A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

Labor Productivity Gains and Spillover Effects from International Migration

Clara Pinto Neves da Conceição Bento, 24107

Work project carried out under the supervision of:

Professor Cátia Batista

January 3rd, 2020

Acknowledgements

I would like to express my sincere gratitude to Professor Cátia Batista for her guidance and helpful advice throughout the semester and for giving me the opportunity to work with the dataset. I would also like to thank my family and friends for their invaluable support and thoughtful suggestions. Thank you, with all my heart, to Z., my everyday companion.

Labor Productivity Gains and Spillover Effects from International Migration

Abstract

Can international migration increase labor productivity in the country of origin? Data from a detailed household survey conducted in Cape Verde was used to examine the effects of having current and/or return emigrants in the household and in the locality on the labor earnings of resident families. To account for migrant self-selection, instrumental variables were used, including migration networks and macro shocks in destination countries. The estimated effects are positive and significant for household- and locality-level variation, suggesting that families with migrants and those living in regions with a higher proportion of migrants tend to have increased labor income.

Keywords: migration, productivity, labor income, development impact of migration

1.1. Introduction and Motivation

Although the consequences of emigration and immigration in terms of labor market outcomes have been extensively studied in the literature – both for migrants and nationals, as well as for host and sending countries –, those of return migration are comparatively understudied. Temporary migration spells can encourage positive change in origin countries once emigrants return, as they usually bring back valuable resources and human capital accumulated during their stay abroad, including various skills and knowledge, training, formal education, principles and values, unobservable skill gains and financial savings. Therefore, it may be that return migrants are potentially more productive than individuals who did not migrate, which could result in a wage premium and, perhaps, in better job occupations upon returning. Furthermore, while abroad, emigrants usually send back financial and non-financial

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

remittances, which in turn can positively influence stayers' knowledge and competencies in various fields. As such, the interaction between returnees or emigrants and their families and communities back home may give rise to improved labor productivity within the household and in regions with a higher share of international migrants, that is, spillover effects over non-migrants.

Research on the consequences of emigration for the sending country has focused substantially on the 'brain drain' versus 'brain gain' debate. Whereas some authors[1] argue in favor of a 'brain drain' effect – through which the loss of the most skilled or educated individuals leads to negative effects on the origin country's economy –, others (see Batista, Lacuesta and Vicente, 2012) argue that the 'brain gain' resulting from the higher rate of human capital accumulation over the prospect of own future migration may compensate the hypothetical brain drain loss. Return migration can help compensate the possible 'brain drain' effect if returnees are able to take advantage of the knowledge, financial and social network resources acquired abroad to positively contribute to the development of their country of origin. Moreover, while abroad, emigrants might also improve the productive capabilities of their peers in the home country through knowledge transfers (for example on different management processes and organizational competences). Other studies using micro evidence – presented in the next section – specifically analyzed if there are 'returns to returning', namely with regard to productivity, one of the main drivers of economic growth. However, the results are contradictory. Indeed, depending on the context and on the countries under analysis, there may or may not be a wage premium resulting from international migration.

A key issue in the quest to obtain unbiased estimates when investigating the effects of migration on several outcomes from the source country's perspective is the self-selection of migrants, arising from the fact that these individuals tend to be different from non-migrants in

[1] Gruber and Scott (1966) and Bhagwati and Hamada (1974) developed considerably this theory's literature.

terms of unobserved characteristics. Selectivity may occur at both the outward and inward migration movements, with instrumental variable estimation widely used in the literature to induce exogenous variation and enable the proper identification of the regressors of interest[2].

This work project focuses on the impacts of international migration on the country of origin, specifically on the productivity gains stemming from it, as measured through migrants' labor earnings compared to those of non-migrants, at the household level. When examining productivity differentials, not only should one focus on the effect coming from return migration, but also on current migrants. In fact, the influence of migration networks can explain part of the effects, since they constitute an important mechanism through which emigrants transmit information and knowledge to their friends and family back home. Additionally, as aforementioned, it is relevant to understand if households living in regions with a higher share of international migrants – including both current and past migrants – present, on average, higher wages and better occupations – that is, jobs in more productive sectors. Hence, this study also assesses whether there are significant regional spillover effects resulting from return or current migration – and the two together – in Cape Verde, a migration-dependent country, following the identification and estimation strategies in Batista and Vicente (2011)[3], but considering a different main outcome: households' labor income.

1.2. Literature Review: Migration, Selection and Labor Productivity

Analyzing detailed data on current, return and non- migrants in Cape Verde, Batista et al. (2012) report a 4 p.p. 'brain gain' effect when increasing the likelihood of own migration in the future by 10 p.p., contributing positively to the discussion about the consequences of emigration and potential return. In fact, human capital theory predicts that investing in

[2] Examples of this can be found in the Literature Review section.
[3] Batista and Vicente (2011) study the impact of international emigration on the demand for political accountability at the locality-level in Cape Verde. They examine, respectively, the direct and indirect influence of return and current migrants and find positive and robust effects. The present work project is mainly based on the empirical strategy and econometric framework used in this paper.

education is expected to increase productivity levels in the future. Regarding the estimation of causal effects from migration, the authors point to the importance of taking educational upgrading after migration into account upon estimation to guarantee unbiasedness. In a different setting, considering Egyptian households, Wahba (2015) provides evidence on the need for controlling for distinct selectivity biases – i.e., emerging from the choices of emigration, return, labor force participation, and occupation following return –, and finds negative and positive selection[4] regarding return and emigration decisions, respectively. The author documents a wage premium of 16% for returnees when compared to non-migrants, although one quarter of this gap is explained by individual unobservable characteristics.

Concerning the areas of political institutions and governance in sending countries, some studies have shed light on the importance of networks and resulting knowledge transfer between migrants and stayers, for instance in Mozambique and in Cape Verde (Batista et al., 2019; Batista and Vicente, 2011). On the one hand, migrants contribute to rise demand for political accountability at home, especially those who lived in countries with better governance; on the other hand, political participation was proved to be enhanced through peer effects arising from social networks of migrants and non-migrants. These findings corroborate the potential impact of international migration on the sending regions, particularly pointing towards the existence of spillover effects on individuals who did not migrate. Adding to the topic of spillovers, Hausmann and Nedelkoska (2018) examined the impact of Albanian migrants who returned from Greece on the wages of non-migrants and found that low-skilled individuals tend to face increases in earnings, whereas no significant effect was found for skilled workers. Notwithstanding, the authors conclude that both the labor market participation and the employment chances of stayers tend to improve with return migration.

[4] For a detailed explanation about the meaning of positive/negative selection bias, please refer to "Return Migration, Self-selection and Entrepreneurship", by Batista et al. (2017).

The question of whether returnees are rewarded for the experience, training or education obtained abroad in the form of higher earnings has been thoroughly addressed in the literature. Return migrants tend to be more entrepreneurial than comparable non-migrants[5] and to find occupations in higher paying sectors upon return (Ammassari, 2004; Lacuesta, 2006; Barrett and Goggin, 2010; Mahuteau and Tani, 2011; Wahba and Zenou, 2012). The higher share of business creation may reflect managerial and organizational skills obtained during the migration spell; likewise, occupational mobility can be an indicator of increased productivity. For instance, Wahba (2015) finds that returnees have a higher probability of working in skilled occupations, as 56% of this group faced an upward occupational change relative to their premigration job. Concerning labor earnings, migration data from the capital cities of seven West African countries presented wage increases for return migrants who have been in OECD countries and higher productivity efficiency for those who become entrepreneurs (De Vreyer et al., 2010).

There is mixed evidence regarding the mechanisms underlying the previous wage gap, as with respect to the role of education. Positive wage changes exist for the different education levels (Barrett and Goggin, 2010), although they may differ depending on the returnee's educational attainment (Wahba, 2015). Human capital accumulated abroad seems like a logical contributor to the wage difference, but the evidence is contradictory: studying a temporary migration programme for unskilled workers in New Zealand, Gibson and McKenzie (2014) document that, on average, returnees accumulate human capital during the time spent abroad; however, Lacuesta (2006) argues that the higher wages of individuals who return to Mexico from the US are caused by pre-migration productivity disparities[6].

[5] Note that the group of non-migrants with similar observable characteristics (such as age, education, labor market experience, etc.) represents the counterfactual of the empirical estimation in the mentioned studies.
[6] Looking at the duration of migration spells, the author concludes that the wage gap of return migrants versus non-migrants is observed for short periods abroad, which didn't allow the time for human capital accumulation.

The inherent problem of migrants' selectivity in the absence of randomized controlled trials demands econometric analysis to account for self-selection both at the initial outward and final inward migration stages, on observable and unobservable characteristics. To mitigate this endogeneity concern, research on both return and current emigration resorts to instrumental variable estimation; economic conditions in destination countries reflected in macroeconomic indicators are commonly used, as well as proxies for migration networks[7].

This work project builds on previous literature regarding the impacts of international migration from the sending country's perspective, in terms of labor market outcomes. Due to its migration-dependent setting and large dataset from a detailed household survey, Cape Verde provides an interesting country for the purpose of this study. The main contribution of the analysis resides in the estimation of spillover effects on non-migrants with respect to productivity, a relatively understudied matter which can have important policy implications.

1.3. Cape Verde: Country Context

Cape Verde is an island country located in Africa, off the west coast, composed by 10 islands with 520,500 inhabitants living in 9 of them. The country became independent from Portugal in 1975 and presents a stable democratic system since 1991. In economic terms, it is considered by the World Bank a "lower-middle income" economy, and it grew at a rate of around 5.5% per year between 1991 and 2014 (a higher rate than the sub-Saharan Africa's average growth rate of 4.5%)[8]. In 2017, it registered a GDP per capita (PPP adjusted) of 7,027$, and an exports rate of 46% of GDP. Considering labor market performance, nearly half of the country's labor force works in the service sector, which represents about two thirds of the economy (values of 2014). The unemployment rate reached 12% in 2014 and has been rather

[7] Examples of this kind of instruments can be found in Acosta (2006), Beaudouin (2005), McKenzie and Rapoport (2006), Batista et al. (2012), Batista and Vicente (2011), Batista and Cestari (2016).

[8] Data from the UN report: "Common Country Analysis: Cabo Verde" (2017). GDP growth decelerated during 2009-2015 due to the global financial crisis, but it started recovering in 2016 (World Bank, March 2019).

stable ever since; according to the ILO, the total unemployment as a percentage of the total labor force in Cape Verde during the period of 2000-2006 was also, on average, about 11.5%.

Repeatedly facing episodes of droughts and famines, and being a country that is scarce on natural resources, Cape Verde has had a history of massive emigration for several decades; it was estimated that in 2012 there were over 100,000 Cape Verdean nationals living abroad, almost 20% of the population (Batista et al., 2012). According to the World Bank, net migration in 2017 (5-year period estimates) was -6,709, whereas in 2007 this indicator equals -11,341, almost twice the previous value. Consequently, it is not surprising that the value of remittances received in the country as a percentage of GDP is among the highest in sub-Saharan Africa (around 12% in both 2017 and 2006). Furthermore, remittances have been exceeding foreign direct investment for decades (Batista and Vicente, 2011), highlighting its importance for the country. Noticeably, the relevance of studying migration's impact in the context of Cape Verde's economic outlook is still very much verified: emigration is highly present, as there are more citizens living abroad than in Cape Verde itself, and the labor market and economic situations are not very different from those observed in the early 00's, when data was collected.

2.1. Data: Household Survey and Descriptive Statistics

The present research is based on a survey covering 1,066 resident households in 5% of the 561 census areas in Cape Verde, composing a representative sample. It was conducted from December 2005 to March 2006 by Cátia Batista and Pedro Vicente, based at the CSAE at the University of Oxford. The data collected[9] includes information about resident nonmigrants, return migrants and also current emigrants: the questionnaire presented detailed questions about the migration characteristics and full migration history of each household, as well as socio-demographic characteristics of all individuals belonging to the household[10], the family's

[9] The respondent was at least 30 years old, and provided information on every member of the household.
[10] In total, the dataset has over 7,000 individual observations.

economic situation – living standards, income earned, remittances received, etc. Since the goal of the empirical analysis is to study labor productivity, observations without information on labor income (the proxy for productivity) were not included in the analysis; hence, the sample used had at most 1,360 individual observations. This procedure creates a sample selection problem that will be discussed in the Final Remarks.

In Table A1 (in Appendix) one can consult relevant descriptive statistics regarding migratory experience, remittances, mean labor earnings and occupations of households. More than 40% of the surveyed families have a migration experience, with at least one international current or return migrant. Concerning the outcome of interest, the mean labor income per active individual (age between 15 and 64) of households with returnees is higher than in those with migrants, which in turn is considerably higher than that of households without migrants.

Table 1.1. Characteristics of Households and Households' Heads (Means)

	HH with migrant	HH with returnee	HH without migrants
Characteristics of Households			
Sample size	573	228	787
Maximum years of education	11.85	10.83	10.18
Member with foreign studies	0.05	0.11	-
Number of working-age individuals	2.74	2.52	1.38
Household size	4.55	4.61	5.62
Daily HH average expenditure	1125	1241	927
Number of children (dependents)	1.7	1.9	2.9
Area of residence (%)			
Urban	43.80	41.67	43.07
South	69.98	75.44	68.49
Characteristics of Heads			
Male (%)	43.26	54.43	37.35
Age	55.8	53.5	41.5
Married (%)	42.70	50.63	16.36
Years of Education	4.29	5.08	5.46
Job Experience (%)			
Only one job	66.48	64.94	75.79
Several safe jobs	8.52	12.99	7.86
Several part-time jobs	1.70	2.60	3.77
Unemployed	11.93	9.09	11.95
Retired	11.36	10.39	0.63

Regarding household heads' occupations, around 20% work in the sectors of health, education or public administration, which are considered the most prestigious and better-paid in Cape

Verde. Only 11% of the individuals included in the sample are outside the age interval previously mentioned and around 43% live in urban areas. The mean of the maximum years of education completed in the family varies between 10 and 11 years across households with and without migrants.

Table A2 (in Appendix) shows the most relevant destinations of returnees and current migrants. Considering the whole dataset, most individuals lived in Portugal (45% and 48%, respectively) or in the United States (21% and 17%, respectively), and a significant share emigrated to European countries[11].

2.2. Econometric Framework and Empirical Strategy: Direct and Spillover Effects

In this section, the empirical strategy underlying the estimation of the impact of international migration[12] on Cape Verdean households' labor productivity will be explored, together with the econometric model specifications and identification strategy. The analysis will be divided in two: firstly, the impact on labor earnings of having or not migrants in the family will be assessed by comparing households with and without migrants; secondly, the potential spillover effects at the regional level will be estimated, considering all households that live in a given region and assessing how the proportion of migrants relative to residents within that locality affects the labor market performance of those families. The second analysis aims at investigating if there are spillover effects coming from a higher share of households with migrants residing in the region, or in other words, if we observe improved labor earnings in regions with a higher share of return and/or current migrants.

The direct effect on own labor productivity of having or not a migrant in the family must be isolated from the indirect effect of having households with migrants in the region. Therefore,

[11] For precise numbers, please refer to the table in Appendix.
[12] Locality or household level variation of remittances will almost fully overlap that of migration, since these two go side by side. Thus, when measuring the effects of international migration, we are actually considering those of both migration and remittances together, not disentangling the effect of these two variables. We are estimating the joint effect of migration and remittances, and for that reason the latter won't be included in any regression.

two separate regressions were estimated: one measures the household-level effect and the other the overall effect at the regional level. The reason why we consider two distinct econometric specifications is because the correlation between both variables of interest – dummy for whether or not the household has a migrant and percentage of households with migrants per region – is expected to be strong, since the individual decision to emigrate is highly correlated with the proportion of migrants in the region of residence, due to network effects[13].

The spillover effect can be explained by information and knowledge transmission from migrants to non-migrants, influencing the residents' labor market choices either through a more frequent, face-to-face, contact (the case of return migrants) or a rather indirect contact (the case of current migrants). Some previous studies estimated a similar effect; for instance, Batista et al. (2019) investigate whether migrant social networks have an impact on political participation in Mozambique, focusing on the existing links between an individual who is living in the country and the migrants in her network. Batista and Vicente (2011) look into the role of international migration in increasing the demand for better governance in Cape Verde, considering a proxy for the number of times a local is expected to be in contact with migrants (either returnees or current migrants) or with their peers in her locality: the percentage of migrants among all individuals within the household's area. As such, to account for the possible influence, on the outcomes of interest, of both returnees and emigrants whom remain in contact with the ones left behind, firstly, the joint effect of return and current migration will be tested; afterwards, the influence of return versus that of current emigrants will be estimated separately.

2.2.1. Econometric Specification(s)

The direct effect is estimated for international migration as a whole (current and return migrants) and for returnees and emigrants separately, respectively, according to the models:

[13] Network effects are usually a strong predictor of migration in a given locality.

$$ln(L_i) = \beta_0 + \beta_1 intmig_i + \beta_2 X_i + u_i \quad (1)$$

$$ln(L_i) = \beta_0 + \beta_1 intret_i + \beta_2 intpres_i + \beta_3 X_i + u_i \quad (2)$$

where L_i is the outcome of interest; a proxy for labor productivity at the household-level, *household's labor earnings*, measured by the mean labor income per active individual in the family. With this variable, one intends to obtain an aggregate measure of productivity for the family. $intmig_i$ is a dummy variable equal to 1 if the household has at least one migrant (current or return), $intret_i$ equals 1 for households with at least one return migrant, and $intpres_i$ for households with at least one current migrant. X_i is a vector of controls[14] including individual, household and regional characteristics that are relevant in explaining the dependent variable, namely household heads' characteristics such as age, gender, education, number of dependents, previous labor market experience; at the household and regional levels it comprises household size, maximum education attained, local unemployment rate, average expenditure per capita, urban and south dummies. This vector also includes region fixed effects to control for the effect of time-invariant regional characteristics.

To evaluate whether there is a spillover effect of international migrants on resident households without migrants, the following empirical question will be addressed: "*does the proportion of international migrants within the household's area of residence influence its labor market performance?*". The answer will be positive if we find a statistically significant impact of the explanatory variable of interest "proportion of international migrants within each region" on the dependent variable "labor productivity of households in each region", following the model:

$$ln(L_i) = \beta_0 + \beta_1 PM_i + \beta_2 X_i + u_i \quad (3)$$

in which PM_i equals the number of households with at least one international migrant – current or return – in each of the 30 enumeration areas (EAs, i.e., neighbourhoods or villages) divided

[14] The choice was based on Wahba (2015), Batista et al. (2017), Batista and Vicente (2011), Batista et al. (2019).

by the total number of households per EA, X_i is the vector of control variables described previously but without region fixed effects, and u_i is the disturbance term.

Following the same reasoning as for the direct impact, the regional level model for the separate influence of return migration and current migration is the following:

$$ln(L_i) = \beta_0 + \beta_1 PR_i + \beta_2 PP_i + \beta_3 X_i + u_i \quad (4)$$

where PR_i and PP_i represent, respectively, the proportion of families within each EA with at least one international return migrant and with at least one international current migrant.

Outcome Variable

Productivity will be proxied through wages, based on the well-known economic mechanism, $[MPL = w/P]$ – marginal productivity of labor equals real wages, in equilibrium. Although this productivity measure presents several caveats, namely the fact that there are many market imperfections which break the equality, there is no available data on the production side (i.e., companies or microenterprises[15]). For the former reason, and because other researchers in development economics treat changes in wages as productivity changes, mean labor income per working-age individual at the household level is the main proxy for labor productivity in our analysis.

2.2.2. Identification Strategy: OLS and Instrumental Variables (IVs) Estimation

Relying on regional-level variation reduces endogeneity concerns related with self-selection coming from unobservable factors. The example most commonly mentioned in the literature has to do with unobserved ability or personal drive, which may positively influence the migration (and return) decisions of an individual, while also affecting positively her labor earnings or occupational choice. Indeed, it is likely that unobserved individual characteristics are correlated within the same household and perhaps even across families living in the same

[15] De Vreyer et. al (2010) estimate microenterprise production functions in an attempt to quantify past migration's productivity gains of becoming an entrepreneur, comparing returnees with non-migrants.

locality, but not across regions, so that focusing on regional variation allows to average out unobserved heterogeneity (Batista and Vicente, 2011). Also, regarding the proportion of migrant-households in each region, and adding to the reasoning concerning unobservable characteristics, it is also reasonable to state that regional-level omitted variables are unlikely to be a threat to external validity, since Cape Verde is a homogeneous country that is rather small[16].

On the other hand, for the case of household-level variation, estimating through a simple OLS regression the impact of migration on a certain outcome poses selectivity issues. These may concern both decision stages, resulting in biased estimates, unless we are in the presence of a randomized field experiment that randomly assigns individuals to migrate and, then, from the pool of emigrants, randomly chooses some to return[17]. Moreover, there may be problems of reverse causality if, for instance, low levels of labor productivity lead to the decision of temporarily migrating to countries with better labor market prospects or, from a different perspective, lead to the impossibility of emigration due to its high associated costs if an individual does not have other sources of income. Since it is not possible to simultaneously observe the outcome variable for a household with a migrant and for the same family had this individual not migrated, i.e., the counterfactual, one needs to estimate the average difference in outcomes between those who did and did not emigrate by getting rid of the selection bias. Thus, there is the need to resort to Instrumental Variable estimation, using the two-stage least-squares (2SLS) method.

Bearing in mind the previous explanation, IV estimation is used to estimate the direct effect of having an international migrant in the family on own labor earnings, while regressions (3) and (4) are estimated by OLS. A good IV corresponds to a variable such that changes in the

[16] As stated in Batista and Vicente (2011), this fact "rules out the most obvious (potentially omitted) factors that could promote migration and (the outcome variable) simultaneously at the locality level."

[17] In this case, there would be no selection effect by construction: $R_i \perp (L_{1i}, L_{0i})$, where L_{1i} represents the labor market outcomes of an individual who migrated and returned ($R_i = 1$) and L_{0i} of one who did not migrate ($R_i = 0$).

instrument are associated with changes in the potentially endogenous variable but do not lead directly to changes in the dependent variable, allowing for exogenous variation of the migration decision. As such, it needs to obey two conditions: it must be exogenous, meaning that it can only influence y through x (the variable of interest) – i.e., it needs to be uncorrelated with the model's error term u_i; and it needs to be relevant, i.e., strongly correlated with x. Then, the instruments allow for the identification of the causal effect of interest, by only influencing labor productivity through the migration experience channel.

Since international migration presents two potential selection moments, the outward and inward stages, the selected IVs must be correlated with the decisions to emigrate to the foreign country and to return to the home country. Furthermore, the IVs need to induce enough variation to identify the effects of interest at the household level. Accordingly, three instruments are proposed: migrant networks, proxied by lagged stocks of emigrants in each EA, and economic conditions in main destination countries[18], measured by nominal GDP per capita or unemployment rates.

2.2.3. Choice of IVs

Many authors have investigated the role of migrant networks in developing countries (see, for example, McKenzie and Rapoport, 2010; Wahba and Zenou, 2005). Arguably, having friends, family or peers in a foreign country should improve the amount of information gathered about living there, such as living conditions and work prospects, and may facilitate the integration period and adaptation to the new culture. Consequently, migration networks decrease the costs and risks associated with emigrating, making network effects an important outward movement's determinant. However, the fulfilment of the first condition for instrument validity described before is not so straightforward, since a network's formation may be related

[18] Portugal, USA, France, Netherlands and Luxembourg.

with community level characteristics that may also be affecting the dependent variable. Nevertheless, as Cape Verde presents a long-standing migration tradition, these migration networks were established in the past, and were mainly motivated by natural phenomena such as droughts. Accordingly, due to their historical dimension and the exogeneity of these shocks, one may argue that they do not influence individuals' labor earnings in the survey's year, except through the migration channel. Hence, following Batista and Vicente (2011), the instrument for the outward migration stage comprises five-year lagged regional migration stocks[19], proxying for migration networks, since by definition a network is something that was formed in the past. That is, an individual's acquaintances who emigrated previously are going to be the ones influencing her decision to leave in a given moment. To induce variation at the household level, the following was done: for households with migrants, this variable's value was calculated for the year of the household's first emigration; for households without any international migration experience, the value was computed for the year in which the head reaches the average emigrating age in the sample, 24 years old.

The second IV set is based on Batista et al. (2012) and includes one out of two variables, or both, depending on the regression specification: nominal GDP per capita can be thought as an indicator of the average wage in the host country as well as of its standard of living; unemployment rates represent labor market conditions and easiness of finding a job abroad. These aspects strongly influence an emigrant's decision to stay in the foreign country. Both macroeconomic indicators aim to represent bilateral shocks between Cape Verde and destination countries and are therefore expected to play an important role in the decision to return to the home country. Besides, it is legitimate to state that they are not related with labor income earned in Cape Verde, due to their macroeconomic nature residing in a foreign country. Variation at the household level is achieved by matching the values of GDP or unemployment

[19] Computed as the number of emigrants in each EA relative to the number of residents in that same EA, in each year.

in each destination country with the year in which the household head turns 36 (the average age of returning in the sample), assuming that this would be the year in which an emigrant would be more prone to return. Since the two variables are expected to be highly correlated, each of them was tested separately as an instrument for return migration.

3. Results

This section describes the estimation results of the empirical models. Firstly, the main variables of interest in each specification will be examined and then an overall analysis of other interesting results will be made. Notice that since all regressions present the dependent variable in logarithm, interpretations will be made in percentage terms.

Table 1 reports the results for the regressions estimated by OLS for the effect of having an international migrant in the household (current or return) on the outcome of interest and of the effect of each migration channel separately, i.e., having at least one return emigrant in the family versus having at least one current emigrant in the family. One can see from the coefficients in columns (3) and (6) that households with either returnees or current emigrants or both have, on average, higher mean labor income per active individual than those without, keeping everything else constant. Specifically, the aggregate effect (see Table 1) of having at least one international migrant in the family is of 19.5%; in other words, when the household has at least one emigrant/returnee, then mean labor income per active individual is expected to be 19.5% higher, *ceteris paribus*. Furthermore, the effect coming from current migration seems to be the most relevant in explaining the outcome variable in all specifications, as reported by the coefficients' magnitude in the last three specifications – when considering the emigrants' and returnees' individual influence, we obtain a coefficient estimate of 0.22 for the first and of 0.16 for the latter (see regression 6). The inclusion of regional fixed effects to account for the influence of time-invariant regional characteristics only changes the results slightly, decreasing the main coefficient's values.

Table 1: Direct Effects – Impact of Having International Migrants (current and return) in the Household on Labor Productivity and Separated Effects from Return Migrants and Current Emigrants

Mean labor income per active individual in HH	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
International emigrant(s) (current or return) in HH	0.4504***	0.2347***	0.195***			
	(0.0733)	(0.0593)	(0.066)			
International return emigrant(s) in HH				0.2548**	0.1767*	0.1654*
				(0.1031)	(0.0935)	(0.0956)
International current emigrant(s) in HH				0.4227***	0.2466***	0.2225***
				(0.0788)	(0.0734)	(0.0797)
Age		0.0074***	0.00742**		0.0062**	0.00604**
		(0.0024)	(0.0028)		(0.0025)	(0.0029)
Gender		0.2718***	0.2438***		0.2618***	0.2348***
		(0.0415)	(0.0396)		(0.0411)	(0.0394)
Schooling		0.0945***	0.0926***		0.0937***	0.0916***
		(0.0128)	(0.0131)		(0.0127)	(0.0128)
Maximum education		0.0318**	0.0351***		0.0309**	0.0337***
		(0.0116)	(0.0114)		(0.0117)	(0.012)
Number of children		-0.0538**	-0.0642**		-0.0603***	-0.0708***
		(0.0217)	(0.0257)		(0.0218)	(0.0254)
Household size		0.0222	0.0307		0.0286	0.0368
		(0.0284)	(0.0312)		(0.0276)	(0.0302)
Previous labor market experience		0.1549*	0.1218		0.161*	0.126
		(0.0848)	(0.0845)		(0.0814)	(0.0824)
Urban area		0.1282	0.2134***		0.1404	0.1822**
		(0.1010)	(0.07)		(0.102)	(0.0729)
Southern islands (Santiago and Fogo)		0.2160**	0.3934***		0.2065*	0.4063***
		(0.1036)	(0.0288)		(0.106)	(0.0289)
Regional average unemployment rate		0.9321	2.5656***		0.9355	2.545***
		(1.1272)	(0.2375)		(1.128)	(0.226)
Regional average household expenditure		1.5160	1.2979**		1.4139	1.415**
		(1.2934)	(0.4787)		(1.315)	(0.521)
Regional Fixed Effects	NO	NO	YES	NO	NO	YES
Constant	5.3244***	3.4487***	3.1865***	5.3221***	3.5028***	3.226***
	(0.0597)	(0.3306)	(0.1318)	(0.0578)	(0.339)	(0.134)
Observations	1,297	1,195	1,195	1,297	1,195	1,195
Adjusted R^2	0.0744	0.288	0.328	0.087	0.297	0.337
F	37.74	68.48	.	23.90	67.16	.

* $p < .1$, ** $p < .05$, *** $p < .01$. Robust standard errors in parentheses, clustered at the EA level[20].

[20] The standard errors' clustering procedure was done to account for some potential correlation at the regional level.

All the described results are significant at 1 percent significance level, except for the return migration's coefficient.

To avoid endogeneity biases related to self-selection of migrants, as discussed in the Econometric Framework's section, regression model (1) was estimated using the 2SLS approach. For this purpose, we used instrumental variables for the initial outward migration and the final inward return migration stages. Table 2 shows the results using each instrument separately, to test for their strength, as well as three combinations between the IV for the exiting movement with the two proposed IVs for return migration.

It is important to highlight the small difference in the coefficient of interest's estimates between specifications (1), (4), (5) and (6), which depict the regression with the migration networks' instrument (IV A) alone and regressions with this variable together with either the GDP per capita IV set (IV set B), the unemployment rates IV set (IV set C), or the two together. Moreover, when comparing these point estimates with the ones obtained with OLS, all of them more than doubled in absolute terms. For instance, looking at the 2SLS regression specification which includes both the lagged regional migration stocks and the GDP per capita values in the main host countries, one can see that families with at least one migrant tend to have an average mean labor income per active individual 49% higher than families without migrants, whereas when analyzing the specification with IV set C and IV A this value equals 42%. The previous results are significant at 1 percent level. Therefore, the overall positive effect of migration (current and return) on Cape Verdean households' labor productivity increases when one accounts for the unobservable self-selection in the outward and inwards phases, and that leads us to conclude that there is negative unobservable migration self-selection.

Regarding the validity of the chosen IVs, instrument relevance must be tested through the F-stat in the first-stage regression, to infer about the correlation between the IVs and the endogenous regressor.

Table 2: Direct Effect – Impact of Having International Migrants (current and return) in the Household on Labor Productivity (2SLS, 2nd stage regression reported)

Mean labor income per active individual in the HH	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
International emigrant(s) (current or return) in HH	0.408**	1.068**	0.606**	0.493***	0.420***	0.460***
	(0.196)	(0.421)	(0.280)	(0.149)	(0.146)	(0.144)
Age	-0.00690*	-0.0272**	-0.0158**	-0.0130**	-0.0119**	-0.0127**
	(0.00371)	(0.0110)	(0.00791)	(0.00534)	(0.00574)	(0.00575)
Gender	0.228***	0.144**	0.167**	0.182***	0.179***	0.176***
	(0.0510)	(0.0721)	(0.0686)	(0.0627)	(0.0665)	(0.0666)
Schooling	0.154***	0.121***	0.124***	0.117***	0.122***	0.123***
	(0.0153)	(0.0199)	(0.0198)	(0.0180)	(0.0192)	(0.0193)
Maximum education	-0.00514	-0.0125	0.000357	0.00567	0.00557	0.00443
	(0.0101)	(0.0169)	(0.0136)	(0.0114)	(0.0120)	(0.0120)
Number of children	-0.0706***	-0.114***	-0.0961***	-0.0985***	-0.0907***	-0.0919***
	(0.0251)	(0.0349)	(0.0335)	(0.0303)	(0.0321)	(0.0322)
Household size	-0.0121	0.0549	0.0333	0.0257	0.0237	0.0258
	(0.0222)	(0.0377)	(0.0332)	(0.0303)	(0.0323)	(0.0326)
Previous labor market experience	0.188**	0.377***	0.382***	0.356***	0.376***	0.378***
	(0.0776)	(0.104)	(0.107)	(0.0971)	(0.106)	(0.106)
Urban area	0.365*	0.0388	0.144	0.354	0.215	0.200
	(0.220)	(0.327)	(0.288)	(0.236)	(0.259)	(0.259)
Southern islands (Santiago and Fogo)	-0.155	-0.0626	0.0091	0.126	0.0702	0.0568
	(0.281)	(0.412)	(0.378)	(0.328)	(0.351)	(0.354)
Regional average unemployment rate	0.388	4.043	3.631	3.299	3.438	3.480
	(1.974)	(2.919)	(2.697)	(2.427)	(2.571)	(2.595)
Regional average household expenditure	2.820	7.581**	5.241*	3.981*	4.274*	4.486**
	(1.781)	(3.345)	(2.746)	(2.045)	(2.195)	(2.190)
Regional Fixed Effects	YES	YES	YES	YES	YES	YES
IV set	A	B	C	A+B	A+C	A+B+C
Observations	1,247	734	671	734	671	671
F-stat of 1st stage	125.51	9.25	9.63	30.55	32.68	20.64
Kleibergen-Paap F-Stat	125.51	9.25	9.63	30.55	32.68	20.64
Hansen J (P-value)	.	0.37	0.02	0.153	0.03	0.13

* $p < .1$, ** $p < .05$, *** $p < .01$. Robust standard errors in parentheses.

Regional fixed effects and constant were partialled-out in all specifications. IV set A: migration networks; IV set B: nominal GDP pc in main destination countries; IV set C: nominal unemployment rates in main destination countries. The three IV sets are described in detail in the main text.

Additionally, and since there are several specifications with more than one instrument, the overidentifying restrictions' test should also be examined to understand if the exogeneity condition for instrument validity is supported by the data. The reported 1st stage F-stat of

regressions (1), (4), (5) and (6) is comfortably above 10, thus consistent with the rule of thumb F>10. However, this same statistic in the regressions considering only one IV for return migration, either the GDP or the unemployment, is slightly below 10 (values 9.25 and 9.63, respectively). This shouldn't constitute much of a problem since the values are very close to 10, ruling out any serious weak instrument's problem.

Another test was performed to see whether the instruments are not just weakly correlated with the endogenous variable, which would cause the 2SLS estimator to be biased towards OLS'. Hence, considering the models (4) and (6), the most suitable ones, and looking at the reported Kleibergen-Paap F-stats – 30.55 and 20.64 –, we can conclude that they are both greater than the Stock and Yogo's critical values for a 20% maximal IV size (11.72 and 15.56), and in model (4) also for a 10% size (29.18). Accordingly, when jointly used, the IVs considered in these two model specifications are indeed relevant. Lastly, the Hansen J statistic for the overidentification test of all instruments has a p-value of 0.153 and 0.13, respectively, for the two models. Thus, the null hypothesis that the over-identifying restrictions are valid is not rejected, which supports the economic reasoning behind the instruments' exogeneity.

It was not possible to estimate the 2SLS regression of model (2) from the previous section, since the various sets of proposed IVs did not have enough explanatory power when applied to the potentially endogenous variables *intret* and *intpres*, that is, when estimating the household-level effects of current migration and return migration separately (not reported). Consequently, the above specifications estimated by OLS must be interpreted with caution, as they just point to a certain positive effect, since the estimates will be biased due to selectivity. It is important to note that instrument relevance was only an issue for one endogenous regressor, the return migrant(s) binary variable. A plausible explanation for this may lie in the inseparability of both effects and their high correlation, since the channels underlying the impact of having emigrants or returnees in the family on labor market performance are

essentially the same. Nevertheless, one must bear in mind that although we can conclude on the impact of having at least one migrant in the family, independently of her being currently away or having returned, it is not possible to take conclusions about whether this effect is coming from the individuals who live abroad or the ones who returned to Cape Verde.

The next table presents the estimation results for the impact of the proportion of households with international migrants within a given family's area of residence relative to the total number of resident-households on its labor market performance. As before, the second set of regressions, (4)-(6), distinguishes the effects of current and return migration so there will be two independent variables of interest, one for the effect of the proportion of families with returnees and the other for that of families with current emigrants. The aggregate effect is positive and statistically significant at 5 percent level. So, the overall effect – both the direct and the spillover effects together – coming from the percentage of households with migrants who live in the same region as a given household on its labor earnings seems to be strong and positive. Specifically, when the proportion of households with at least one emigrant/returnee increases by 1 p.p., then the expected mean labor income per active individual increases by 51.7%, on average, with all other factors held constant.

When separating the variable of interest in the two distinct regressors, the only statistically significant result is obtained for the percentage of resident-households with current emigrants (coefficient estimate equal to 0.52). Interestingly, this point estimate is almost exactly the same as the one for the aggregate variable, which means that current migration almost fully explains the outcome variable, in the particular case under analysis. Thus, the higher the percentage of families with emigrants in the locality, the higher the mean labor income per active individual of households in that region. This result is consistent with the reasoning explored in previous sections regarding knowledge transferability from emigrants to their peers back home during the period spent in the foreign country. Notwithstanding, Cape Verde is a

Table 3: Overall Effect – Impact of the Proportion of International Migrants (current and return) in the Region on the Labor Productivity of Households

Mean labor income per active individual in the HH	(1)	(2)	(3)	(4)	(5)	(6)
Percentage of resident-HHs with international migrants	0.5385***	0.5385	0.517**			
	(0.1861)	(0.401)	(0.244)			
Percentage of resident-HHs with return emigrants				0.4095	0.1491	0.1491
				(0.9251)	(0.3421)	(0.5728)
Percentage of resident-HHs with current emigrants				0.4125	0.5223**	0.5223
				(0.6430)	(0.2650)	(0.4296)
Age			0.0104***		0.01023***	0.0102***
			(0.00264)		(0.0015)	(0.0026)
Gender			0.275***		0.2750***	0.2750***
			(0.0408)		(0.0409)	(0.0410)
Schooling			0.0938***		0.094***	0.094***
			(0.0132)		(0.0121)	(0.0130)
Maximum education			0.0395***		0.0396***	0.0396***
			(0.0105)		(0.0065)	(0.0104)
Number of children			-0.0565**		-0.0551***	-0.0551**
			(0.0252)		(0.01994)	(0.0254)
Household size			0.0176		0.01689	0.01689
			(0.0310)		(0.01727)	(0.0312)
Previous labor market experience			0.148*		0.1479**	0.1479*
			(0.0822)		(0.06351)	(0.0857)
Urban area			0.243***		0.2499***	0.2499***
			(0.0768)		(0.0443)	(0.0864)
Southern islands (Santiago and Fogo)			0.252**		0.2534***	0.2534**
			(0.0986)		(0.0475)	(0.0995)
Regional average unemployment rate			0.464		0.5446	0.5446
			(0.986)		(0.4986)	(0.9961)
Constant	5.3109***	5.3109***	3.373***	5.3231***	3.3722***	3.3722***
	(0.0723)	(0.1669)	(0.226)	(0.1538)	(0.1739)	(0.2359)
Observations	1,297	1,297	1,195	1,297	1,195	1,195
Adjusted R^2	0.0052	0.0052	0.276	0.0043	0.2743	0.2743
F	8.37	1.80	85.98	1.02	43.11	80.20

Robust standard errors in parentheses, clustered at the EA level in regressions (2), (3), (4) and (6).
* $p < .1$, ** $p < .05$, *** $p < .01$

country with a high prevalence of migrants abroad and so it makes sense that they exert substantial influence over the labor earnings of resident households, through financial remittances that they can invest to increase labor productivity and through knowledge assimilation or skill gains that are passed on to them.

Focusing on other relevant variables included in the different model specifications, some interesting findings are worth mentioning. Residing in the south is positively related with the outcome of interest, as well as living in urban areas, which owes to the fact that southern (and urban) regions in Cape Verde are economically more active. Besides, being a male, years of schooling, having previous work experience, and maximum education attained in the household also contribute positively to the labor earnings of families in our sample. The number of children in the household is always negatively related with labor market performance, which is an intuitive result. A puzzling coefficient sign in the reported tables is the one of age in the 2SLS estimation, which turns out to be negative, contrarily to OLS regressions, although its magnitude is rather small. Finally, the difference in the number of observations on Table 2 has to do with the years considered for the macroeconomic variables (1980 and above) and the cases in which the head is younger than 36 or turns 36 after 2006.

4. Final Remarks

This work project contributes to the literature on the effects of international migration on the countries of origin of migrants. It departs from the hypothesis that labor earnings may depend not only on an individual's own decision to emigrate – and to return – but also on the migration experience of family members or of the people living in the same locality. To test this hypothesis, an instrumental variable approach was used to address both potential outward and inward self-selection effects inherent to migrating and returning decisions. The relevance of the proposed research questions is relatively undisputed, especially in the socio-economic context under analysis, since Cape Verde is a country with very high emigration rates.

The obtained results suggest that this relationship is positive and significant, both for the household-level analysis and for the regional overall effect previously explained in the main text. An interesting result that can be interpreted by comparing the OLS regressions' estimates with the ones presented by 2SLS estimations using different instrumental variables is the sign

of the self-selection bias. A negative selection bias occurs when, for instance, the individuals who do not present observable qualifications, as education, or work less hard are those who decide to move abroad (or also to return). This is in line with the results from the present analysis, in which OLS estimates underestimate the true labor productivity gains over resident households coming from migrants, current and return, reflected in the IV estimations. Related to these findings is the 'brain gain' theory empirically tested by Batista et al. (2012), that explores an important spillover effect from international migration: educational attainment of those who stay behind. In light of this argument, there is a substantial impact of current emigrants on the incentives for those who have stayed to get education. This work project can be regarded as complementary to the latter mechanism, as education and labor earnings usually go side by side.

Some relevant questions were left unanswered, namely what are the mechanics that explain the observed results (Batista and Vicente, 2011). Thus, further studies should focus on trying to identify the reasons underlying the suggested positive impact of international migration on labor productivity. Moreover, due to data limitations, it was not possible to study occupational mobility, which would imply having information on individuals' jobs before and after the migration experience, both for migrants and non-migrants. This would further complement the analysis regarding labor market outcomes, since the type of occupation a person has is usually very much related with her productivity level. Also, the present econometric analysis is not robust to sample selection, as labor income constitutes a selected sample. This is because some individuals choose to participate in the labor market whereas others decide not to, which makes it impossible to observe the wages of those who do not work. Finally, the proper identification of the distinct effects from return and current migration would be useful to help policy makers design incentive programs focused on specific objectives. Indeed, the results suggest that maintaining high levels of international emigration will improve

local labor earnings, but also that stimulating the return of nationals who decide to live abroad may induce labor productivity gains. Nonetheless, it remains open to debate whether it is the first or the latter which contributes most to the improvement of residents' labor productivity.

5. References

Ammassari, Savina. 2004. "From nation-building to entrepreneurship: the impact of élite return migrants in Côte d'Ivoire and Ghana." *Population, Space and Place,* 10(2): 133–154.

Barrett, Alan M. and Goggin, Jean. 2010. "Returning to the Question of a Wage Premium for Returning Migrants." IZA Discussion Paper No. 4736.

Batista, C. and Cestari, Francesco. 2016. "Migrant Intentions to Return: The Role of Migrant Social Networks." NOVAFRICA Working Paper 1602.

Batista, C. and Vicente, Pedro. 2011. "Do Migrants Improve Governance at Home? Evidence from a Voting Experiment." *The World Bank Economic Review*, 25(1): 77–104.

Batista, C., Lacuesta, A. and Vicente, Pedro. 2012. "Testing the 'brain gain' hypothesis: Micro evidence from Cape Verde." *Journal of Development Economics,* 97(1): 32–45.

Batista, C., McIndoe-Calder, T. and Vicente, Pedro. 2017. "Return Migration, Self-selection and Entrepreneurship." *Oxford Bulletin of Economics and Statistics,* 79(5): 797–821.

Batista, C., Seither, J. and Vicente, Pedro. 2019. "Do migrant social networks shape political attitudes and behavior at home?." *World Development*, 117: 328–343.

Bhagwati, J., Hamada, Koichi. 1974. "The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis." *Journal of Development Economics,* 1: 19–42.

De Vreyer, P., Gubert, F. and Robilliard, A. 2010. "Are There Returns to Migration Experience? An Empirical Analysis using Data on Return Migrants and Non-Migrants in West Africa." *Annals of Economics and Statistics*, 97: 307–328.

Gibson, J. and McKenzie, D. 2014. "Development through Seasonal Worker Programs: The Case of New Zealand's RSE Program." World Bank Policy Research Working Paper 6762.

Gruber, Herbert, Scott, Anthony. 1966. "The international flow of human capital." *The American Economic Review,* 56: 268–274.

Hausmann, Ricardo and Nedelkoska, Ljubica. 2018. "Welcome home in a crisis: Effects of return migration on the non-migrants' wages and employment." *European Economic Review*, 101(C): 101–132.

International Labour Organization. 2019. ILOSTAT. https://ilostat.ilo.org/.

Lacuesta, Aitor. 2006. "Emigration and Human Capital: Who Leaves, Who Comes Back and What Difference Does it Make?" Banco de España Working Papers 0620.

Mahuteau, S. and Tani, M. 2011. "Labour market outcomes and skill acquisition in the host country: north African migrants returning home from the European Union." IZA Discussion Paper No. 5441.

McKenzie, D. and J. Sasin, M. 2007. "Migration, Remittances, Poverty, and Human Capital: Conceptual and empirical challenges." World Bank Policy Research Working Paper 4272.

McKenzie, D. and Rapoport, Hillel. 2010. "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks." *Review of Economics and Statistics*, 92(4): 811–821.

Thomas, Kevin J. A. 2008. "Return Migration in Africa and the Relationship between Educational Attainment and Labor Market Success: Evidence from Uganda." *International Migration Review*, 42(3): 652–674.

Wahba, J. and Zenou, Y. 2005. "Density, social networks and job search methods: theory and application to Egypt." *Journal of Development Economics*, 78(2): 443–473.

Wahba, J. and Zenou, Y. 2012. "Out of sight, out of mind: Migration, entrepreneurship and social capital." *Regional Science and Urban Economics*, 42(5): 890–903.

Wahba, J. 2015. "Selection, selection, selection: the impact of return migration." *Journal of Population Economics*, 28(3): 535–563.

World Bank. 2019. World Development Indicators. Washington, DC: World Bank.

6. Appendices

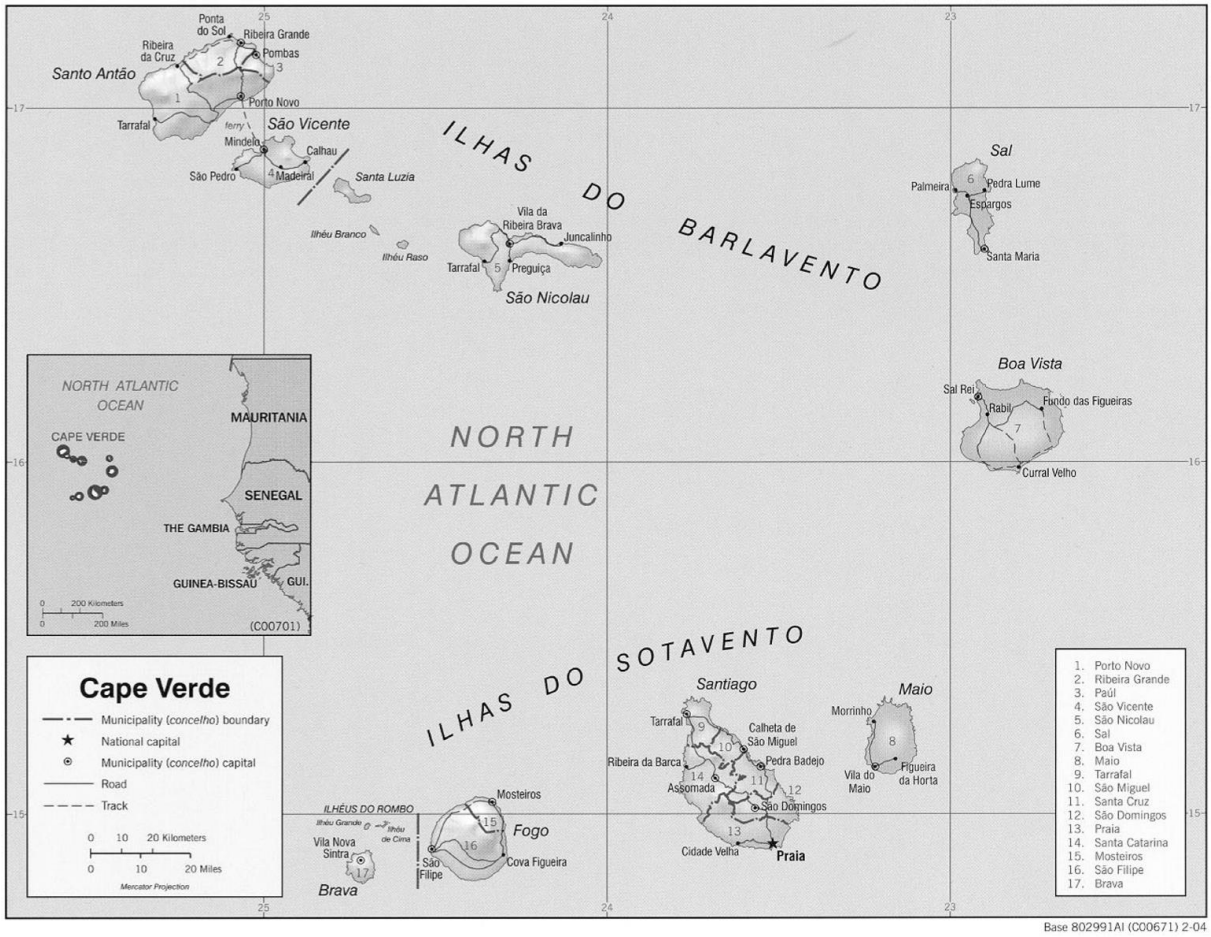
Appendix 1 – Table A1. Household Characteristics (all households and percentages)

	Migratory Experience (in percentage of total number of households)			
HHs with at least one international migrant	42.13			
HHs with at least one international return migrant	16.76			
HHs with at least one international current migrant	35.51			
	Remittances			
HHs who received international remittances	24.19			
	Mean Labor Income per Active Individual			
	Mean	Std. Dev.	Min	Max
HHs with at least one international migrant	390.77	325.18	17.28	1949.38
HHs with no migrants	278.73	314.35	24	3480
	Occupation of HH head			
Health, Education, Public Administration	20.36			
Agriculture	7.26			
Fishing, Industry, Transportation, Domestic, Self-employment	8.67			
Construction, Commerce	23.59			
Other Services	30.04			
Retired, Student, Unemployed, Other	10.08			

Appendix 2 – Table A2. Destinations of International Return and Current Migrants

Country	Return Migrants (%)	Current Migrants (%)
Portugal	45	48
USA	21	17
Netherlands	7	5.35
France	8.5	15.72
Luxembourg	2.5	4.35
Italy	2	3
São Tomé e Príncipe	6.5	0
Other	7.5	6.7
Total Observations	200	598

Appendix 3 – Figure 1. Map of Cape Verde



Source: Geographic Guide