483)	(0.0484)	(0.0511)
Observations	16,164	16,164	16,164

Table 3 continued

Variable	(1)	(2)	(3)
Number of Countries	17	17	17
R-squared	0.152	0.151	0.151

Standard errors, reported below the coefficients, are robust and clustered by region. Country fixed effects are included
***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level

Table 4 Two-stage least squares

Variable	2SLS (1)	2SLS (2)	2SLS (3)
Borrowed	−0.146**	−0.144**	−0.136*
	(0.0721)	(0.0713)	(0.0702)
Secondary edu or higher	0.0952***	0.0965***	0.0963***
	(0.0150)	(0.0150)	(0.0150)
Primary edu	0.0307**	0.0307**	0.0304**
	(0.0139)	(0.0141)	(0.0139)
National Institution Index	−0.0006***	−0.0006***	−0.0006***
	(0.0002)	(0.0002)	(0.0002)
Law-and-Order Index	−0.0014***	−0.0014***	−0.0014***
	(0.0002)	(0.0002)	(0.0002)
Corruption Index	0.0005***	0.0005***	0.0005***
	(0.0002)	(0.0002)	(0.0002)
Asset Safe	−0.0224*	−0.0224*	−0.0227*
	(0.0134)	(0.0133)	(0.0134)
Gender	0.0588***	0.0580***	0.0581***
	(0.0109)	(0.0109)	(0.0109)
Married	−0.0588***	−0.0597***	−0.0553***
	(0.0111)	(0.0108)	(0.0107)
Age	−0.00301*	−0.0032**	−0.0053***
	(0.0016)	(0.0016)	(0.0005)
Age^2	−0.00003	−0.00003	
	(0.00002)	(0.00002)	
Depend on remittances	0.0361*	0.0361*	0.0358*
	(0.0196)	(0.0193)	(0.0196)
Relative abroad	0.167***	0.167***	0.167***
	(0.0170)	(0.0169)	(0.0171)
ln(Income)	−0.0104**	−0.0105**	−0.0101*
	(0.0051)	(0.0051)	(0.0056)
Self employed	−0.0132		
	(0.0167)		

Table 4 continued

Variable	2SLS (1)	2SLS (2)	2SLS (3)
Wealth		−0.0125	
		(0.0316)	
HH size			0.0002
			(0.0014)
Constant	0.518***	0.522***	0.554***
	(0.0519)	(0.0516)	(0.0533)
Observations	16,164	16,164	16,164
Number of Countries	17	17	17
F-statistic	23.1434	23.2946	23.8682
p value	0	0	0
Sargan overid test	1.2605	1.2185	1.138
p value	0.2616	0.2697	0.2861
Basmann overid test	1.258	1.2161	1.1359
p value	0.262	0.2701	0.2865
Wooldridge overid score test	1.2348	1.1947	1.1166
p value	0.2665	0.2744	0.2906

Standard errors, reported below the coefficients, are robust and clustered by region. Country fixed effects are included
***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level

Table 5 First-stage results of the two-stage least squares from Table 4

Variable	(1) 1st stage of 2SLS	(2) 1st stage of 2SLS	(3) 1st stage of 2SLS
Bank Acc	0.1967***	0.1966***	0.1993***
	(6.83)	(6.82)	(6.87)
Banking Costs	−0.0323*	−0.0323*	−0.0344*
	(1.61)	(1.62)	(1.72)

1st of 2SLS refers to the first-stage estimates for the 2SLS estimator in Table 4
Standard errors, reported below the coefficients, are robust and clustered by region. Country fixed effects are included
Exogenous variables are included and are the same as the corresponding columns in Table 4
***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level

2. Define w as some or all elements in x and then using x and $(w-\tilde{w})\hat{v}$ as instruments, estimate the coefficients α and β_1, using linear 2SLS. Note that $\tilde{w}$ is the mean of w.

In his initial work, Lewbel (2012) does not show that the identifying assumptions are satisfied for when m or b is binary, but later on, Lewbel (2018) shows that they can be satisfied when the aforementioned variables are discrete, thus applicable to our case.

The key additional assumptions for identification using Lewbel (2012) are that $\text{Cov}(w, vu) = 0$ and $\text{Cov}(w, v^2) \neq 0$, where w is some or all elements in x Baum and Lewbel (2019). One condition that supports these assumptions is that w must be correlated with the squared error in stage 1 (v^2) so that the instrument is correlated with the endogenous variable b. To verify this condition in our estimation, we run the Breusch and Pagan (1979) test which has a null hypothesis of homoscedasticity. We reject the null in favor of heteroscedasticity for our vector of w variables (listed in Table 6) in the first-stage residuals. We also find that Hansen J test for overidentification restrictions is insignificant providing further evidence of the validity of the instrument. Results are presented in Table 6 and deliver a negative relationship between migration and credit.

Under identification through heteroscedasticity, the estimated coefficient for the effect of borrowing on migration is -.076 (as can be seen in Table 6, column 1). On the other hand, using the external instruments, the biprobit estimate of the average marginal effect of borrowing on migration is – .0946 and the 2SLS estimate is – .146 (as can be seen in column 1 of Tables 8 and 4, respectively). The absolute value of these three estimates exceeds the absolute value of the OLS estimate (which is – .0385, as can be seen in column 1 of Table 3). Therefore, the various estimation strategies generate different magnitudes for the marginal effect of borrowing on migration. If the Lewbel coefficient is the more accurate estimate, this would indicate that our identification strategy using external instruments in Sect. 3.1 did not sufficiently deal with endogeneity or omitted variables issues, such that the estimated coefficients using external instruments (in 2SLS and biprobit) overestimate the impact of borrowing on migration.

3.3 Geographical instrumental variables

Our third methodological approach in tackling endogeneity is using the "leave one out instrument," widely utilized in the literature. We instrument for having borrowed, using the average number of individuals who have borrowed in one's home country, excluding the respondent's own region. Therefore, the instrument varies at the regional level. Such a geographical instrument has been used by Bai et al. (2019), Dang and La (2019), Ansell (2008) and Acemoglu et al. (2019). The identification strategy assumes that regional differences in borrowing levels are driven by the same shock process, which implies a strong first-stage effect (Bai et al. 2019). In other words, differences in regional borrowing levels have a common cause such that regional variation in average borrowing of other regions would instrument for actual borrowing in the respondent's region (and the regional variation would in turn identify the effect of borrowing on migration for respondents). For example, if low borrowing is driven by the lack of financial regulation or high interest rates, then the variation in regional borrowing (other than the respondent's region) would pick up these effects and would be correlated with the respondent's own borrowing levels. A key identification assumption is that international migration desire in any region is determined by the region's conditions, rather than variations in borrowing elsewhere. We argue that this condition holds in our data. Indeed, Dustmann and Okatenko (2014) show that international migration

Table 6 IV estimation using Lewbel (2012)

Variable	(1)	(2)	(3)
Borrowed	−0.0758**	−0.0758**	−0.0911**
	(0.0339)	(0.0338)	(0.0443)
Secondary edu or higher	0.0905***	0.0902***	0.0929***
	(0.0135)	(0.0134)	(0.0138)
Primary edu	0.0292**	0.0291**	0.0293**
	(0.0139)	(0.0138)	(0.0139)
National Institution Index	−0.0006***	−0.0006***	−0.0006***
	(0.0002)	(0.0002)	(0.0002)
Law-and-Order Index	−0.0014***	−0.0014***	−0.0014***
	(0.0002)	(0.0002)	(0.0002)
Corruption Index	0.0005***	0.0005***	0.0005***
	(0.0002)	(0.0002)	(0.0002)
Asset Safe	−0.0235*	−0.0235*	−0.0234*
	(0.0135)	(0.0135)	(0.0135)
Gender	0.0587***	0.0588***	0.0581***
	(0.0107)	(0.0107)	(0.0108)
Married	−0.0602***	−0.0602***	−0.0570***
	(0.0111)	(0.0111)	(0.0107)
Age	−0.0037**	−0.0037**	−0.0054***
	(0.0015)	(0.0015)	(0.0005)
Age^2	−0.00002	−0.00002	
	(0.00002)	(0.00002)	
Depend on remittances	0.0353*	0.0354*	0.0355*
	(0.0197)	(0.0197)	(0.0195)
Relative abroad	0.165***	0.165***	0.165***
	(0.0172)	(0.0173)	(0.0172)
ln(Income)	−0.0115**	−0.0114**	−0.0110**
	(0.0052)	(0.005)	(0.005)
Self employed	−0.0135	−0.0133	
	(0.0165)	(0.0166)	
Wealth	−0.0121		
	(0.0316)		
HH size	−0.0001		
	(0.00130)		

Table 6 continued

Variable	(1)	(2)	(3)
Observations	16,164	16,164	16,164
Number of countries	17	17	17
Hansen J statistic	7.212	7.231	6.880
p value	0.782	0.780	0.737
Breusch–Pagan test for heteroscedasticity (*p* value)	0.00	0.00	0.00

Standard errors, reported below the coefficients, are robust and clustered by region. Country fixed effects are included

***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level

In each column, variables included in Z are all those in X variables except income, as described in Sect. 3.2

Breusch–Pagan tests for variables in Z create heteroscedastic residuals in the first stage

desire itself is driven by local conditions of respondents' regions such as quality of local labor markets and amenities, which we control for.

Nonetheless, a threat to identification can arise if a respondent is aware of regions with better borrowing opportunities and therefore plans to migrate internally instead of internationally, suggesting that borrowing in other regions can reduce international migration but raise internal migration. However, given the large upfront costs of international migration and the higher risk associated with it, internal and international migration cannot be considered perfect substitutes, an observation that would support our identification strategy. Our instrument passes the cutoff F-statistic suggested by Stock and Yogo (2005), exhibits strong first-stage results and delivers similar conclusions to the previous identification methods detailed above. The results are presented in Table 7.[15]

As can be seen in Table 7, column 1, the estimated coefficient for the effect of borrowing on migration is −.35, which deviates in magnitude from our previous estimates in Sects. 3.1 and 3.2 (although all the estimates have a similar negative sign). This deviation casts a doubt on the validity of the geographical IV. Comparing the results of our different identification strategies, we find that identification through heteroscedasticity delivers the weakest estimate for the impact of borrowing on migration (−.0758), while the marginal effect delivered by biprobit and 2SLS estimates points to a stronger impact of borrowing on migration desire (−.0946 and −.146, respectively).[16] All identification strategies deliver a stronger effect than that predicted by OLS (−.0385).

[15] The first-stage coefficient of the geographical IV model is negative. One explanation could be that limited funds are available for borrowing. Therefore, if many are borrowing elsewhere, few loans are left for the respondent's region. Alternatively, if certain central regions have bank branches (which can facilitate borrowing in said regions), then neighboring (less central) regions are less likely to have bank branches and residents are less likely to borrow.

[16] See column 1 of Tables 4, 6, and 8.

Table 7 Two-stage least squares—geographical instrument estimation

Variable	2SLS (1)	2SLS (2)	2SLS (3)
Borrowed	−0.350**	−0.350**	−0.349**
	(0.144)	(0.144)	(0.142)
Secondary edu or higher	0.122***	0.122***	0.122***
	(0.0195)	(0.0192)	(0.0191)
Primary edu	0.0322*	0.0326*	0.0322*
	(0.0169)	(0.0172)	(0.0170)
National Institution Index	−0.0004*	−0.0004*	−0.0004*
	(0.0002)	(0.0002)	(0.0002)
Law-and-Order Index	−0.0012***	−0.0013***	−0.0013***
	(0.0002)	(0.0002)	(0.0002)
Corruption Index	0.0006**	0.0006**	0.0006**
	(0.0002)	(0.0002)	(0.0002)
Asset Safe	−0.0328*	−0.0327*	−0.0336**
	(0.0169)	(0.0168)	(0.0170)
Gender	0.0683***	0.0683***	0.0683***
	(0.0140)	(0.0136)	(0.0136)
Married	−0.0642***	−0.0640***	−0.0529***
	(0.0135)	(0.0137)	(0.0142)
Age	0.0002	0.0002	−0.0047***
	(0.0019)	(0.002)	(0.0006)
Age^2	−0.00006***	−0.00006***	
	(0.00002)	(0.00002)	
Depend on remittances	0.0518**	0.0513**	0.0496**
	(0.0228)	(0.0225)	(0.0235)
Relative abroad	0.180***	0.180***	0.178***
	(0.0204)	(0.0203)	(0.0203)
ln(Income)	−0.0119*	−0.0119*	−0.00787
	(0.0066)	(0.0066)	(0.0078)
Self employed	0.0034		
	(0.0174)		
Wealth		−0.0328	
		(0.0410)	
HH size			0.0025
			(0.0017)
Constant	0.345***	0.340***	0.381***
	(0.0622)	(0.0633)	(0.0722)

Table 7 continued

Variable	2SLS (1)	2SLS (2)	2SLS (3)
Observations	16,164	16,164	16,164
Number of countries	17	17	17
		(First stage)	
Average borrowing in other regions	−3.7347***	−3.7311***	−3.7411***
	(0.1971)	(.1991)	(.2158)
F-statistic	359.811	351.909	301.231
P value	0	0	0

Standard errors, reported below the coefficients, are robust and clustered by region. Country fixed effects are included

***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level

Instrument included is described in Sect. 3.3

4 Marginal effects

Using the results of the biprobit model, we calculate the average treatment effect of having borrowed on the desire to migrate and list our estimates in column 1 of Table 8. For the full sample, we find that an individual who borrows is on average about 9.5 percentage points less likely to report a desire to migrate. We also calculate the average treatment effect of having borrowed for each country in our sample and report it in column 1 of Table 8. South Africa has the smallest average marginal effects, while Ghana and Malawi have the highest. Examining the descriptive statistics in Table 1, we find that for Ghana and Malawi, the average number of respondents reporting an intention to move is among the highest. Therefore, the larger marginal effect in these countries may indicate that the impact of credit on migration is stronger when a substantial fraction of the population desires to move.

Next, we focus on the effect of credit access on skilled workers' migration desires since brain drain is a concern for developing countries (Czaika and Parsons 2017; Bhagwati and Hamada 1974; Miyagiwa 1991; yiu Wong and Yip 1999). Additionally, international migration involves costs for physical relocation and requires information collection and paperwork submission in order to migrate. Such hurdles are more likely overcome by skilled workers rather than unskilled (Mayr and Peri 2008). Therefore, skilled workers may be better able to translate their desire to move into a reality.[17] We find that despite being more likely to report a positive intention to migrate, when they access credit, this desire decreases. Note that the coefficient on education levels is positive perhaps highlighting the brain drain phenomenon. However, access to credit reduces the migration desire of skilled workers in our data set. We report the average

[17] Brain drain is loosely defined as the international migration of skilled workers. Although some gains exist to the home economies from skilled emigration such as remittances, improved international networks, and expertise brought home through return migration, the extent of these gains is hard to quantify and the empirical evidence of such gains is mixed. While Beine et al. (2001) and Beine et al. (2008) document evidence of brain gain effects, Faini et al. (2003) fails to spot such an effect.

Table 8 Average marginal effects from bivariate probit

Average marginal effect of borrowing on the predicted probability of reporting a desire to move		
	(1)	(2) If respondent has primary, secondary or tertiary edu
Sample	−0.0946*	−0.105*
	(0.0517)	(0.0596)
Nigeria	−0.111*	−0.116*
	(0.0626)	(0.0670)
Kenya	−0.0906*	−0.104*
	(0.0490)	(0.0582)
Tanzania	−0.0978*	−0.109*
	(0.0536)	(0.0619)
Ghana	−0.112*	−0.115*
	(0.0634)	(0.0672)
Uganda	−0.108*	−0.115*
	(0.0607)	(0.0663)
Malawi	−0.111*	−0.116*
	(0.0631)	(0.0671)
South Africa	−0.0703*	−0.0853*
	(0.0363)	(0.0463)
Mali	−0.0969*	−0.109*
	(0.0526)	(0.0612)
Niger	−0.0738*	−0.137***
	(0.0386)	(0.0481)
Senegal	−0.111*	−0.116*
	(0.0625)	(0.0670)
Zambia	−0.0882*	−0.102*
	(0.0478)	(0.0572)
Cameroon	−0.102*	−0.112*
	(0.0559)	(0.0636)
Zimbabwe	−0.0861*	−0.1000*
	(0.0450)	(0.0548)
Burundi	−0.0824*	−0.0967*
	(0.0441)	(0.0538)
Chad	−0.0738*	−0.0887*
	(0.0400)	(0.0500)

Table 8 continued

Average marginal effect of borrowing on the predicted probability of reporting a desire to move		
	(1)	(2) If respondent has primary, secondary or tertiary edu
Congo (Kinshasa)	−0.104*	−0.113*
	(0.0568)	(0.0642)
Ivory Coast	−0.0947***	−0.107***
	(0.0350)	(0.0409)

Standard errors reported below the coefficients calculated by delta method. For each country's marginal effect, other countries' dummies are set to zero
***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level
Marginal effects are based on regression in column (1) of Table 10

marginal effect of borrowing on the predicted probability of reporting a desire to move when the respondent has primary, secondary or tertiary education in column 2 of Table 8.

Previous studies show that initially as income rises, migration rises, but then at high levels of income, migration decreases (Dao et al. 2018; Clemens 2014; Abramitzky et al. 2013). Intuitively, as income increases enough to finance moving costs, migration increases, but at high levels of income, migration decreases as the opportunity cost of emigration increases (since income can be allocated to investment and consumption smoothing at home Bazzi 2017). In our estimations, as income rises, migration desire decreases since the coefficient on income is negative (as can be seen in Table 10). Our estimations show that borrowing, which provides additional funds, further decreases migration intentions. Therefore, for our sample, respondents may be more likely to use the funds for investment or consumption smoothing rather than migration.

4.1 Banking access and migration

We investigate the extent to which access to banking can change the marginal effect of borrowing on migration desire. We find that the marginal effect of borrowing for those with a bank account is larger in magnitude (i.e., absolute terms) than for those without a bank account. Additionally, for those who feel that their assets are safe and have a bank account, the marginal effect of borrowing has an even larger magnitude than for those who do not. The changes in the marginal effects are shown in Fig. 2 and are based on the biprobit estimates from column 1 in Table 10.

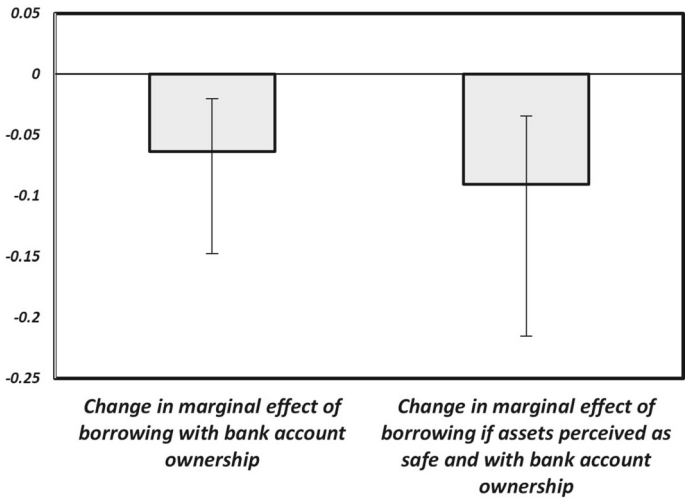

Fig. 2 Change in marginal effect of borrowing on reporting a desire to migrate: with/without bank account, assets safe/unsafe*. *Based on biprobit estimates from column 1 in Table 10. Vertical lines represent 95% confidence internals

5 Conclusion and discussion

Our work documents a direct link between borrowing and the desire to migrate in Sub-Saharan Africa. Having borrowed reduces the likelihood of reporting wanting to migrate, especially for those with higher levels of education and those who have a lower level of income initially. The likelihood of reporting a desire to migrate is also lower for those who borrowed while having a bank account and perceive that their assets are safe. We document this direct effect using Gallup survey data for the years 2009 and 2010. In order to deal with endogeneity issues, arising because of possible unobservable factors and/or reverse causality, we implement several identification strategies. All techniques, namely traditional IV estimation, Lewbel's (2012) instrumental variable method, the "leave one out IV" identification and bivariate probit, confirm the negative effect.

Theoretically, there are different channels through which borrowing could affect migration. First, borrowing offers additional liquidity and as a result eases budget constraints through an income effect. Hence, borrowing can facilitate migration of households or individuals who are liquidity-constrained. However, individuals may use borrowed funds to purchase assets or undertake investments, which can increase the opportunity cost of migration and keep them at home. Of course, there is the possibility that borrowing imposes collateral or other guarantee restrictions, burdening individuals and cementing their presence in the country of origin. Alternatively, if borrowing in the country of origin is used to smooth consumption against negative shocks, then being able to borrow reduces migration. This last channel is discussed in the seminal work of Stark and Lucas (1988), who show that as a part of a bargaining problem or due to altruistic reasons, families invest in helping a member migrate as a way to

mitigate risks of domestic markets.[18] In this case, migration and remittance are a part of a family contract that fills in the gap for the lack of formal insurance mechanisms at home.

Although we cannot ascertain the motivation for borrowing in our data set, given that we find a negative effect on migration, generally it is likely that the funds provided an income boost or liquidity for individuals to smooth consumption, or invest, which in turn can increase respondents' lifetime utility at home, and consequently reduce migration intentions.

In conclusion, we should note that our analysis comes with limitations. Although the overidentification tests did not reject the validity of our external IVs, they still may be invalid. Therefore, caution should be exercised over the implications of our results. If our instruments are in fact invalid, then the magnitude of the biprobit and 2SLS estimates may be biased (in either direction). The differences in the magnitudes of the coefficients across our specifications in Sects. 3, 3.2, and 3.3 may also be a cause for concern. Future research should identify stronger identification strategies, limiting the possibility of bias in the estimates.

Declarations

Conflict of 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.

A: Wealth Index construction

Using joint correspondence analysis (JCA) on the following Yes or No questions, we construct a wealth index such that higher values imply higher wealth.

- Have there been times in the past 12 months when you did not have enough money to provide adequate shelter or housing for you and your family?
- Have there been times in the past 12 months when you did not have enough money to buy food that you or your family needed?
- Does your home have a cellular phone?
- Does your home have a television?

The proportion of variance explained by the first component is 68.37. It is worth noting that in Dustmann and Okatenko (2014) the authors use additional variables capturing internet access and electricity. However, our data do not include such information. Additionally, while Dustmann and Okatenko (2014) use polychoric principal component analysis to construct their index, we use JCA.

B: Two tests for weak instruments

We run two more tests to check whether the IVs are weak and report them in Table 9. We report the test results for specifications in columns 1, 2, and 3 of Table 4. As

[18] See also Amery and Anderson (1995).

Table 9 Weak instruments tests (Appendix B)

Ho: instruments are weak					Associated reg. column	Test statistic
Largest relative bias tolerated	5%	10%	20%	30%		
2SLS relative bias	19.93	11.59	8.75	7.25	(1)	443.854
2SLS relative bias	19.93	11.59	8.75	7.25	(2)	443.887
2SLS relative bias	19.93	11.59	8.75	7.25	(3)	457.372
Largest rejection rate of a nominal 5% Wald	10%	15%	20%	25%		
2SLS Size of nominal 5% Wald	8.68	5.33	4.42	3.92	(1)	443.854
2SLS Size of nominal 5% Wald	8.68	5.33	4.42	3.92	(2)	443.887
2SLS Size of nominal 5% Wald	8.68	5.33	4.42	3.92	(3)	457.372

Associate regression estimates are reported in Table 4

suggested by Stock et al. (2002), the first test checks for relative bias in our 2SLS estimator compared to the standard OLS. We find that our 2SLS estimator suffers at most 5% of the bias delivered by the OLS estimator. This is evident by the fact that our test statistic reported in Table 9 is larger than the critical value for 5% relative bias. Therefore, our 2SLS estimator is superior to OLS and we can thus conclude that our instruments are not weak. The second test checks for the presence of large size distortions in our 2SLS estimates that can arise if our instruments are weak. As can be seen in Table 9, our test statistic is larger than the critical value, indicating that a Wald test at the 5% level can have an actual rejection rate of no more than 10%. Therefore, we supply additional evidence rejecting the null hypothesis of weak instruments.

C: Biprobit model

We rewrite the two equations given by 1 and 2 as follows:

$$\text{prob}[m_{ij} = 1] = \Phi[x'_{ij}\beta + b_{ij}\alpha + c'_j\delta_1] \tag{3}$$

$$\text{prob}[b_{ij} = 1] = \Phi[z'_{ij}\gamma + c'_j\delta_2], \tag{4}$$

where Φ is the evaluation of the standard normal cumulative distribution function (cdf). Then, we estimate Eqs. 3 and 4 jointly by Maximum Likelihood (ML) as a biprobit. To do so, we assume that the random errors u_{ij} and v_{ij} follow a bivariate normal distribution and allow that $\text{cov}(v_{ij}, u_{ij}) = \rho$. Given our discussion above on omitted variables bias and reverse causality, we suspect that b_{ij} is correlated with the error term u_{ij} and therefore is not exogenous, we expect the parameter determining the correlation between the aforementioned random errors (ρ) to not be zero. If b_{ij} were exogenous, then the correlation would be zero. Although we can use a two-step procedure to estimate a linear probability model using 2SLS, we cannot replicate this procedure here since the endogenous variable in our model is binary (Wooldridge 2010). In any case, Angrist (1991) shows that 2SLS and biprobit can deliver similar results.

The bivariate probit model is econometrically identified without an exclusion restriction. This provides another avenue to evaluate the identification arising from the exclusion restriction by comparing the estimated model that includes the exclusion restriction, to the estimated model that is identified purely from functional form. Therefore, we calculate the Akaike's information criterion (AIC) and Bayesian information criterion (BIC) for two different models. The first model is the one with the exclusion restriction where we use bank account ownership and cost of banking variables as instruments. The second model is one with the migration and borrowing equations determined by the same exogenous variables and no exclusion restrictions are imposed (variables are listed in Table 10, column 1). We find that the model with the exclusion restrictions delivers a smaller AIC and BIC (AIC = 26753.39, BIC = 27153.29), compared to the model without an exclusion restriction (AIC = 27275.9,

Table 10 Bivariate probit estimation (Appendix C)

Variables	(1)		(2)		(3)	
	$m_{i,j}$	$b_{i,j}$	$m_{i,j}$	$b_{i,j}$	$m_{i,j}$	$b_{i,j}$
Borrowed	−0.319*		−0.319*		−0.321*	
	(0.185)		(0.189)		(0.185)	
Secondary edu or higher	0.313***	0.247***	0.309***	0.250***	0.315***	0.252***
	(0.0499)	(0.0587)	(0.0505)	(0.0589)	(0.0502)	(0.0571)
Primary edu	0.132***	0.158***	0.132***	0.157***	0.133***	0.160***
	(0.0509)	(0.0596)	(0.0507)	(0.0597)	(0.0514)	(0.0587)
National Institution Index	−0.002***	0.00004	−0.002***	0.00005	−0.002***	0.00003
	(0.0006)	(0.0008)	(0.0006)	(0.0008)	(0.0006)	(0.0008)
Law-and-Order Index	−0.0044***	0.0007	−0.0044***	0.0007	−0.0044***	0.0007
	(0.0006)	(0.0008)	(0.0006)	(0.0008)	(0.0006)	(0.0008)
Corruption Index	0.0018***	0.0004	0.0018***	0.0004	0.0018***	0.0005
	(0.0006)	(0.007)	(0.0006)	(0.0007)	(0.0006)	(0.0007)
Asset Safe	−0.0720*	0.104**	−0.0717*	0.104**	−0.0716*	0.105**
	(0.0410)	(0.0511)	(0.0410)	(0.0511)	(0.0409)	(0.0508)
Gender	0.179***	−0.0307	0.182***	−0.0332	0.179***	−0.0330
	(0.0343)	(0.0380)	(0.0340)	(0.0384)	(0.0341)	(0.0370)
Married	−0.182***	0.0747*	−0.179***	0.0721*	−0.182***	0.0741*
	(0.0330)	(0.0381)	(0.0337)	(0.0379)	(0.0330)	(0.0380)
Age	−0.0006	0.0526***	−0.00002	0.0522***	−0.0005	0.0528***
	(0.0056)	(0.0072)	(0.0057)	(0.0073)	(0.0056)	(0.0072)
Age^2	−0.0002***	−0.0006***	−0.0002***	−0.0005***	−0.0002***	−0.0006***
	(0.00007)	(0.00009)	(0.00007)	(0.00009)	(0.00007)	(0.00009)

Table 10 continued

Variables	(1)		(2)		(3)	
	$m_{i,j}$	$b_{i,j}$	$m_{i,j}$	$b_{i,j}$	$m_{i,j}$	$b_{i,j}$
Depend on remittances	0.109*	0.0557	0.109*	0.0564	0.108*	0.0540
	(0.0595)	(0.0494)	(0.0602)	(0.0497)	(0.0588)	(0.0489)
Relative abroad	0.486***	0.0821*	0.485***	0.0819*	0.487***	0.0841*
	(0.0489)	(0.0468)	(0.0490)	(0.0469)	(0.0487)	(0.0466)
ln(Income)	−0.0352**	0.0298	−0.0350**	0.0295	−0.0352**	0.0295
	(0.0174)	(0.0239)	(0.0174)	(0.0240)	(0.0173)	(0.0238)
HH size	0.00113	0.0172***	0.00117	0.0171***	0.00118	0.0173***
	(0.00443)	(0.0063)	(0.0045)	(0.0063)	(0.0044)	(0.0063)
Self employed			−0.0444	0.045		
			(0.0530)	(0.0471)		
Wealth					−0.0640	−0.192
					(0.101)	(0.171)
Bank Acc		0.867***		0.868***		0.868***
		(0.0841)		(0.0838)		(0.0841)
Costly Banking		−0.228**		−0.229**		−0.230**
		(0.108)		(0.109)		(0.107)
Constant	−0.0508	−3.664***	−0.0444	−3.649***	−0.0554	−3.673***
	(0.183)	(0.273)	(0.0530)	(0.274)	(0.183)	(0.273)
Observations	16,164		16,164		16,164	
ρ	.118		.118		.119	
p value	0.2226		0.2334		0.2165	

Standard errors, reported below the coefficients, are robust and clustered by region. Country fixed effects are included
***Indicates a statistical significance at 1% level. **Indicates a statistical significance at 5% level. *Indicates a statistical significance at 10% level
The table shows estimated coefficients. For the marginal effects, please see Table 8

BIC = 27660.43). Therefore, the model with the exclusion restriction fits the data best according to these criteria.[19]

D: Descriptive statistics

In Table 11, we report the average, standard deviation, minimum, and maximum of all variables used in the analysis.

Table 11 Descriptive statistics (Appendix D)

Variable	Mean	Std. Dev	Min	Max
Desire to move	0.3316	0.4708	0	1
Borrowed	0.0972	0.2962	0	1
Secondary edu or higher	0.5288	0.4992	0	1
Primary edu	0.2876	0.4526	0	1
National Institution Index	49.5577	35.6369	0	100
Law-and-order Index	61.1895	33.3298	0	100
Corruption Index	82.8632	32.9780	0	100
Asset Safe	0.6416	0.4795	0	1
Male	0.5268	0.4993	0	1
Married	0.5155	0.4998	0	1
Age	33.0454	13.9435	15	99
Depend on remittances	0.1311	0.3375	0	1
Relative abroad	0.2066	0.4049	0	1
ln(Income)	6.1825	1.2084	1.2812	12.0314
Self employed	.1876	.3904	0	1
Wealth Index	−0.0440	0.1461	–.4771	0.2738
HHsize	6.7055	3.8926	1	29
Bank acc	0.1901	0.3924	0	1
Costly banking	0.9717	0.1659	0	1

References

Abramitzky R, Boustan LP, Eriksson K (2013) Have the poor always been less likely to migrate? evidence from inheritance practices during the age of mass migration. J Dev Econ 102:2–14

Acemoglu D, Naidu S, Restrepo P, Robinson JA (2019) Democracy does cause growth. J Polit Econ 127(1):47–100

Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50(2):179–211

Ajzen I, Fishbein M (2005) The influence of attitudes on behavior. Handb Attitudes 173(221):31

Allen F, Carletti E, Cull R, Qian J, Senbet L, Valenzuela P (2021) Improving access to banking: evidence from kenya. Rev Finance 25(2):403–447

[19] We also find that a model where bank account ownership and cost of banking variables are included in both the migration and borrowing equation delivers a higher AIC and BIC than our model with an exclusion restriction (AIC = 26755.45, BIC = 27170.74).

Amery HA, Anderson WP (1995) International migration and remittances to a Lebanese village. Canadian Geographer/Le Géographe Canadien 39(1):46–58
Angelucci M (2015) Migration and financial constraints: evidence from Mexico. Rev Econ Stat 97(1):224–228
Angrist JD (1991) Instrumental variables estimation of average treatment effects in econometrics and epidemiology. NBER technical working papers 0115, National Bureau of Economic Research, Inc
Ansell BW (2008) Traders, teachers, and tyrants: democracy, globalization, and public investment in education. Int Organ 62(2):289–322
Aryeetey E (2005) Informal finance for private sector development in sub-Saharan Africa. J Microfinance 7(1):3
Bai J, Jayachandran S, Malesky EJ, Olken BA (2019) Firm growth and corruption: empirical evidence from Vietnam. Econ J 129(618):651–677
Barajas A, Sahay R, Kyobe A, N'Diaye P, Mitra S, Cihak M, Yousefi S, Mooi Y (2015) Financial inclusion: Can it meet multiple macroeconomic goals? Technical report, International Monetary Fund
Baum CF, Lewbel A (2019) Advice on using heteroskedasticity-based identification. Stand Genomic Sci 19(4):757–767
Bazzi S (2017) Wealth heterogeneity and the income elasticity of migration. Am Econ J Appl Econ 9(2):219–55
Beck T, Demirguc-Kunt A, Martinez Peria MS (2008) Banking services for everyone? barriers to bank access and use around the world. World Bank Econ Rev 22(3):397–430
Beine M, Docquier F, Rapoport H (2001) Brain drain and economic growth: theory and evidence. J Dev Econ 64(1):275–289
Beine M, Docquier F, Rapoport H (2008) Brain drain and human capital formation in developing countries: winners and losers. Econ J 118(528):631–652
Bhagwati J, Hamada K (1974) The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis. J Dev Econ 1(1):19–42
Buera FJ, Kaboski JP, Shin Y (2021) The macroeconomics of microfinance. Rev Econ Stud 88(1):126–161
Cai S (2020) Migration under liquidity constraints: evidence from randomized credit access in China. J Dev Econ 142:102247
Clemens MA (2014) Does development reduce migration? Edward Elgar Publishing, Cheltenham
Creighton MJ (2013) The role of aspirations in domestic and international migration. Soc Sci J 50(1):79–88
Czaika M, Parsons C (2017). The gravity of high-skilled migration policies. Demography 54
Dabla-Norris E, Deng Y, Ivanova A, Karpowicz I, Unsal F, VanLeemput E, Wong J (2015) Financial Inclusion; zooming in on Latin America. IMF Working Papers 15/206, International Monetary Fund
Dang DA, La HA (2019) Does electricity reliability matter? Evidence from rural Vietnam. Energy Policy 131:399–409
Dao TH, Docquier F, Parsons C, Peri G (2018) Migration and development: dissecting the anatomy of the mobility transition. J Dev Econ 132:88–101
Docquier F, Machado J, Sekkat K (2015) Efficiency gains from liberalizing labor mobility. Scand J Econ 117(2):303–346
Dustmann C, Okatenko A (2014) Out-migration, wealth constraints, and the quality of local amenities. J Dev Econ 110:52–63
English C (2008) Many Africans count on community savings groups. Technical report, GALLUP
Faini R, et al (2003) Is the brain drain an unmitigated blessing. In: UNU-WIDER discussion paper no 2003/64. Citeseer
Fishbein M, Ajzen I, et al (1975) Intention and behavior: an introduction to theory and research
Gazeaud J, Mvukiyehe E, Sterck O (2023) Cash transfers and migration: theory and evidence from a randomized controlled trial. Rev Econ Stat 105(1):143–157
Hajilee M, Stringer DY, Metghalchi M (2017) Financial market inclusion, shadow economy and economic growth: new evidence from emerging economies. Q Rev Econ Finance 66:149–158
Hale JL, Householder BJ, Greene KL (2002) The theory of reasoned action. The persuasion handbook: developments in theory and practice 14:259–286
Hoop T, Howlett M, Meysonnat A, Desai S, Anderson CL, Schmidt C, Kaminsky M, Sidhu A (2020). Measuring savings group participation rates in Africa: data assessment and recommendations. Technical report, Evidence Consortium on Women's Groups
Kaestner R, Malamud O (2014) Self-selection and international migration: new evidence from Mexico. Rev Econ Stat 96(1):78–91

Lewbel A (2012) Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. J Bus Econ Stat 30(1):67–80
Lewbel A (2018) Identification and estimation using heteroscedasticity without instruments: the binary endogenous regressor case. Econ Lett 165:10–12
Liebig T, Sousa-Poza A (2004) Migration, self-selection and income inequality: an international analysis. Kyklos 57(1):125–146
Mahendra E (2014) Financial constraints, social policy and migration. Technical report, IMI Working Paper, 101
Marchal L, Naiditch C (2020) How borrowing constraints hinder migration: theoretical insights from a random utility maximization model. Scand J Econ 122(2):732–761
Marlar J (2010a) Few Sub-Saharan Africans aware of local microfinance options. Technical report, GALLUP
Marlar J (2010b) Sub-Saharan Africans bank on family for business loans. Technical report, GALLUP
Mayr K, Peri G (2008). Return migration as a channel of brain gain. Working paper 14039, National Bureau of Economic Research
Mesnard A (2004) Temporary migration and capital market imperfections. Oxf Econ Pap 56(2):242–262
Miyagiwa K (1991) Scale economies in education and the brain drain problem. Int Econ Rev 32(3):743–59
Piracha M, Saraogi A (2017) Remittances and migration intentions of the left-behind. Migr Dev 6(1):102–122
Poggi C (2019) Credit availability and internal migration: evidence from Thailand. J Dev Stud 55(5):861–875
Rapoport H (2002) Migration, credit constraints and self-employment: a simple model of occupational choice, inequality and growth. Econ Bull 15(7):1–5
Rubalcava L, Teruel G (2005). Conditional public transfers and living arrangements in rural Mexico. California Center for Population Research
Singh R et al (2018) Credit constraints and rural migration: evidence from six villages in Uttar Pradesh. Migr Lett 15(3):389–399
Stark O, Lucas RE (1988) Migration, remittances, and the family. Econ Dev Cult Change 36(3):465–481
Stock J, Yogo M (2005) Testing for weak instruments in linear IV regression. Cambridge University Press, New York, pp 80–108
Stock JH, Wright JH, Yogo M (2002) A survey of weak instruments and weak identification in generalized method of moments. J Bus Econ Stat 20(4):518–529
Sy MAN, Maino MR, Massara MA, Saiz HP, Sharma P (2019). FinTech in Sub-Saharan African countries: a game changer? International monetary fund
Van Dalen HP, Henkens K (2008). Emigration intentions: mere words or true plans? Explaining international migration intentions and behavior. In: Explaining international migration intentions and behavior (June 30, 2008)
Van Dalen HP, Henkens K (2013) Explaining emigration intentions and Behaviour in the Netherlands, 2005–10. Popul Stud 67(2):225–241
Van Dalen HP, Groenewold G, Fokkema T (2005a) The effect of remittances on emigration intentions in Egypt, Morocco, and Turkey. Popul Stud 59(3):375–392
Van Dalen HP, Groenewold G, Schoorl JJ (2005b) Out of Africa: What drives the pressure to emigrate? J Popul Econ 18(4):741–778
Van Rooyen C, Stewart R, De Wet T (2012) The impact of microfinance in Sub-Saharan Africa: a systematic review of the evidence. World Dev 40(11):2249–2262
Wooldridge JM (2010) Econometric analysis of cross section and panel data. The MIT Press, New York
Wooldridge JM (2015) Introductory econometrics: a modern approach. Cengage learning
yiu Wong K, Yip CK (1999) Education, economic growth, and brain drain. J Econ Dyn Control 23(5):699–726