

# Return Migration, Self-Selection and Entrepreneurship\*

Catia Batista<sup>†</sup>, Tara McIndoe-Calder<sup>‡</sup>, and Pedro C. Vicente<sup>§</sup>

Forthcoming *Oxford Bulletin of Economics and Statistics*

December 2016

## Abstract

Are return migrants more entrepreneurial? Existing literature has not addressed how estimating the impact of return migration on entrepreneurship is affected by double unobservable migrant self-selection, both at the initial outward migration and at the final inward return migration stages. This paper exploits exogenous variation provided by the civil war and the incidence of agricultural plagues in Mozambique, as well as social unrest and other shocks in migrant destination countries. The results lend support to overall negative unobservable return migrant self-selection, which results in an under-estimation of the effects of return migration on entrepreneurial outcomes when using a 'naïve' estimator that does not control for self-selection at both the initial migration and at the final return migration stages.

**JEL Codes:** F22; L26; O15.

**Keywords:** international migration; return migration; entrepreneurship; self-selection; business ownership; migration effects in origin countries; household survey; Mozambique; sub-Saharan Africa.

---

\* We thank the editor, James Fenske, and two anonymous referees for helpful suggestions on improving the paper. We also thank Randy Akee, Ron Davies, Elaine Liu, David McKenzie, Cormac O'Grada, Monica Parra, Matloob Piracha, Panu Poutvaara, Jackie Wahba, Karl Whelan, and other participants in seminars and conferences at the 2010 MIT NEUDC Meetings; 2011 NORFACE/CREAM Conference on International Migration; 2011 Oxford CSAE Conference; 2011 Mexico City World Bank/IZA Conference on Employment and Development; Nova University of Lisbon, Trinity College Dublin and University College Dublin for useful comments. We thank Julia Seither for providing us with data and comments on agricultural shocks in Mozambique. We wish to acknowledge financial support from the DFID - Department for International Development (UK), in the context of the International Growth Centre, from the Department of Economics and IIS at Trinity College Dublin, and from Nova Forum at Nova University of Lisbon. IRB approval for the fieldwork on which this paper is based was provided by the University of Oxford.

<sup>†</sup> Nova School of Business and Economics - Universidade Nova de Lisboa, CREAM, IZA and NOVAFRICA. Email: [catia.batista@novasbe.pt](mailto:catia.batista@novasbe.pt).

<sup>‡</sup> Central Bank of Ireland. Email: [tara.mcindoealder@centralbank.ie](mailto:tara.mcindoealder@centralbank.ie).

<sup>§</sup> Nova School of Business and Economics - Universidade Nova de Lisboa, BREAD, CSAE-University of Oxford, and NOVAFRICA. Email: [pedro.vicente@novasbe.pt](mailto:pedro.vicente@novasbe.pt).

## 1. Introduction

International emigration has been traditionally regarded as detrimental to the origin countries of migrants. Most concerns relate to the type of ‘brain drain’ issues originally proposed by Gruber and Scott (1966) and Bhagwati and Hamada (1974), and refer to the loss of the most educated nationals of a country, which causes the disappearance of a critical mass in production, research, public services (notably health and education) and political institutions. This negative effect would be compounded by the presence of positive production externalities or complementarities between human capital and other factors of production. In addition, fiscal losses would occur in the form of foregone tax revenue when educated nationals leave the country.

The effects of international migration on the economic development of migrant sender countries have, however, lately attracted renewed and considerable interest. In fact, recent studies have emphasized that emigration seems to have a positive impact on the educational attainment of both migrants and non-migrants, as well as on the demand for improved political institutions and on community engagement in the home country, as well as on international trade and FDI between the origin and destination countries of migrants.<sup>1</sup>

It can be argued that an additional channel through which migration may directly benefit home countries is through the return of migrants, who can bring new productive skills (such as education or managerial capacity) acquired abroad, as well as financial resources provided by past remittances and accumulated savings.

---

<sup>1</sup> See, for instance, Batista et al. (2012), Batista and Vicente (2011), Batista et al. (2016), Beine et al. (2008, 2011), Docquier et al. (2016), Gallego and Mendola (2013), Kugler and Rapoport (2007) and Javorcik et al. (2011).

While there are currently no systematic data on worldwide return migration, recent literature has focused on the international movements of students - the growing ‘brain circulation’ phenomenon.<sup>2</sup> UNESCO (2011) numbers show that the stock of foreign tertiary students in countries for which data are available was greater than 3 million in 2009, which doubles the corresponding number in 1999. Rosenzweig (2007) moreover argues that the proportion of foreign students who remain in the United States as permanent immigrants is only around 20% for the average sending country, which leaves a large room for ‘brain circulation’, i.e. the return of educated migrants to their origin country. In a different line of research, Gibson and McKenzie (2014) study New Zealand’s Recognized Seasonal Employer program, a temporary migration program that targets mainly unskilled workers. They accordingly find that migrants who return home tend to acquire human capital while abroad.

Despite the recent intensified interest regarding both the development impact of international migration for migrant countries of origin, and the temporary nature of some international migratory movements, there has only been limited research on the entrepreneurial effects of return migration – a literature discussed towards the end of this section. Most importantly, the existing literature evaluating the entrepreneurial impact of return migration has not taken into account the role of unobservable migrant self-selection, both at the initial migration and at the return migration stages, which this paper shows to be a serious impediment to a causal estimation of this impact.<sup>3</sup>

---

<sup>2</sup> Rosenzweig (2007) and Nyarko (2011) focus on the magnitude and effects of “brain circulation” from Asia and Ghana, respectively.

<sup>3</sup> Migrant self-selection on observable characteristics, notably education, has been a central topic of research since Borjas (1987) seminal work, notably followed by Borjas and Bratsberg (1996) for return migration and Chiquiar and Hanson (2002) emphasizing the importance of migration costs. More recent work has focused on migrant self-selection based on unobservable characteristics of migrants. See, for instance, Coulon and Piracha (2005), Batista (2008), Akee (2010) and Bertoli et al. (2013) using instrumental variable techniques, and McKenzie et al. (2010) using quasi-experimental

In this paper we propose to examine the question of whether return migrants contribute to entrepreneurship in the origin country. For this purpose, we conducted a representative household survey in four provinces of Mozambique during September and October 2009, when 1766 respondents were interviewed for this purpose. The retrospective nature of our dataset, as well as the characteristics of the Mozambican context that has migrants departing to different locations subject to a variety of exogenous shocks, allows us to address the issue of unobservable self-selection of return migrants both at the (outward) initial migration and at the (inward) final return migration stages, unlike previous literature. The data we collected and use in this analysis also facilitates an examination of predominantly south-south migration flows (between Mozambique and neighboring sub-Saharan African countries), which have been mostly ignored due to data unavailability in the past economics migration literature.

‘Naïve’ estimates of the entrepreneurial impact of return migration that do not take self-selection into account indicate that having a return migrant in the household contributes to increasing the probability of business ownership by nearly 13 percentage points (pp). However, because we are focusing on entrepreneurial outcomes, our estimates are likely to be affected by unobservable self-selection of individuals, at both the initial migration and at the final return migration stages: potentially, migrants and return migrants will differ substantially from non-migrants in terms of unobservable characteristics such as ability or entrepreneurial motivation, for instance, which should be correlated with entrepreneurial outcomes. Our results indeed highlight that the ‘naïve’ estimation results hide substantial unobservable self-selection bias. When we exclude the effect of migrant unobservable self-selection, both at the outward initial migration and at the inward

---

evidence. Note, however, that all these articles control for self-selection using income data only. An exception is the work by Fairlie and Woodruff (2007, 2010) that examine in detail patterns of observable self-selection of Mexican migrants and their self-employment decisions.

return stages, the impact of return migration on the probability of owning a business is estimated to be significantly larger: between 22 pp and 27 pp, depending on the method of estimation and source of variation that is used.

Note that, in order to identify migrant self-selection at the various stages, we use different sources of variation, such as displacement caused by wars and other violent events, agricultural plagues and macroeconomic shocks affecting differently origin and destination countries. Using these different sources of variation and also various estimation methods, namely next-neighbor matching and instrumental variable estimation, we obtain robust supportive evidence of an overall positive entrepreneurial effect of return migration, which increases after accounting for outward and inward unobservable self-selection. This implies that there is overall negative unobservable return migration self-selection.

Our work is most importantly related to a few relatively recent articles exploring the relationship between migration and entrepreneurship. Similarly to Dustmann and Kirchkamp (2002), we examine the occupational choice of return migrants, although we compare the decisions of return migrants and non-migrants instead of focusing on the determinants of the decisions to return and to become an entrepreneur. We are closer to McCormick and Wahba (2001), Mesnard (2004), Mesnard and Ravallion (2006) and Amuedo-Dorantes and Pozo (2006) in that they focus on examining the role of migration in overcoming wealth and credit constraints for businesses ownership. However, we take a broader perspective in that we look at the overall importance of return migration in promoting business ownership and explicitly tackle self-selection issues.

Piracha and Vadean (2010) and Wahba and Zenou (2012) both find, for Albania and Egypt, respectively, that return migration seems to promote entrepreneurship, particularly after an initial migrant re-integration period. However, even though both these papers take into account the problem of the

endogeneity of migration, they simply instrument the initial decision to migrate and never address the fact that there are multiple stages of self-selection in the decisions of a return migrant that may complicate causal estimation of the effects of return migration on entrepreneurship. This is exactly the focus and novelty of our paper, which discusses unobservable self-selection at both the initial migration and the return stages, while controlling for this problem using different sources of variation and estimation methods.

Finally, Yang (2008) explores exogenous variation in Filipino migrant income caused by the 1997 Asian financial crisis to find a positive impact of migrant income on investment and entrepreneurial activities in the home country. He however recognizes that this positive impact may be mediated by a number of channels, namely remittances, migrant savings or return migration. In this paper, we attempt to isolate the impact of return migration. In addition, further to controlling for self-selection in the decision to return, we also attempt to control for self-selection in the initial decision to migrate.

The remainder of the paper is organized as follows. In the next section, we begin by presenting a brief overview of Mozambique. We then proceed, in section 3, by describing the household survey we conducted and use in our empirical work, including a discussion of descriptive statistics. In section 4, we present the econometric model and identification strategy adopted in our empirical analysis. Section 5 discusses the main empirical findings, including a variety of robustness checks. Finally, section 6 summarizes our findings and presents policy implications.

## 2. Mozambique: Country Context

Mozambique, a country with 22.4 million inhabitants, is one of the poorest countries in the world with a GDP per capita of 838 USD in 2008.<sup>4</sup> Indeed, it ranks 161 in 189 countries (latest available years) in terms of GDP per capita.<sup>5</sup> Without important natural resources until recently, and with 81% of the population directly dependent on agriculture,<sup>6</sup> it has been an aid-dependent country for many years, with official aid assistance accounting for 22% of GNI in 2008.<sup>7</sup>

Politically, Mozambique became independent from Portugal in 1975, after an independence war that started in 1964 and officially ended in 1974. FRELIMO (Frente de Libertação de Moçambique), the independence movement, then started a single-party, socialist regime supported by the former Soviet Union and its allies. Starting in 1977, Mozambique suffered a devastating civil war fought between FRELIMO and RENAMO (Resistência Nacional Moçambicana). RENAMO was supported by Apartheid South Africa and, in the context of the Cold War, by the United States. The civil war ended in 1992 with an agreement to hold multi-party elections. FRELIMO has won all presidential elections since then.

Migratory movements from Mozambique were traditionally labor-driven mainly from the southern Mozambican provinces to South African mines and commercial farms. More recently, emigration from Mozambique has frequently been related to political instability. At independence, in 1975, most Portuguese citizens residing in Mozambique until this time returned to Portugal. During the subsequent civil war, mainly in the 1980s, large refugee movements were generated into neighboring countries. After 1992, peace in Mozambique attracted

---

<sup>4</sup> World Development Indicators, 2009.

<sup>5</sup> World Development Indicators, 2009.

<sup>6</sup> CIA World Factbook, 2010.

<sup>7</sup> World Development Indicators, 2009.

back over 1.7 million of its refugees and former combatant emigrants. More recently, in May and June 2008, xenophobic attacks in South Africa, against some of the poorest foreign immigrants (mostly Mozambican and Zimbabwean) resulted in the deaths of more than 60 people and prompted further substantial return migrant movements. Official reports point to 40,000 people fleeing back to Mozambique immediately after the onset of the violence.<sup>8</sup>

### **3. Data description**

#### **3.1. Household survey**

This study is based on a representative household survey including modules on business ownership and international migration. The survey was conducted in four provinces of Mozambique (Cabo Delgado, Zambezia, Gaza, and Maputo-Province) from September 2009 to October 2009 by the CSAE at the University of Oxford.<sup>9</sup>

The locations covered in the survey, 161 in total, were selected following a standard two-stage clustered representative sampling procedure - first on provinces, then on enumeration areas. The sampling framework was the 2004 electoral map of the country using as weights the number of registered voters per polling location (usually schools) as provided by the CNE/STAE (2004) in their 2004 elections (disaggregated) electoral data electronic publication.<sup>10</sup> This sampling procedure implies that all registered voters in the universe under consideration had the same probability of being sampled. The survey is based on a sample of 1763 resident households (including both non-migrants and return migrants), and also provides information on a large sample of current emigrants. Sampling in each enumeration

---

<sup>8</sup> Red Cross of Mozambique (2009).

<sup>9</sup> Figure A1 in Appendix illustrates the geographical coverage of the household survey.

<sup>10</sup> Comissão Nacional de Eleições - Secretariado Técnico de Administração Eleitoral (2004). Note that the 2009 electoral map only became available when fieldwork was already ongoing.

area followed standard household representativeness (n<sup>th</sup> house calls). However, only household heads or their spouses, one per household, were interviewed. Interviews were also conditional on ‘having access to a cell phone’ for receiving or sending calls and text messages. This included cases in which there was no ownership of cell phones in the household, but easy access to a neighbor or family member allowing cell phone usage.<sup>11</sup>

### **3.2. Descriptive statistics**

The dataset highlights the importance of international migration in Mozambique. Table 1 shows that 33% of all households in the sample have at least one member who is currently or has been an international migrant - while 23% of all sampled families have at least one return migrant living in the family home. Table 1 also shows that, in terms of business ownership, 28% of families in our sample report owning a business - 14% of which businesses are owned by return migrants.<sup>12</sup>

[Table 1 about here.]

Table 2 indicates that an overwhelming fraction of return migrants (72%) travelled to South Africa from Mozambique. There are, however, significant numbers of return migrants that departed to Tanzania (9%) and Malawi (7%). Most other migrant destinations are in Africa, while less than 5% of Mozambican migrants head to Europe (mostly Germany and Portugal). This geographic pattern of migration implies that this paper will essentially examine south-south migration flows.

---

<sup>11</sup> According to UNCTAD (2010), more than 80% of the Mozambican population had cell phone coverage in 2009. During fieldwork, having access to a cell phone proved an undemanding requirement on respondents, with only 3% of interviews not being completed due to lack of access to a cell phone.

<sup>12</sup> Table A3 in the online appendix presents the distribution of the different types of businesses present in our sample. The most prevalent business is street vending (46.9%), followed by agricultural businesses (33.7%) and services (15.2%). Return migrants however are significantly more likely than non-migrant households to own stores, and less likely to own agricultural businesses.

[Table 2 about here.]

Table 3 shows the summary statistics for all the variables used in the regression analysis that follows and not yet described in Tables 1 and 2. We find that the surveyed households are predominantly rural (only 29% are within 5km of a town), have relatively young household heads with low levels of education (close to 6 years of schooling, on average), expenditure (approximately 4 USD/day) and asset ownership. Further, around 15% of households report receiving remittances.

[Table 3 about here.]

## 4. Econometric Framework and Identification Strategy

### Econometric framework

Given that one cannot simultaneously observe the actual and the counterfactual entrepreneurial outcomes for each individual in our sample given their return migration status (and hence one cannot directly measure the *individual* entrepreneurial gain of return migration for this individual), we need to estimate an *average* entrepreneurial effect of return migration. This effect can be described as:

$$\begin{aligned} E[E_i | R_i = 1] - E[E_i | R_i = 0] &= \\ &= E[E_{1i} - E_{0i} | R_i = 1] + E[E_{0i} | R_i = 1] - E[E_{0i} | R_i = 0] \end{aligned} \quad (1)$$

where  $E_i$  and  $R_i$  are binary variables denoting, respectively, the entrepreneurial outcome and return migration status of individual  $i$ ;  $E_{1i}$  denotes the entrepreneurial outcome for a return migrant ( $R_i = 1$ ); and  $E_{0i}$  represents the entrepreneurial outcome for a non-migrant ( $R_i = 0$ ).

Equation (1) shows that estimating average entrepreneurial effects can be problematic. Indeed, this expression makes clear that simply comparing the average difference in entrepreneurial outcomes between return migrants and non-migrants will not identify a causal effect of return migration on entrepreneurship.

Indeed, the causal effect of interest, the *Average Treatment on the Treated (ATT)* effect,  $E[E_{1i} - E_{0i} | R_i = 1]$ , is shown to be masked by a *Selection Bias* that highlights that there would be differences in entrepreneurial outcomes between non-migrants and return migrants even if the latter had chosen not to emigrate in the first place.

An example of selection bias occurs when those who choose to emigrate are broadly more “able”, which could mean that they are more educated, motivated and driven than those who do not emigrate – all characteristics that should improve their entrepreneurial outcomes. In this instance, there is a positive selection bias, which implies that simply comparing average differences in entrepreneurial outcomes between return migrants and non-migrants exaggerates the true entrepreneurial skill gains of return migrants. Conversely, a negative selection bias (occurring if, for example, it is those individuals who lack observable qualifications, such as education, or who are less hard-working that decide to leave the origin country and, afterwards, return home) will understate the true entrepreneurial skill gains of return migrants when simply comparing average differences in entrepreneurial outcomes between return migrants and non-migrants. Note that the sign of self-selection is very much an empirical question: it is *a priori* equally possible to have negative self-selection or positive self-selection of migrants.

An additional issue is that the self-selection of migrants at any of the two relevant stages (initial or return migration) might occur based on observable or unobservable variables. Up to recently, the literature on migrant self-selection as started by Borjas (1987) based on Roy (1951), focused exclusively on self-selection based on observable characteristics, such as education and income. As

examined by more recent migration research<sup>13</sup>, unobservable migrant self-selection often operates based on unobservable personality traits, for instance, which are very likely to be correlated with our outcome of interest, entrepreneurship.<sup>14</sup>

### **Identification strategy**

In order to devise an identification strategy for our parameter of interest, it is important to examine the nature of the selection bias potentially affecting migrants in our sample. The thought experiment we have in mind is: “*What would be the estimated impact of return migration on entrepreneurial outcomes if we could choose to send abroad and bring back individuals who were randomly selected from the population of non-migrants residents in the home country?*”

This phrasing makes clear that there are two implicit selection stages in this thought experiment: first, randomly selecting non-migrants and sending them abroad; second, from the pool of randomly selected migrants, randomly choosing some of them to return to the origin country. This thought experiment would then avoid the two types of selection issues arising with return migration: (1) *(outward) self-selection at the initial migration stage*, which refers to the potentially idiosyncratic characteristics of those who decide to leave the country; (2) *(inward) self-selection at the return migration stage*, which refers to the potentially idiosyncratic characteristics of those migrants who decide to return to the sending country.

Given the expected self-selection of individuals into migration and return migration, the identification challenge is then to first find comparable return migrants and non-migrants in terms of observable and unobservable characteristics

---

<sup>13</sup> See, for instance, Coulon and Piracha (2005), Batista (2008), Akee (2010), McKenzie et al. (2010) or Bertoli et al. (2013).

<sup>14</sup> Batista and Umblijs (2014) present robust evidence that less risk-averse immigrants tend indeed to be more entrepreneurial.

*before the initial migration* decision is made; and second, within this restricted sample of migrants, to find return and current migrants that are similar in terms of observable and unobservable characteristics *before the return migration* decision is made.

We propose to use war events in the history of Mozambique to create the exogenous variation needed to simulate a randomly selected sample of outward migrants from our sample. Given the political context described in Section 2, it seems reasonable to expect that individuals who left Mozambique at the time of the independence and civil wars were migrating primarily as a result of events beyond their control - they were hence likely forced to leave the country independently of their characteristics, unlike economic migrants in non-war times.<sup>15</sup>

Still with the same purpose of randomly choosing migrants from the existing pool of non-migrants, we propose an alternative identification strategy that uses the exogenous variation provided by the geographic incidence of plagues over time in our sample.<sup>16</sup> Similarly to war events, plagues severely disrupt agricultural activities in a setting where families' lives depend on subsistence farming, leading to widespread outmigration from affected areas. This type of emigration is likely exogenous in the sense that plagues are outside individual control and ensuing migration should not, therefore, be systematically correlated with observed or unobserved migrant characteristics. The instrumental variable we construct for this purpose is a binary variable taking value 1 when a plague has occurred in any of

---

<sup>15</sup> In Table A2 in the online appendix, we present evidence that war migrants are on average younger and include a large proportion of females, consistent with whole families leaving the country as a result of the war, as opposed to non-war economic migrants that are typically working age males. In terms of education, which is most likely to positively correlate with unobservable characteristics related to entrepreneurship, we see no significant differences in education before migration, although it was significantly higher for war migrants after return to Mozambique.

<sup>16</sup> We make use of the disaster dataset provided by DesInventar, the UNISDR Disaster Information Management System, publicly available at <http://www.desinventar.net/DesInventar/index.jsp>. The plagues instrument we construct corresponds to any biological disaster identified in this database, including animal incidences, diseases or insect infestations as defined by the United Nations.

the 15 years prior to 2006, i.e. allowing for the average duration of migration spells in our sample before the year of the survey.<sup>17</sup>

Our identification strategy to remove the unobserved self-selection component from migration flows out of Mozambique is therefore to use both war events and incidence of plagues as exclusion restrictions. This strategy should provide a robust approach as the two restrictions use totally distinct sources of variation that occurred at different points in time (thereby avoiding potential cohort effects), and it should therefore mitigate any local identification concerns.

In order to generate a random sample of return migrants from the existing pool of return migrants that allows excluding self-selection at the return migration stage, we use events of forced return migration. In particular, we restrict the sample of return migrants to those who returned from South Africa immediately after the sudden eruption of the violent xenophobic riots against immigrants described in Section 2, as well as to those who were deported due to their illegal migration status<sup>18</sup>, and also to those who return to the origin country because of illnesses or deaths in the family. All of these return motives are likely to be exogenous in the sense that they are typically unanticipated and outside an individual's control. They should hence be uncorrelated with the individual's entrepreneurial outcomes except through the fact that these motives prompted the return itself.

An alternative identification strategy with the purpose of randomly choosing return migrants from the existing pool of migrants is to use the exogenous variation provided by changes in the GDP per capita difference between destination and origin countries, as well as the distance between the migrant origin

---

<sup>17</sup> Disaggregated disaster data quality improved massively from 1990, hence our choice of the 15 year window. We nevertheless constructed similar instrumental variables using a 20-year window and results were similar, although the instrument strength became weaker as could be expected. Similar results were obtained when allowing for a 10-year window or a different duration of migration spells.

<sup>18</sup> Note that illegal migration status is widespread in the Mozambican immigrant community residing in South Africa.

and destination areas. GDP differentials provide economic incentives to move back to the origin country as incomes change between origin and destination, whereas distance between migrant origin and destination also has predictive power for return migration decisions. Both these variables are exogenous in the sense that they are completely outside individual control and should not, therefore, be systematically correlated with migrant characteristics.

### **Estimation strategy**

The simplest possible estimate of the entrepreneurial gains to return migration would be obtained from a regression of the following form:

$$E_i = \alpha_0 + \alpha_1 R_i + \alpha_2' X + \varepsilon_i \quad (2)$$

where  $E_i$  is a proxy for entrepreneurship by individual  $i$  in our sample, such as business ownership or self-employment;  $R_i$  denotes whether individual  $i$  is a return migrant; and  $X$  denotes a set of observable individual, household and geographical characteristics that potentially affect entrepreneurial activity.

Following the discussion of the econometric framework summarized by (1), we know that an estimate for  $\alpha_1$  will only be equal to the causal effect of interest if the selection bias disappears after conditioning on observable characteristics  $X$ , i.e. if  $E[E_{0i} | X, R_i = 1] = E[E_{0i} | X, R_i = 0]$ . This is, however, unlikely to be the case, as the return migrant status,  $R_i$ , is most often correlated with the error term  $\varepsilon_i$ , which may include unobservable characteristics such as motivation, ambition, work diligence or risk preferences. These unobservable characteristics can be expected to affect both the actual entrepreneurial outcomes of non-migrants and the counterfactual outcomes of return migrants had they

decided not to migrate and return, hence creating an unobservable self-selection bias in this ‘naïve’ estimate.

Following the identification strategy discussed above, we will therefore pursue a few alternative estimation strategies in order to obtain estimates of the causal parameters of interest regarding the entrepreneurial outcomes of return migrants relative to those of non-migrants.

First, we will estimate a Linear Probability Model (LPM) on samples restricted following the identification exercises proposed in the previous subsection, so as to isolate selection bias effects and obtain an estimate for the effect of return migration on entrepreneurship.

Second, an alternative estimation method to obtain the effect of return migration on entrepreneurship (that can also be used to evaluate robustness of the results of running LPM on the restricted samples) will be to conduct nearest-neighbor matching (NNM) estimations.

An additional method to estimate the overall effect of interest, which can also be used to examine the robustness of our LPM and NMM estimates, is to perform a two-stage least square estimation of equation (2) using the instrumental variables proposed in the identification strategy described in the previous section. According to McKenzie et al. (2010), provided good instrumental variables can be found, the two stage least squares method is the best at excluding self-selection biases relatively to a random natural experiment.

The outcomes of these estimation strategies are discussed in the next section.

## 5. Empirical analysis

In this section, we summarize the main empirical results in this paper. In particular, we present, interpret and discuss the robustness of our estimates of the ‘entrepreneurial gains’ of return migration.

### Main empirical results

The entrepreneurial outcome we examine in the baseline results is business ownership at the household level. Table 4 displays the LPM estimates of the likelihood of business ownership for households that have at least one return migrant relative to households with no migrants.<sup>19</sup> These estimates are obtained while controlling for: (i) characteristics of the household and household head that may affect business ownership, such as the age and gender of the household head, as well as maximum completed education and number of persons belonging to the household; (ii) household level proxies for financial resource availability that may limit the possibilities of opening and running a business, such as household expenditure and asset ownership (where we focus on the most durable and precisely measured, namely home, land and car ownership); and, finally, (iii) geographical control variables such as migration destination, urban area and province fixed effects.<sup>20</sup>

[Table 4 about here.]

Column (1) in Table 4 shows that having a return migrant in the family is associated with a significant increase in the probability of owning a business. Our LPM estimates point to an average increase in the probability of business creation of 12.5 pp when there is a return migrant in the household. The magnitude and

---

<sup>19</sup> Note that we present LPM estimates for simplicity of interpretation. Running the same regressions using Probit yields essentially the same results.

<sup>20</sup> Fixed effects were only included for the two main migrant destinations in our sample: South Africa and other African countries.

statistical significance of this estimate is unaffected when we include controls for current migrants and remittances being received in the household, as shown in Column (2) of Table 4. While it could be argued that the entrepreneurial effects of current migrants and remittances could be captured to some extent by the return migrant variable, this does not seem to be the case in any of the specifications we run, where the estimated coefficient on return migration is always pretty much unaffected by the inclusion of these control variables.

As discussed in the previous section describing the identification strategy we use, the ‘naïve’ LPM estimate of a 12.5 pp increase in the probability of business creation when there is a return migrant in the household is likely combining the true effects of return migration on business ownership with the effects of unobservable self-selection of migrants both at the initial migration, and at the subsequent return migration stage. For this reason, we next estimate our LPM model using restricted samples following the identification strategy proposed in the previous section. This strategy has the purpose of excluding the two types of migrant selection bias effects, in order to obtain an estimate of the effect of return migration on entrepreneurship.

Following the identification strategy discussed in Section 4, we start by restricting the return migrant sample to “war migrants”, i.e. those return migrants who left Mozambique during wartime, as this migration decision is much less likely to be influenced by unobservable characteristics than the typical migration decision. This exercise should allow evaluating the effects of *unobservable outward self-selection at the initial migration stage*. This analysis can be done by simply comparing the ‘naïve’ LPM estimates for the whole sample to the LPM estimates for the restricted sample. In this setting, there is negative self-selection in the initial migration stage if the estimates obtained in the restricted sample are higher than the ‘naïve’ estimates run on the full sample. The intuition is that we are

focusing on a restricted sample of quasi-randomly chosen migrants, unlike the excluded self-selected migrants whose less entrepreneurial unobservable characteristics lower the average 'naïve' estimates. Therefore, excluding the self-selected migrant group increases the average returns to migration relative to the 'naïve' estimates. Positive self-selection occurs in the opposite situation, when the restricted sample estimates are lower than the 'naïve' estimates.

Columns (3) and (4) in Table 4 present the estimates of the LPM given by equation (2) when restricting the subsample of return migrants to those who left the country during war time. The estimated results show that there seems to be evidence of overall *negative outward self-selection* as the estimated impact of return migration on entrepreneurship is significantly raised when we restrict the estimation sample to war migrants compared to the estimation based on the whole sample of return migrants. These results can be understood in the context of the long history of Mozambican migration to South African mines and farms in non-war times. This history implies that strong migrant networks can lower migration costs and improve employment prospects even for migrants with lower unobservable ability. Our results show that these selected migrants were less able than a randomly selected war migrant to gain entrepreneurial skills during their migration experience.

Following our estimation strategy, we now proceed to further restricting the estimation sample of war migrants to include only individuals whose return was forced by exogenous motives - including the sudden eruption of violent xenophobic riots against foreign immigrants in South Africa, as well as deportation due to illegal migration status (widespread in the Mozambican immigrant community), or illnesses and deaths in the family. This further sample restriction should allow us to evaluate the *overall unobservable self-selection at both the*

*initial and return migration stages*, as well as to isolate *unobservable inward self-selection at the return migration stage*.

Similarly to our analysis of self-selection at the initial migration stage only, we will find that there is overall negative unobservable self-selection of return migration if the estimates obtained in the sample restricted to both war migrants and exogenous returns are higher than the ‘naïve’ estimates run on the full sample. The intuition is again that we are focusing on a restricted sample of quasi-randomly chosen migrants. Therefore, excluding the self-selected migrant group increases the average returns to migration relative to the ‘naïve’ estimates. Overall positive self-selection would occur in the opposite situation, when the restricted sample estimates are lower than the ‘naïve’ estimates.

Columns (5) and (6) in Table 4 show that return migrants who left Mozambique in war times and were forced to return by an exogenous motive are 24 pp more likely to own a business than non-migrants. This implies that there is *overall strong negative unobservable self-selection* when we consider both the initial and return migration stages, as the 12.5pp coefficient from the ‘naïve’ LPM estimation nearly doubles relative to the restricted sample estimation.

This 24 pp estimate is our proposed empirical counterpart to the ideal counterfactual thought experiment of assessing the true entrepreneurial gains of return migration by picking a random sample of non-migrants to emigrate and then picking a random subsample of those emigrants to bring back to the origin country. This is, hence, our proposed *baseline estimate for the true entrepreneurial gains from return migration excluding unobservable self-selection at both migration stages*.

Comparing this 24 pp estimate to the 14 pp estimate when restricting the sample of return migrants to those who emigrated during war times, we can evaluate the *unobservable inward self-selection of return migrants applicable to a*

*sample of non-selected emigrants*, such as war migrants. This self-selection pattern is clearly negative, meaning that the higher entrepreneurial gains obtained by randomly selected return migrants imply that it is less able war migrants that self-select to return to the origin country.<sup>21</sup>

Our results underscore the importance of controlling for both types of unobservable self-selection in estimating the entrepreneurial effect of return migration.

### **Robustness check: Nearest-Neighbor Matching Estimation**

A possible concern with the estimation strategy used in our baseline results might be that the use of a linear probability model (or indeed of a probit model, which yields very similar results) imposes linear assumptions that are too restrictive to adequately identify our parameter of interest. To this effect, we redo our estimation of the average treatment effect (ATE) of return migration on the probability of owning a business, but now using a non-parametric matching method. The purpose of this procedure is to investigate whether results are sensitive to the linear approximation embedded in the LPM.

To implement this approach, we rely on the nearest-neighbor matching (NNM) procedure proposed by Abadie and Imbens (2006). This matching approach ensures that return migrants are only compared to non-migrants who are sufficiently similar to them in terms of observables.

[Table 5 about here.]

---

<sup>21</sup> Note that we are not able to precisely estimate the entrepreneurial gains of return migration using a restricted subsample that includes only those return migrants who were forced to return from the sample of all emigrants (i.e. including emigrants who left the country at war and non-war times). We are therefore unable to infer the pattern of unobservable inward self-selection for (selected) return migrants in general using this strategy.

The results shown in Table 5 are obtained from carrying out the same estimations as presented in Table 4 but using NNM methods. The NNM results confirm the LPM results. The estimated average treatment effects are very similar in magnitude and statistical significance to those produced using LPM estimation for the various estimation samples used.

Specifically, in the unrestricted sample, Column (1) of Table 5 shows that the ‘naïve’ NNM estimate for the increase in the probability of return migrants owning a business is 11 pp, which compares to 12.5 pp in the LPM estimation.

When restricting the sample to those return migrants who initially left the country in times of war, this coefficient is raised to 19 pp, as can be seen in Column (2) of Table 5. This point estimate is higher than the 14 pp provided by the LPM results, but similarly provides evidence supportive of *negative unobservable selection at the initial migration stage*.

After restricting the sample to those migrants who were forced to leave due to war and forced to return due to reasons beyond their control, the estimated effect of migration on entrepreneurial outcomes becomes 27 pp, as shown in Column (3) of Table 5. This compares to the 24 pp estimate using LPM, although the NNM is a slightly less statistically significant estimate due to the reduction in sample size imposed by the common support imposed in the NNM estimation. This still underscores the importance of controlling for both types of migration self-selection, and strengthens the findings from the LPM estimated coefficients according to which *overall unobservable self-selection at both stages of migration is negative, and a compound effect of both negative unobservable self-selection at the initial and at the return migration stages*.<sup>22</sup>

---

<sup>22</sup> Similarly to what happens when estimating a LPM on our sample of migrants, NNM cannot provide statistically significant estimates for the entrepreneurial gains from exogenous return

### **Robustness check: Instrumental Variable Estimation**

In order to verify the robustness of our empirical findings, we now provide alternative sources of variation and a different (instrumental variable) estimation approach to identify the entrepreneurial gains of return migrants in our sample, as well as implied migrant self-selection patterns.

In a first step, we use the variation provided by agricultural plagues affecting Mozambique since approximately 1990, as described in Section 4. In a setting where most individuals practice subsistence agriculture, this type of shock is likely to induce strong migration movements, uncorrelated to individual unobservable characteristics, which provides an alternative way to quasi-randomly select migrants out of Mozambique and hence to investigate selection patterns at the initial migration stage.

In order to tackle self-selection problems at the return migration stage, and randomly choose return migrants from the existing pool of migrants to evaluate the unobservable inward self-selection bias occurring at the return migration stage, we need to find instruments that are strongly correlated with the decision to return and uncorrelated with the decision to own a business in the home country. We propose to use the set of exclusion restrictions described in the identification strategy section of the paper. Namely, we construct instrumental variables that make use of the exogenous variation provided by shocks to GDP per capita in the migrant destination countries, as well as by differences in the distance between areas of residence and migrant destinations.

Variation at the individual level for the instrumental variable summarizing the shocks to relative GDP per capita at migrant destination relative to the origin is

---

migration from the pool of existing migrants (including both those who migrated at war and at non-war times). We therefore still cannot infer the pattern of unobservable self-selection of return migrants in general using our sample.

achieved in the following way. The instrumental variable is computed as a weighted average of the changes in the GDP per capita difference between destination countries and Mozambique, where weights are constructed in order to reflect the relative size of the Mozambican migrant population in each migrant destination – a proxy for existing migration networks at each destination.<sup>23</sup> This migration-weighted GDP variable is matched to individual migrants in the sample in the year that they turn 30, which is the age at which individuals are most likely to start a business.<sup>24</sup> Note that the IPUMS census information from the Minnesota Population Centre (2010) is only available for migrants to South Africa, Tanzania and Portugal, which together account for 82% of the total number of return migrants in the sample.<sup>25</sup>

It seems reasonable to expect that changes in relative GDP per capita between origin and destination countries are unanticipated and outside the migrants' control. This relative income variable should hence be uncorrelated with the migrants' choice to own a business at origin, except through the fact that this motive prompts the return itself – as it is likely that these relative income changes provide economic incentives for return migration decisions. In the same way, it is also expected that the distance between survey districts and migrant destinations has predictive power for return migration, but should not be directly correlated with the decision to own a business, except through the fact that it prompts the return migration decision itself.

---

<sup>23</sup> These weights are computed as the number of Mozambican nationals between the ages of 25 and 34 resident at destination, as a proportion of the non-migrant resident population at the same destination. The weights are established using Mozambican and national IPUMS census data from the Minnesota Population Centre (2010).

<sup>24</sup> Note that changing the year at which individual migrants are matched with the GDP weighted variable by one or two years does not make a difference for the validity of the instrumental variables nor the estimation results that are obtained.

<sup>25</sup> The distance measures are naturally available for all migrants, while in constructing the migration-weighted GDP variable we are constrained by the IPUMS census data availability, which is restricted to South Africa, Tanzania and Portugal only among the Mozambican migrant destinations.

[Table 6 about here.]

The statistics obtained in our sample and displayed in Table 6 provide evidence supportive for these IV rationales. Looking at the full sample, the agricultural plagues, the destination-distance and the relative GDP per capita instruments are each and both together significantly correlated with the decision to return, as required of a strong instrument and confirmed by the Kleinberg-Papp F-statistics reported in Columns (2) – (5) of Table 6. These instruments also seem uncorrelated with the decision to own a business in the origin country and indeed pass the tests of over-identifying restrictions shown in Columns (4) and (5) of Table 6.

The two stage least square estimation results obtained using all instruments are shown in Column (5) of Table 6. The point estimate of 23.9 pp for the incremental impact of return migration on business ownership is positive and statistically significant. This is our preferred IV specification in that it uses all instruments and thereby attenuates concerns regarding the local validity of each of these instruments when used separately. It is remarkably close to the 24.3 pp estimate obtained using the restricted sample LPM approach.

We next turn to restricting the IV estimation sample to those migrants who initially left at war times, in order to check the robustness of our estimate using a different estimation approach. In particular, focusing on a sample of migrants who are less likely to be self-selected than the average migrant in our sample, as discussed before, while also accounting for self-selection at the return migration stage by using an IV strategy, should allow us to identify the effect of return migration on business ownership while minimizing unobservable self-selection concerns at both the initial and return migration stages. As shown in column (6) of Table 6, while the lower number of observations decreases the F-statistics for the distance instrument, the estimation results are very much in line with the ones

obtained using the two alternative estimation methods we proposed before. Indeed, the coefficient of 21.7 pp obtained for the increased probability of a return migrant owning a business relative to a non-migrant is not far from the 24pp obtained when using the restricted sample LPM estimation, and also the full IV strategy, or even the 27pp obtained when using the war restricted sample NNM estimation.

### **Discussion of Results**

The (LPM and NNM) ‘naïve’ estimates not accounting for unobservable self-selection of initial emigration are lower than the war restricted sample (LPM or NNM) estimates, the agricultural plague IV full sample estimate and the war restricted sample IV estimate. This is clear evidence of negative unobservable self-selection bias at the initial migration stage, in the sense that typical economic migrants in our sample are much less likely to open businesses than forced migrants (either because of war events or agricultural plagues). In other words, even though it is likely that forced migrants are much more similar to the non-migrant population in general, which is sufficient to support our claim of negative unobservable migrant self-selection, one way of interpreting our results is to say that forced migrants have unobservable characteristics that make them more likely to open businesses than the typical self-selected economic migrants.

The negative unobservable migrant self-selection at the initial migration stage we estimate can be understood in the context of an historical tradition to emigrate to South Africa mines and farms since the end of the XIX century, and the consequent prevalence of migrant networks, which are likely to substantially lower migration costs. This context is in this sense similar to that faced by Portuguese migrants to France, where Batista (2008) also finds evidence of negative unobservable migrant self-selection. This is a situation rather different from that of Mexican or Tongan emigrants, who face high migration costs and positive

unobservable self-selection patterns as obtained by Chiquiar and Hanson (2008) and McKenzie et al. (2010).

The direction of the unobservable self-selection bias at the final return migration stage is not as clear. The restricted sample estimation approach (both LPM and NNM) points toward negative self-selection of return migrants, meaning that typical self-selected return migrants are generally less likely than current migrants to have the unobservable characteristics that make them entrepreneurial. The opposite is suggested by the instrumental variable estimation approach that shows a decline in the estimated coefficient when controlling for both types of self-selection - column (5) of Table 6 - relatively to when controlling for the initial migration self-selection only – in column (1) of Table 6. This would mean that it is actually the most entrepreneurial migrants that decide to return, when they are not pushed to do so.

Overall our empirical results lend support to negative unobservable self-selection of return migrants. This unobservable self-selection pattern means that the effects of return migration on entrepreneurial outcomes when using ‘naïve’ estimators not accounting for self-selectivity are under-estimated. From our analysis it clearly comes out that there seems to be an overall positive entrepreneurial effect of return migration, ranging between 22 and 27pp after accounting for both outward and inward unobservable self-selection using a variety of sources of variation and estimation approaches.

## **6. Concluding remarks**

This paper examines broad research questions related to the entrepreneurial impact of return migration. More precisely, it evaluates the effect of sending random national residents abroad and then randomly bringing them back to the origin country, while also evaluating the benefits of bringing an average existing

migrant back. For this purpose, we conducted and make use of a tailored representative household survey of Mozambique. This is a typical Southern African country in the sense that migratory flows are mostly directed at neighboring countries. We hence also indirectly ask the important question of whether Southern Africa's growth promise may be supported by the current substantial south-south migration flows in the area.

Our methodological contribution highlights the importance of going beyond simple 'naïve' comparisons of entrepreneurial outcomes between return migrants and non-migrants in order to avoid estimation biases related to unobservable self-selection of migrants. Indeed, because we are focusing on entrepreneurship, our estimates could potentially be rather sensitive to unobservable self-selection of individuals into initial migration and subsequent return migration: migrants and return migrants may differ substantially from non-migrants in terms of unobservable characteristics correlated with entrepreneurial outcomes, such as ability or motivational drive, for instance.

In order to identify migrant self-selection at these various stages, we use different sources of variation, namely displacement caused by wars and other violent events, as well as agricultural plagues, but also macroeconomic shocks at destination and physical distance between migrant origin and destination. Using these different sources of variation and also various estimation methods, we obtain evidence of overall negative unobservable self-selection of return migrants, when controlling for unobservable self-selection at the initial and return migration stages.

Conducting 'naïve' LPM estimation shows that being a return migrant is associated with a significant increase of 13 percentage points in the probability of owning a business relative to non-migrants. However, when we exclude the effect of unobservable self-selection, both at the outward initial migration and at the inward return stages, we obtain that this effect is significantly larger and significant

- between 22 pp and 27 pp, depending on the method of estimation and source of variation that are used. The fact that we use alternative, totally distinct sources of variation that yield the same qualitative results and quantitative estimates of similar magnitude is reassuring regarding the validity of our analysis.

The external validity of our estimates is likely to be stronger for other sub-Saharan African countries with similar contexts of underdeveloped national labor markets and low international migration costs, and particularly for those countries where labor movements are often forced and happen as a result of violence and conflict. It is important to underline that our estimates are obtained in the context of one of the poorest African countries, where there are few wage earning occupations to be taken as an alternative to self-employment or subsistence agriculture. These characteristics of the setting of our study imply that the entrepreneurial gains from return migration are likely to be higher in our study than in settings with more alternatives to entrepreneurship.

Furthermore, the Mozambican context is one where there is a widespread tradition of international migration, especially to South Africa, where there is an important Mozambican diaspora that is likely to diminish migration costs. If migrant self-selection is not taken into account, these low migration costs are likely to result in negative self-selection patterns that tend to reduce 'naïve' estimates of the entrepreneurial gains from return migration.

In a business setting like the Mozambican, with few large employer firms, even basic entrepreneurship can be valuable as an economic diversification tool, especially if business owners have valuable foreign experience that may help them manage their businesses in a more effective way. Survival rates of microenterprises in Mozambique and elsewhere in sub-Saharan Africa are typically very low, and migratory experience may be a way to increase business survival, namely by providing effective business skills and by ensuring privileged access to foreign

markets through diaspora networks. Measuring how migration experience may affect the performance of business owners is outside the scope of our work, and is left for future research.

In light of our results, policymakers should feel encouraged to design temporary migration programs that can better serve the interests of both origin and destination countries. For example, lotteries granting migrant visas for temporary work permits could plausibly be proposed for sub-Saharan African countries that already see large numbers of migrants moving mostly informally to South Africa and the Middle East. Moreover, while further studies on other source countries are required to make a stronger general argument, for this case at least, keeping the doors of richer countries open to migration may be regarded as a form of “efficient aid”.

## References

- Abadie, Alberto, and Guido Imbens (2006). “Large Sample Properties of Matching Estimators for Average Treatment Effects”, *Econometrica*, 74 (1), 235-267.
- Akee, Randall (2010). “Who’s Leaving? Deciphering Immigrant Self-Selection From a Developing Country”, *Economic Development and Cultural Change*, 58 (2): 323-344.
- Batista, Catia (2008). “Why Doesn’t Labor Flow from Poor to Rich Countries? Micro Evidence from the European Integration Experience”, *Oxford Working Paper* 402.
- Batista, Catia, Aitor Lacuesta, and Pedro C. Vicente (2012). “Testing the ‘Brain Gain’ Hypothesis: Micro Evidence from Cape Verde”, *Journal of Development Economics*, 97 (1): 32-45.
- Batista, Catia, Julia Seither, and Pedro C. Vicente (2016). “Migration, Political Institutions, and Social Networks in Mozambique”, *NOVAFRICA Working Paper*.
- Batista, Catia, and Janis Umblijs (2014). “Migration, Risk Attitudes, and Entrepreneurship: Evidence from a Representative Immigrant Survey”, *IZA Journal of Migration*, volume 17 – 3.
- Batista, Catia, and Pedro C. Vicente (2011). “Do Migrants Improve Governance at Home? Evidence from a Voting Experiment”, *World Bank Economic Review*, 25 (1): 77-104.
- Beine, Michel, Frederic Docquier, and Hillel Rapoport (2008). “Brain drain and LDCs’ growth: winners and losers”, *Economic Journal*, 118:631-652.

- Beine, Michel, Cecily Oden-Defoort, and Frederic Docquier (2011). “A Panel Data Analysis of The Brain Gain”, *World Development*, 39 (4): 523-532.
- Bertoli, Simone, Jesus Fernandez-Huertas Moraga, and Francesc Ortega (2013). “Crossing the Border: Self-Selection, Earnings and Individual Migration Decisions”, *Journal of Development Economics*, 101 (1): 75-91.
- Borjas, George (1987). “Self-Selection and the Earnings of Immigrants”, *American Economic Review*, 77(4): 531-53.
- Borjas, George, and Bernt Bratsberg (1996). “Who Leaves? The Out-Migration of the Foreign-Born”, *The Review of Economics and Statistics*: 78, 165–176.
- Chiquiar, Daniel, and Gary Hanson (2005). “International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States”, *Journal of Political Economy*, 113(2): 239-281.
- Comissão Nacional de Eleições - Secretariado Técnico de Administração Eleitoral (2004). *Resultados Desagregados das Eleições Presidenciais*. Maputo, Moçambique: CNE-STAE.
- Coulon, Augustin, and Matloob Piracha (2005). “Self-Selection and the Performance of Return Migrants: The Source Country Perspective”, *Journal of Population Economics*, 18: 779–807.
- Docquier, Frédéric, Elisabetta Lodigiani, Hillel Rapoport and Maurice Schiff (2016). “Emigration and Democracy”, *Journal of Development Economics*, 120, 209-223.
- Dustmann, Christian, and Oliver Kirchkamp (2002). “Migration Duration and Activity Choice After Re-Migration”, *Journal of Development Economics*, 67: 351-372.
- Fairlie, Robert, and Christopher Woodruff (2007). “Mexican Entrepreneurship: A Comparison of Self-Employment in Mexico and the United States”, in George Borjas (ed.), *Mexican Immigration*, Chicago: University of Chicago Press and National Bureau of Economic Research.
- Fafchamps, Marcel, and Forhad Shilpi (2013). “Determinants of the Choice of Migration Destination”, *Oxford Bulletin of Economics and Statistics*, 75(3): 388-409.
- Fairlie, Robert, and Christopher Woodruff (2010). “Mexican-American Entrepreneurship”, *The BE Journal of Economic Analysis and Policy*, 10 (1): 1.
- Gallego, Juan, and Mariapia Mendola (2013). “Labour Migration and Social Networks Participation in Southern Mozambique”, *Economica*, 80: 721-759.
- Gibson, John, and David McKenzie (2014). “The Development Impact of a Best Practice Seasonal Worker Policy”, *Review of Economics and Statistics*, Forthcoming.
- Heckman, James, and Richard Robb (1985). “Alternative methods for evaluating the impact of interventions: An overview”, *Journal of Econometrics*, 30 (1-2): 239-267.
- Imbens, Guido, and Joshua D. Angrist (1994). “Identification and Estimation of Local Average Treatment Effects”, *Econometrica*, 62 (2): 467-475.
- Javorcik, Beata, Çağlar Özden, Mariana Spatareanu and Cristina Neagu (2011). “Migrant Networks and Foreign Direct Investment”, *Journal of Development Economics*, 94 (2): 231-241.

- Kugler, Maurice, and Hillel Rapoport (2007). "International Labor and Capital Flows: Complements or Substitutes?", *Economics Letters*, 92 (2): 155-162.
- McCormick, Barry, and Jackie Wahba (2001). "Overseas Work Experience, Savings and Entrepreneurship Amongst Return Migrants to LDCs", *The Scottish Journal of Political Economy*, 48(2): 164–178.
- McKenzie, David, John Gibson and Steven Stillman (2010). "How Important Is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration", *Journal of the European Economic Association*, 8(4): 913-945.
- Mesnard, Alice (2004). "Temporary migration and capital market imperfections", *Oxford Economic Papers*, 56: 242–262.
- Mesnard, Alice, and Martin Ravallion (2006). "The Wealth Effect on New Business Startups in a Developing Economy", *Economica*, 73: 367-392.
- Minnesota Population Center (2010). *IPUMS - Integrated Public Use Microdata Series, International: Version 6.0*. Minneapolis: University of Minnesota.
- Nyarko, Yaw (2011). "The Returns to the Brain Drain and Brain Circulation in Sub-Saharan Africa: Some Computations Using Data from Ghana", *NBER Working Paper 16813*.
- Piracha, Matloob, and Florin Vadean (2010). "Return Migration and Occupational Choice: Evidence from Albania", *World Development*, 38(8): 1141–1155.
- Rauch, James, and Vitor Trindade (2002). "Ethnic Chinese Networks in International Trade", *Review of Economics and Statistics*, 84: 116–130.
- Red Cross of Mozambique (2009). *International Federation's Disaster Relief Emergency Fund (DREF) Final Report*. Operation No. MDRMZ004.
- Rosenzweig, Mark (2007). "Higher education and international migration in Asia: brain circulation", in *ABCDE World Bank Conference*, ed. Francois Bourguignon. New York: Oxford University Press.
- Roy, A. D. (1951). "Some Thoughts on the Distribution of Earnings", *Oxford Economic Papers*, 3:135-146.
- Ruiz, Isabel, and Carlos Vargas-Silva (2015). "The Labor Market Impacts of Forced Migration", *American Economic Review*, 105(5): 581-86.
- Staiger, D., and J. Stock (1997). "Instrumental variables regression with weak instruments", *Econometrica*, 65(3): 557-86.
- Stock, James and Motohiro Yogo (2005). "Testing for Weak Instruments in Linear IV Regression", in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. Andrews and J. Stock, 80-108. Cambridge University Press.
- UNCTAD (2010): "Information Economy Report 2009: Trends and Outlook in Turbulent Times," Tech. rep., UNCTAD.
- Wahba, Jackie, and Yves Zenou (2012). "Out of Sight, Out of Mind: Migration, Entrepreneurship and Social Capital", *Regional Science and Urban Economics*, 42(5): 890–903.
- World Development Indicators (2009). Washington, DC: The World Bank.
- Yang, Dean (2008). "International Migration, Remittances, and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks", *The Economic Journal*, 118: 591-630.

## Tables and Figures

Table 1: Household Characteristics. All Households. Percentages.

	Migratory Experience
Households with at least one migrant	32.73
Households with at least one current migrant	15.77
Households with at least one return migrant	23.03
	Business Ownership
Households with at least one business	28.42
Households with at least one business-owning return migrant	3.97

Source: Own survey.

Table 2: Destination Countries of Return Migrants.  
Most Recent Migration Journey. Percentages.

---

South Africa	72.38
Tanzania	8.78
Malawi	6.64
Swaziland	2.57
Zimbabwe	2.36
Other African	1.07
Germany	1.93
Portugal	1.28
Other European	1.07
Cuba	1.07
Other	0.86

---

Number of Observations	467
---------------------------	-----

---

Source: Own survey.

Table 3: Summary Statistics. All Households.

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Business ownership, %	1,763	28.42	45.11	0	100
HH Head Age	1,747	37.62	13.58	15	88
HH Maximum Years of Schooling	1,760	6.62	4.08	0	19
HH Size, Total Number of Persons	1,762	5.87	2.85	0	29
HH Expenditure, MZN per day	1,674	128.83	164.03	0	2,381
Asset Ownership, %	1,762	7.78	26.79	0	100
Remittance Receipts, %	1,763	15.48	36.19	0	100
Province, Cabo Delgado, %	1,763	25.41	43.55	0	100
Province, Zambezia, %	1,763	24.96	43.29	0	100
Province, Maputo, %	1,763	24.9	43.26	0	100
Province, Gaza, %	1,763	24.73	43.16	0	100
Proximity to urban area $\leq$ 5Km, %	1,763	28.59	45.2	0	100

Source: Own survey.

Table 4: Probability of owning a business. All households. LPM estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Full Sample	War Migrants	War Migrants	War Migrants + Forced Return	War Migrants + Forced Return
Return migrant in household, dummy	0.125*** (0.047)	0.125*** (0.048)	0.138** (0.063)	0.141** (0.063)	0.240** (0.099)	0.243** (0.100)
Age of household head, years	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)
Gender of household head, female	-0.048** (0.022)	-0.046** (0.022)	-0.045* (0.023)	-0.046** (0.022)	-0.056** (0.022)	-0.058*** (0.022)
Max education in household, years	-0.008** (0.004)	-0.008** (0.004)	-0.009** (0.004)	-0.009** (0.004)	-0.008** (0.004)	-0.009** (0.004)
Household size, persons	0.007* (0.004)	0.007* (0.004)	0.006 (0.004)	0.005 (0.005)	0.006 (0.005)	0.005 (0.005)
Expenditure, MZN/day	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Asset ownership, dummy	0.159*** (0.049)	0.161*** (0.049)	0.191*** (0.053)	0.194*** (0.053)	0.182*** (0.063)	0.184*** (0.063)
Migrant destination: South Africa, dummy	-0.145** (0.062)	-0.139** (0.062)	-0.151* (0.089)	-0.155* (0.090)	-0.067 (0.204)	-0.056 (0.210)
Migrant destination: Other African, dummy	-0.422*** (0.071)	-0.426*** (0.072)	-0.437*** (0.094)	-0.440*** (0.094)	-0.495*** (0.107)	-0.501*** (0.107)
Current migrant in household, dummy		0.017 (0.035)		0.041 (0.039)		0.068 (0.044)
Remittances received, dummy		-0.056* (0.033)		-0.044 (0.045)		-0.089* (0.050)
Urban, Province Fixed Effects	Included	Included	Included	Included	Included	Included
Observations	1,658	1,658	1,416	1,416	1,305	1,305

Standard errors reported in parentheses, clustered at the enumeration area level. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

Table 5: Probability of owning a business. NMM estimates.

	(1)	(2)	(3)
	Full Sample	War Migrants	War Migrants + Forced Return
Return migrant - ATE	0.11** (0.06)	0.19** (0.09)	0.27* (0.15)
Observations	1149	1083	505

Note: All regressions use a matching estimator proposed by Abadie and Imbens (2006) and implemented in Stata through the nmatch command. Common support is established prior to each matching exercise by refining the sample after running logit regressions on restricted sub-samples. Controls are the same as in the fuller OLS specifications. Standard errors reported in parentheses.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

Table 6: Probability of owning a business. 2SLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	War Migrants				
Return migrant in household, dummy	0.967** (0.387)	0.177** (0.083)	0.464 (0.309)	0.213** (0.095)	0.239** (0.093)	0.217** (0.108)
Age of household head, years	-0.004*** (0.001)	-0.003*** (0.001)	-0.003** (0.001)	-0.003** (0.001)	-0.003** (0.001)	-0.003*** (0.001)
Gender of household head, female	-0.010 (0.029)	-0.045** (0.022)	-0.047* (0.025)	-0.052** (0.023)	-0.052** (0.023)	-0.042* (0.022)
Max education in household, years	-0.016*** (0.005)	-0.008** (0.004)	-0.011** (0.005)	-0.009** (0.004)	-0.010** (0.004)	-0.009** (0.004)
Household size, persons	0.005 (0.004)	0.007* (0.004)	0.001 (0.005)	0.002 (0.005)	0.002 (0.005)	0.005 (0.004)
Expenditure, MZN/day	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Asset ownership, dummy	0.168*** (0.054)	0.160*** (0.048)	0.146*** (0.052)	0.145*** (0.050)	0.145*** (0.050)	0.192*** (0.053)
Migrant destination: South Africa, dummy	-0.927** (0.362)	-0.194** (0.084)	-0.500* (0.291)	-0.262** (0.105)	-0.287*** (0.104)	-0.226* (0.120)
Migrant destination: Other African, dummy	-1.148*** (0.337)	-0.467*** (0.089)	-0.862*** (0.285)	-0.634*** (0.096)	-0.658*** (0.095)	-0.512*** (0.123)
Urban, Province Fixed Effects	Included	Included	Included	Included	Included	Included
IV Set	A	B	C	B + C	A + B + C	B
Kleibergen-Paap F-Statistic	14.77	13.7	12.52	11.25	11.03	10.51
Hansen J-Test (p-value)	-	-	-	0.58	0.282	-
Observations	1,658	1,658	1,304	1,304	1,304	1,416

IV Set A: Incidence of agricultural plagues in Mozambique.

IV Set B: Distance between residence in Mozambique and migrant destination.

IV Set C: Ratio between GDP per capita in Mozambique and in migrant destinations.

See text for detailed explanation on IV construction.

Standard errors reported in parentheses, clustered at the enumeration area level.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

## Appendix

Figure A1: Geographical coverage of household survey.



Sources: Diva, 2010  
and ArcGIS.

Table A1: Return migrants by decade of departure, shares.

	All return migrants	War emigrants	Non-war emigrants	Exogenous return migrants	Non-Exogenous return migrants
	%	%	%	%	%
Pre-1980	12.7	24.0	5.9	8.2	13.3
1980-1989	19.3	51.3	-	40.8	16.3
1990-1999	29.8	24.7	32.8	24.5	30.5
2000-2009	38.3	-	61.3	26.5	39.9
Number of observations	410	154	256	49	361

Table A2: Observable characteristics of return migrants by destination.

	All return migrants Mean	War migrants Mean	Non-war migrants Mean	Difference between war and non-war migrants Mean      P-Value		Exogenous return migrants Mean	Non-exogenous return migrants Mean	Difference between exogenous and non-exogenous returns Mean      P-Value	
<i>All migrants</i>									
Age on first migration, years	21.49	17.58	23.90	-6.31	0.00	18.45	21.93	-3.48	0.02
Migrant gender, female %	35.10	41.25	31.95	9.30	0.05	43.86	33.89	9.97	0.14
Education, years	6.36	6.97	6.04	0.93	0.01	6.43	6.36	0.07	0.89
Adjusted education, years	5.88	6.10	5.77	0.34	0.34	5.95	5.87	0.07	0.89
Observations	473	160	313			57	416		
<i>Migrants leaving for South Africa</i>									
Age on first migration, years	22.37	18.33	24.06	-5.73	0.00	21.88	22.42	-0.54	0.78
Migrant gender, female %	35.80	38.37	34.92	3.45	0.57	34.48	35.92	-1.44	0.88
Education, years	5.63	6.06	5.47	0.59	0.18	6.14	5.58	0.56	0.40
Adjusted education, years	5.28	5.35	5.26	0.10	0.82	5.97	5.21	0.75	0.24
Observations	338	86	252			29	309		
<i>Migrants leaving for other African destination</i>									
Age on first migration, years	19.60	16.24	23.71	-7.47	0.00	14.30	21.40	-7.09	0.01
Migrant gender, female %	38.00	47.27	26.67	20.61	0.03	56.00	32.00	24.00	0.03
Education, years	7.41	7.40	7.42	-0.02	0.98	6.67	7.66	-0.99	0.31
Adjusted education, years	6.71	6.46	7.00	-0.54	0.48	5.88	6.99	-1.11	0.21
Observations	100	55	45			25	75		
<i>Migrants leaving for non-African destination</i>									
Age on first migration, years	19.55	18.44	24.00	-5.56	0.03	28.00	19.11	8.89	.
Migrant gender, female %	17.24	29.41	0.00	29.41	0.04	0.00	18.52	-18.52	0.52
Education, years	10.57	10.00	11.45	-1.45	0.34	6.00	10.92	-4.92	0.08
Adjusted education, years	9.43	8.59	10.73	-2.14	0.15	4.50	9.81	-5.31	0.06
Observations	29	17	12			2	27		

Table A3: Type of businesses owned, distribution at household level, per cent

	Full sample Mean	Non-migrants Mean	Return migrants Mean	Difference between migrants and non-migrants	
				Mean	P-Value
Stores	9.18	7.67	13.82	6.15	0.04
Street vendors	46.91	46.03	49.59	3.56	0.49
Services	15.17	15.34	14.63	-0.71	0.85
Goods	7.78	6.88	10.57	3.69	0.19
Agricultural	33.73	36.24	26.02	-10.23	0.04

Note: "Goods" includes workshop; factory; commercial. "Services" includes restaurant; services; technical.