What Makes Brain Drain More Likely? Evidence from Sub-Saharan Africa*

FIRST DRAFT

Romuald Méango,
Munich Center for the Economics of Aging

September, 22 2016

Abstract

In Sub-Saharan Africa, high-skilled workers are 13 times more likely to migrate than low-skilled ones. This sheer number has fueled fears about "Brain Drain" as only 3% of the population obtains tertiary education. Although migration prospects might give *incentives* to invest in schooling, it is still unclear for which households they exist and whether these can compensate for the *selection* of high-skilled workers into migration. This papers measures the *selection*, *incentive* and *net* effects of emigration from DR Congo, Ghana and Senegal to Europe. Institutional contexts and household characteristics are strong determinants of the three effects. Rich households experience a strong selection of high-skilled workers into migration, thereby decreasing the average schooling level in the origin countries. However, stronger incentives to invest in schooling partly or fully compensate for this decrease. By contrast, poor households experience small selection and equally small incentives, except in Senegal, where they exhibit negative incentives to invest in early schooling. This is possibly due to low returns to secondary education in Europe and/or binding liquidity constraints.

Keywords: Migration, Brain Drain, Brain Gain, Sub-Saharan Africa.

JEL: C30, I25, J61.

*The question in the title of the manuscript is one of the "Eight Questions about Brain Drain" stated by Gibson and McKenzie (2011). This research received financial support from the Leibniz Association (SAW-2012-ifo-3). I thank Theresa Koch and Judith Leopold for excellent research assistance. I greatly benefited from discussions with Axel Börsch-Supan, Tabea Bucher-Koenen, Esther Mirjam Girsberger, Marc Henry, Elie Murard, Ismaël Mourifié, Lars Nesheim, Çağlar Özden, François Poinas, Hillel Rapoport, and Frank Windermeijer, as well as participants to the 13th IZA Anuual Migration Meeting, the 9th MPC-EUI Migration and Development Conference, the 3rd IAAE Meeting, and workshops and seminars at the Ifo Institute, the Munich Center for the Economics of Aging, Toulouse School of Economics, and the University of Lausanne. The usual disclaimer applies. Correspondence address: Amalienstr. 33, 80799 Munich, Germany. E-mail: meango@mpisoc.mpg.de

1 Introduction

One adult out of three surveyed would like to emigrate permanently out of Sub-Saharan Africa. One out four potential migrants would like to enter the European Union.[1] For many of them, education has proved the best asset to fulfill this wish. In the early 2000s, 13% of the high-skilled Sub-Saharan African population lived abroad, yet the overall migration rate is only 1%. Individuals with a tertiary degree represented 43% of the migrant population, compared to 3% of the resident population (Easterly and Nyarko, 2009). Close to half of the high-skilled migrants travel to Europe. In these destination countries pressure increases for further selective migration policies. In the meantime, fears about "Brain Drain" and its negative consequences on development worsen in Sub-Saharan Africa.

"Brain Drain" arises as a consequence of the propensity for the more educated workers to migrate. Usually, the term refers to the subsequent reduction of the average level of schooling in the sending country, and could be seen as a *selection effect*.[2] For two decades, economists have pointed out the existence of a counterbalancing effect, by which better returns to schooling abroad and improved migration odds give incentives for more schooling investment in the sending country - this is the *incentive effect* of migration prospects. Thus, from the perspective of a sending country, what matters primarily is the *net effect*: the resulting change in average schooling after incentive and selection effects have taken place.

As Sub-Saharan African sending countries ponder what should be the appropriate policy response to high-skilled migration, economists should provide answers to three essential questions: how strong is the selection, how strong are the incentives, and how does it translate to the net effect? The empirical microeconomic literature has spent much time and effort in establishing the existence of positive incentives, but has made few attempts at measuring the net effect, and even fewer at distinguishing and measuring selection, incentive and net effects at the household level.[3] Moreover, in-depth studies of these effects in major sending countries from Sub-Saharan Africa are still rare.[4] This study aims at filling these gaps.

Acknowledging heterogeneous effects at the household level is relevant in at least three respects. (1) It helps to better understand the microeconomic mechanisms leading to the observed macroeconomic effects. (2) This improved understanding will in turn allow for designing better-informed and well-targeted policy responses. For example, the empirical analysis below shows that rich and poor households' schooling investments respond differently to migration prospects; therefore, any careful policy response should account for this discrepancy. (3) Finally, the distinction allows for better-suited econometric analysis. For example, Batista, Lacuesta, and Vicente (2012) use traditional instrumental variable estimation framework, that is not taking account of the heterogeneity of households. Thus, their estimator captures the incentive effect

[1] Gallup Survey, 2010-2011, Online results, accessed on 26 July 2016, at: http://www.gallup.com/poll/124028/700-million-worldwide-desire-migrate-permanently.aspx

[2] Here, I refer to the difference in the proportions of high-skilled in the resident population before and after migration. The term "brain drain" can also refer to the absolute decrease in the high-skilled population, that is the *number* of high-skilled migrants. Some use the term "brain drain" as synonymous to high skilled migration. I thank Çağlar Özden for this insight.

[3] See the related literature in the next section.

[4] Easterly and Nyarko (2009), and Batista, Lacuesta, and Vicente (2012) are exceptions. Many studies examine the trends and consequences of high skilled migration on the economic development of Sub-Saharan Africa (Clemens, 2007, 2011; Özden and Phillips, 2014; Tankwanchi, Özden, and Vermund, 2013).

on a special part of the population, known as the *compliers* (Angrist, Imbens, and Rubin, 1996). Extrapolation of the results to the whole population might be quite misleading.

The empirical analysis rests on a generalized Roy model for households' schooling investment with a migration option. Households make two simultaneous decisions, one about schooling and one about migration attempt. The result of an attempt is not known before schooling investments are sunk, but the household has a subjective, schooling-dependent probability of success. In this context, each household is endowed with two potential outcomes: the schooling investment when they decide to attempt migration, and the schooling investment when they decide *not to* attempt migration.

A positive *incentive effect* would arise if more schooling increases the odds of migration or wages abroad. It is measured by the difference in the average schooling of the population between the observed (factual) state of the economy and the counterfactual scenario of restricted migration. After schooling investment are sunk, some of those who decided to attempt migration leave the country. A negative *selection effect* would arise if the high-skilled migrate disproportionately more often than the low-skilled, thus decreasing the average schooling in the origin country. This is measured by the gap between the average schooling of non-migrants and the average schooling of the whole population (non-migrants plus migrants). Finally, the *net effect* is the ensuing change in the average schooling of non-migrants between the factual and the counterfactual.

The main challenge in the existing literature is to retrieve households' schooling investment in the counterfactual scenario. The first major contribution of this study is to provide a useful characterization of households' schooling investment in the counterfactual scenario of *closed economy*, that is when no one is allowed to migrate. Within the generalized Roy model, a household's schooling investment in the *closed economy* equals exactly the potential outcome of schooling when this household decides *not to* attempt migration. This simple characterization can be used for the estimation of the three effects through well-established econometric tools, as long as one can observe the decision to attempt migration or not: *Matching*, *Local Instrumental Variable*, *Unrestricted* and *Restricted Bounds* estimators, among others.

The Migration from Africa to Europe (MAFE) project survey contains both detailed information about migration attempts, and actual emigration spells from Sub-Saharan African countries to Europe. The MAFE survey covers three major sending countries: the Democratic Republic of Congo (DR Congo), Ghana and Senegal. Furthermore, it contains detailed information about education, labor market history, socioeconomic and demographic characteristics for both non-migrants and migrants to major destinations in Europe. These features makes it uniquely suited for the present study.

The empirical analysis in this paper shows that together the institutional context and households characteristics determine the direction and magnitude of the selection, incentive and net effects. In the DR Congo, where the migration attempt rate is fairly low and the average schooling level comparatively high, migration prospects have almost no impact on households' schooling investments. In Ghana, where the migration attempt rate is relatively high (mostly among high skilled workers) and the average schooling level is comparatively high, selection of high-skilled workers into migration leads to a decrease of the average human capital. This effect is sizable among rich households and households with a previous migrant. , thereby decreasing the average schooling level in the origin countries. However, migration prospects also give stronger incentives

to invest in schooling. This compensates for the decrease due to selection.

Of the three countries, Senegal is the peculiar case. Migration attempt rates are high in all subgroups of the population, while the average schooling level is low (3 to 5 less schooling years in comparison to Ghana and Congo). The selection of the high-skilled into migration is concentrated among rich households, yet this again is possibly compensated for by positive incentives. However, the poor population might have negative incentives to invest, even in early schooling. In the pessimistic case, this would amount to a 16 to 31% reduction in enrollment at upper secondary level, compared to the *closed economy* scenario. Explanations for this finding are the comparatively low returns associated with secondary education in Europe, and binding liquidity constraints.

The rest of the paper proceeds as follows: Section 2 links the study to the existing literature. Section 3 motivates the measures of the three effects, and characterizes the schooling investment in the counterfactual scenario. Section 4 discusses identification assumptions for four estimators: *Matching*, *Local Instrumental Variable*, *Unrestricted* and *Restricted Bounds* estimators. Section 5 begins the empirical analysis by presenting the data, and some descriptive statistics, and by assessing the validity of the estimation assumptions. Section 6 presents the main results, and Section 7 provides a discussion of these results. Finally, Section 8 concludes. Some technical details of the estimation are presented in the Appendix.

2 Related Literature

During the last two decades, the interest of economists for the "brain drain" has been revived by two important sets of contributions. The first set, mostly led by theoretical contributions, argued for the existence of a potential *incentive effect* that could cancel out, or even overturn, the negative *selection effect* (Mountford, 1997; Stark, Helmenstein, and Prskawetz, 1997; Vidal, 1998). This has been called the "Brain Gain". The second set of contributions provided empirical support for the existence of the incentive effect in some contexts (Batista, Lacuesta, and Vicente, 2012; Beine, Docquier, and Rapoport, 2008; Chand and Clemens, 2008; Shrestha, Forthcoming; Theoharides, 2014). However, in the context of illegal and labor migration for low skill jobs, migration prospects can produce negative incentives for schooling investments (Girsberger, 2014; McKenzie and Rapoport, 2011). This paper differentiates and measures the selection, incentive and net effects across households. Doing so provides a better understanding of the microeconomic mechanisms generating the observed macroeconomic outcomes. Understanding these effects allows for better policy designs to address the concerns raised by high skilled migration.

Since its origin, the empirical literature faces the challenge of identifying the counterfactual schooling investment in the case of restricted migration. Natural experiments offer set-ups to test the theory (Chand and Clemens, 2008; Shrestha, Forthcoming). However, their external validity is questionable. Studies that have used instrumental variable strategy have failed to account for the heterogeneity in households (Batista, Lacuesta, and Vicente, 2012). Since traditional instrumental variable estimations capture effects on special parts of the population, extrapolation might be misleading (Angrist, Imbens, and Rubin, 1996).

This study improves on the previous literature in several respects. The schooling investment of heterogenous households is characterized in the counterfactual scenario of a *closed economy* - a counterfactual largely discussed in the literature, for example Mountford (1997), Stark,

Helmenstein, and Prskawetz (1997), and Beine, Docquier, and Rapoport (2001, 2008). This counterfactual schooling investment is the schooling investment when the individual *does not* attempt migration. The unique data used in this study contains information on migration attempts by the respondents.[5]

Observation of migration attempts allows using several estimation techniques to identify and estimate the counterfactual schooling investment, i.e., matching, local instrumental variable, and bounds. Hence, previous stringent assumptions found in the literature on the functional form of the model equations, the structure of the error terms, and the properties of the instrumental variables are substantially relax. Moreover, the critical assumptions underlying the proposed estimation techniques are assessed in Section 5.4. If the underlying assumptions of the estimation techniques fail, the range of values that the effects of interest can take using worst-case bounds is characterized in Section 4.2.

Return migration and remittances are alternative channels through which the sending country can experience an increase in its human capital (Gibson and McKenzie, 2011; Dinkelman and Mariotti, 2015; Theoharides, 2014). The framework in this paper can isolate the contribution of returned migrants to average schooling level at origin, presented in Section E.2. Conceptually, the same could be done with the contribution of remittances. However, the data do not contain information about remittances at the time of schooling investment. Nevertheless, the discussion of the results addresses the case where the households has q member living abroad at the time of schooling investment (see Section E.1). Finally, much of the public discussion focuses on *absolute measures* of the brain drain, that is, the number of high-skilled that are "lost", rather than the proportion of the resident population. These absolute measures are considered in Section E.3.

3 Measures of the Effects of Migration on Schooling Decision

The net effect of migration on households' schooling investment is measured by comparing the average level of schooling in the observed (factual) state of the economy to the schooling investment in an hypothetical (counterfactual) situation where no migration is possible, the *closed economy* (Section 3.1). Since the factual household's schooling investment is observed, the main challenge is to characterize the counterfactual schooling investment in the case of closed economy; hence, the need for the model described in Section 3.2.

3.1 Empirical Measures of the Selection, Incentive and Net Effect at the household level

I consider a framework based on the human capital literature, where education is considered an investment in future earnings and employment for rationale agents who seek to maximize their lifetime earnings (Willis and Rosen, 1979). The simplest framework has two countries, the origin country (0) and the destination country (1), and two schooling levels, low (l) and high (h). Consider two periods. In the first period, in the origin country, a household with a child makes two choices: a schooling choice $S = \{l, h\}$, and a choice to attempt migration $M^* \in \{0, 1\}$. The

[5] Besides, the data set allows observing migrants in their destination countries, while (Batista, Lacuesta, and Vicente, 2012) have the concern that households who emigrate and leave no one in the origin country are not accounted for. This is a possible source of biases studied by Steinmayr (2014) and Murard (2016).

schooling investment is implemented in the first period. The attempt to emigrate is made in the second period, given the level of schooling. It can be either successful or not. $M \in \{0,1\}$ is the migration status in the second period.

Let X be a set of a household's observable characteristics, u, a set of a household's unobserved characteristics (e.g. child's ability), and (p_l, p_h) a set of household-specific subjective probabilities. p_l (resp. p_h) is the household's subjective probability that the migration attempt succeed when the child has schooling l (resp. h). The set (X, u, p_h, p_l) is the information set of the household at the time it makes the schooling and attempt choices. Given this information set and an attempt decision $M^* = m^*$, the household chooses the schooling level S to maximize the expected return to schooling. In particular, in the counterfactual scenario of a *closed economy*, $p_h = p_l = 0$. Let S_{cf} be the household's schooling choice in this counterfactual scenario. In the next section, I characterize S and S_{cf}. Before doing so, I present measures of the average selection, incentive, and net effect for all households with observable characteristics $X = x$.

The average *selection effect* for households with characteristics $X = x$, say $\Delta_{sel}(x)$, is the difference between the average schooling of residents (with characteristic x) and the average schooling of the whole population (residents and migrants with characteristic x).

$$\Delta_{sel}(x) := \mathbb{E}(S|Y = 0, X = x) - \mathbb{E}(S|X = x) \tag{1}$$

The average *incentive effect* for households with characteristics $X = x$, say $\Delta_{inc}(x)$, is the difference between the average schooling of the whole population (with characteristic x) and the average schooling of the whole population (with characteristic x) in the closed economy.

$$\Delta_{inc}(x) := \mathbb{E}(S|X = x) - \mathbb{E}(S_{cf}|X = x) \tag{2}$$

The resulting average *net effect*, say $\Delta_{net}(x)$ is the sum of the selection and the incentive effect:

$$\Delta_{net}(x) := \Delta_{sel}(x) + \Delta_{inc}(x) = \mathbb{E}(S|Y = 0, X = x) - \mathbb{E}(S_{cf}|X = x) \tag{3}$$

$\Delta_{net}(x)$ is the measure of the net effect in the theoretical models discussed by Mountford (1997), Stark, Helmenstein, and Prskawetz (1997), and Beine, Docquier, and Rapoport (2001, 2008), now defined at the household level.

The proposed measures can be easily modified to additionally account for return migration. Denoting as $\{\mathcal{R}\}$ the pool of never-migrants and returned migrants, the average net effect including returners from households with characteristics $X = x$ is defined as:

$$\Delta^r_{net}(x) \equiv \mathbb{E}(S|\{\mathcal{R}\}, X = x) - \mathbb{E}(S_{cf}|X = x) \tag{4}$$

If $\Delta^r_{net} > 0$ while $\Delta_{net} < 0$, then return migration is important to compensate for the *ex ante* decrease in average schooling.

3.2 Characterization of Households' Schooling Investment

Consider the schooling decision given the choice to attempt migration, which defines two potential outcomes. Let $S(0)$ be the schooling choice when the individual does *not* attempt migration.

Correspondingly, let $S(1)$ be the schooling choice when the individual attempts migration. In the following, I show that $S_{cf} = S(0)$.

Let $\Pi_d^m(x,u)$ be the net return (gains net of the costs) to schooling level $s \in \{l,h\}$ in location $m \in \{0,1\}$.

$$\Pi_s^m(x,u) = \Pi_s^m(x) + u_s^m \tag{5}$$

$\Pi_s^y(X)$ is the average net expected return to schooling s for a household with characteristics x. u_s^m is a latent cost of schooling s that I interpret as the unobserved ability of the child or a private consumption value.

As in Rosenzweig (2008), the returns to schooling depend on the expected location in the second period. Given M^*, the household's expected return to education s is :

$$\begin{aligned} &\underbrace{(1-M^*)\left[\Pi_s^0(X) + u_s^0\right]}_{\text{No attempt}} + \\ &\underbrace{M^*\left[(1-p_s)\times(\Pi_s^0(X) + u_s^0)\right]}_{\text{Unsuccessful Attempt}} + \\ &\underbrace{M^*\left[p_s \times (\Pi_s^1(X) + u_s^1)\right]}_{\text{Successful Attempt}} \end{aligned} \tag{6}$$

The first line is the return to schooling s when the household chooses *not to* attempt migration. The second line is the return when an unsuccessful attempt is made. Finally, the third line is the return when the child migrate to the destination country 1 with education s.

Hence, a household with characteristic x chooses $S(0) = h$ over $S(0) = l$, if and only if :

$$\Pi_h^0(x) + u_h^0 - (\Pi_l^0(x) + u_l^0) > 0 \tag{7}$$

The household chooses $S(1) = h$ over $S(1) = l$, if and only if :

$$\begin{aligned} (\Pi_h^0(X) \quad &+ \quad u_h^0) - (\Pi_l^0(x) + u_l^0) \\ &+ \quad p_h \times (\Pi_h^1(X) + u_h^1 - (\Pi_h^0(X) + u_h^0)) \\ &- \quad p_l \times (\Pi_l^1(X) + u_l^1 - (\Pi_l^0(X) + u_l^0)) > 0 \end{aligned} \tag{8}$$

Equations (7) and (8) together imply that $S_{cf} = S(0)$, since the return to schooling is the same whether $p_1 = p_0 = 0$ or $Y^* = 0$. [6] Hence, in Equations (2) - (4),

$$\mathbb{E}(S_{cf}|X = x) = \mathbb{E}(S(0)|X = x) \tag{9}$$

The next section discusses the identification of the selection, incentive and net effects, in particular, the identification of $E(S(0)|X = x)$.

[6] Appendix A discusses an extension to the case where the budget constraint is binding in the presence of an emigration option.

4 Identification

The schooling choice S, the migration status M, and the characteristics X are all observed in the data for each household. Hence, the average selection effect for each subgroup $X = x$, $\Delta_{sel}(x)$, is identified. Furthermore, M^*, the attempt choice, is also observed. However, $S(0)$ is observed when the household chooses *not to* attempt migration, but unobserved when the household chooses to attempt migration. Thus, identification of counterfactual quantity $\mathbb{E}(S(0)|X)$ is more challenging.

First, Section 4.1 discusses two well-known alternative sets of assumptions leading to point identification (*strong ignorability* and *local instrumental variable*). Then, Section 4.2 shows that informative bounds can be derived with less demanding assumptions.

4.1 Point Identification

The first set of assumptions leading to point identification are known as "strong ignorability" assumptions. The second set of assumptions, that I call "local instrumental variable" assumptions, rests on the existence of an exclusion restriction.

4.1.1 Strong Ignorability

"Strong ignorability" has two components:

SI-1 (Overlap) $P(\tilde{X} = x|M^* = 1) < 1$

SI-2 (Selection-on-observable) Let $\tilde{X}$ be a set of observable characteristics of the household, such that X is a sub-vector of $\tilde{X}$. $S(0)$ is independent of M^* conditional on $\tilde{X}$.

Under SI-1 and SI-2,

$$\mathbb{E}(S(0)|X) = \mathbb{E}\left(\mathbb{E}(S|M^* = 0, \tilde{X})|X\right). \tag{10}$$

The right-hand side of the above equation is (point) identified. Matching is used to implement the result of Equation (10), as the survey provides a rich set of information about the household. In the empirical application below, households who attempt migration are matched to households who do not attempt migration on gender, father's occupation, age, religion, ethnicity, household size (number of siblings), and household's migration network size (number of migrants that the respondent reports as an acquaintance at age 15). More details about the estimation procedure are presented in Appendix D.

4.1.2 Local Instrument Variable

"Local Instrumental Variable" has three components:

LIV-1 (Exclusion restriction) Let Z be random variable such that, $S(0)$ is independent of Z conditional on $\tilde{X}$.

LIV-2 (Selection equation) There exists a random variable U_M such that $M^* = I(P(M^* = 1|X, Z) > U_M)$, where $P(M^* = 1|X, Z)$ is a non-trivial function of X.

LIV-3 (Separability) There exists a random variable U_M and a function μ_{M^*} such that $S = \mu_{M^*}(X) + U_S$

Under the above conditions, Heckman and Vytlacil (2007) show that there exists a real function, K, defined on the unit interval, such that:

$$\mathbb{E}(S|X) = \mathbb{E}(S(0)|X) + K(P(M^* = 1|X, Z)) \tag{11}$$

The first term on the right-hand side is identified, provided sufficient variation of the propensity score $P(M^* = 1|X, Z)$. The estimation of Equation (10) is conducted by gender, father's-occupation, and households'-with-migrant status. In the empirical application, the instrument Z is a measure of labor demand shocks in each European country weighted by the proportion of the household's network based in each of these countries. The identifying assumption is that these weighted demand shocks have no effect on the schooling decision when the individual does not attempt migration.[7] The construction of the instrument is described in Appendix C. More details about the estimation procedure are presented in Appendix D.

Both "strong ignorability" and "local instrument variable" are strong and ultimately untestable assumptions. Section 5.4 discusses their plausibility based on the data. The next Section presents identification results under less demanding assumptions.

4.2 Set Identification

The first set of bounds is the most extreme possible (worst-case bounds). The second set of bounds assumes positive selection and sorting into migration.

4.2.1 Worst-case bounds

Without additional assumption on the model, $\mathbb{E}(S(0)|X)$ must lie between bounds that correspond to two extreme cases:

B-1 (Maximum incentive) If they would have not attempted migration, *none* of those who attempt migration in the current economy would have obtained schooling $S = h$.

B-2 (Minimum incentive) If they would have not attempted migration, *all* of those who attempt migration in the current economy would have obtained schooling $S = h$.

B-1 corresponds to a migration scenario with maximal possible incentive effect, hence, to the maximal possible net effect. B-2 corresponds to a migration scenario the minimal possible incentive effect (possibly negative), hence, to the minimal possible net effect. It follows that:

$$0 \leq P(S(0) = h, M^* = 1|X) \leq P(M^* = 1|X),$$

$$P(S = h, M^* = 0|X) \leq \mathbb{E}(S(0)|X) \leq P(S = h, M^* = 0|X) + P(M^* = 1|X),$$

[7] This is a much weaker exogeneity condition than the one entertained by Batista, Lacuesta, and Vicente (2012). They require that Z is independent of both $S(0)$ and $S(1)$. In fact, if individuals are forward looking it seems plausible that the weighted labor demand shocks Z have an effect on the schooling choice, when one decides to attempt migration $S(1)$.

and:

$$\begin{array}{l} P(S=h|X) \\ -(P(S=h,M^*=0|X)+P(M^*=1|X)) \end{array} \leq \Delta_{inc}(X) \leq \begin{array}{l} P(S=h|X) \\ -P(S=h,M^*=0|X) \end{array},$$

$$\begin{array}{l} P(S=h|M=0,X) \\ -(P(S=h,M^*=0|X)+P(M^*=1|X)) \end{array} \leq \Delta_{net}(X) \leq \begin{array}{l} P(S=h|M=0,X) \\ -P(S=h,M^*=0|X) \end{array}.$$

From the bounds on the net effect, one can test for the existence of a strictly positive net effect (even without an instrument).

4.2.2 Restricted bounds

The "worst-case bounds" result from a completely agnostic approach towards the direction of the selection into migration. However, the economic literature is far from being agnostic on this issue.[8] In the following, hypotheses are introduced that are compatible with both the "Brain Drain", as exposed, for example, by Bhagwati and Hamada (1974) and the "Brain Gain" theories, as exposed by Mountford (1997); Stark, Helmenstein, and Prskawetz (1997); Vidal (1998).

For the "Brain Drain" theory to be valid there is no need that all people who attempt migration would have obtained maximum education, had they not attempted migration. Instead, it is crucial that:

RB-1 (Positive selection) If they would have not attempted migration, those attempting migration would have obtained (on average) at least the same schooling as those not attempting.

In other words, potential migrants are *positively selected*. It follows that:

$$P(S=h|M^*=0,X) \leq P(S(0)=h|M^*=1,X). \tag{12}$$

The "Brain Gain" argument does not object to the previous point; rather, it claims that, (legal) migration provides additional incentives for schooling. Hence:

RB-2 (Positive sorting) If they would have not attempted migration, those attempting migration (legally) would have obtained (on average), at most, as much schooling as they do when attempting migration.

In other words, potential migrants are *positively sorted*. It follows that:

$$P(S(0)=h|M^*=1,X) \leq P(S=h|M^*=1,X). \tag{13}$$

Both conditions have strong support in the literature (Grogger and Hanson, 2011). The restrictions on $\mathbb{E}(S(0)|X)$, $\Delta_{inc}(X)$, and $\Delta_{net}(X)$ trivially follow. Moreover, the positive selection and positive sorting assumptions have an important testable implication. Equations (12) and (13) imply that for all $X=x$,

$$P(S=h|M^*=0,X) \leq P(S=h|M^*=1,X). \tag{14}$$

[8] See for examples Borjas (1987) and Grogger and Hanson (2011).

	Ghana	Senegal	DR Congo
Year of Survey	2009-2010	2008	2009-2010
Destinations [a]	NL, UK	FR, IT, SP	BE, UK
Respondents	1,665	1,668	2,066
(+ restriction)[b]	(1,364)	(1,049)	(1,686)
Migrants EU (%)	30.7	40.3	23.2
(+ restriction)	(30.7)	(42.6)	(20.3)
Men (%)	56.9	48.2	44.2
(+ restriction)	(57.4)	(49.8)	(43.2)

[a] NL: the Netherlands; UK: United Kingdom; FR: France; IT: Italy; SP: Spain; BE: Belgium.
[b] Restricted sample with individuals who are aged between 25 and 60, who have at least some formal schooling, and who have not migrated before age 21.

Table 1: Summary Information about the MAFE Project Survey

Since it has to holds for all cells defined by X, this is a very demanding condition tested subsequently in the data. Overall, there is strong support in the data for the validity of Equation (14). The next Section begins the empirical analysis.

5 Data, Descriptive Statistics and Assessment of Assumptions

5.1 The MAFE Survey

The empirical analysis is based on longitudinal biographical survey data collected in the framework of the Migration between Africa and Europe (MAFE) Project.[9] The survey was conducted in the capital cities of three Sub-saharan African countries (Kinshasa - DR Congo, Accra - Ghana, and Dakar - Senegal). In the following, countries are refered to the countries, rather than capital cities. A representative sample of households was interviewed in each origin country. Then, for households with a migrant member, the migrant was traced and interviewed if he migrated to one of the major destinations in Europe. The sample of migrants was augmented using a snowball sampling methodology. Sampling weights are added to produce a representative sample. Table 1 presents the years of data collection, the European countries were interviews were conducted, the sample size, and the proportion of respondents that are migrants for each origin country. For more details on the MAFE project methodology see Beauchemin (2012).

The survey collects retrospective biographical information about the respondents' demographic and socioeconomic characteristics, and labor force participation history. For each household there is information about: demographic characteristics, past and current migrant network, current financial transfers and living conditions. The major attractiveness of the MAFE survey data is the

[9]The MAFE project is coordinated by the Institut National d'Études Démographiques (INED) (C. Beauchemin) and is formed, additionally by the Université catholique de Louvain (B. Schoumaker), Maastricht University (V. Mazzucato), the Université Cheikh Anta Diop (P. Sakho), the Université de Kinshasa (J. Mangalu), the University of Ghana (P. Quartey), the Universitat Pompeu Fabra (P. Baizan), the Consejo Superior de Investigaciones Cientàficas (A. Gonzàlez-Ferrer), the Forum Internazionale ed Europeo di Ricerche sull'Immigrazione (E. Castagnone), and the University of Sussex (R. Black). The MAFE project received funding from the European Community's Seventh Framework Programme under grant agreement 217206. The MAFE-Senegal survey was conducted with the financial support of INED, the Agence Nationale de la Recherche (France), the Région Ile de France and the FSP programme 'International Migrations, territorial reorganizations and development of the countries of the South'. For more details, see: http://www.mafeproject.com/

information about actual migration history, and (unsuccessful) migration attempts. The survey records: year and destination of attempt, documentation status, and reasons of failure.

In the following analysis the sample is restricted to individuals who never migrated to Europe before age 21 to ensure that they obtained education in Senegal. Individuals aged 60 or more are also excluded because they presumably made schooling investments during colonial years. Table 1 contains summary information about the restricted sample.

5.2 Main Variables of Interest

The general context of emigration from DR Congo, Ghana and Senegal is described by Baizán, Beauchemin, and González-Ferrer (2013); Beauchemin, Sakho, Schoumaker, and Flahaux (2014); Schans, Valentina, Schoumaker, and Flahaux (2013); Schoumaker, Flahaux, and Mobhe (2013). I focus on aspects relevant to the "brain drain" discussion (schooling level, migration attempt and actual migration propensities), stressing similarities and differences between the three countries.

The survey records information about the last year of schooling successfully completed by the respondent.[10] The average schooling level is lowest in Senegal (about 9 years), 4.5 years less than in DR Congo, and 4 years less than Ghana. In all three countries, men are more educated than women, with a gender schooling gap of 2.3 years in DR Congo, 2.6 years in Ghana, and 1.1 years in Senegal.[11]

Figure 1 compares the schooling level distributions of those residing in the main migration destination and of the rest of the population, by country of origin and by gender. The upper panel is for men, the lower panel for women. Education is categorized into four groups: at most some primary education, some lower secondary education, some upper secondary education, and some tertiary education. In DR Congo, the majority of male residents (48%) have obtained some upper secondary education. By contrast, more than 75% of migrants to Europe have obtained some tertiary education. High educated individuals are also over-represented among migrant women. The picture is very similar in Ghana. In opposition to DR Congo and Ghana, "Primary education" is the most important group among residents in Senegal(close to 50%). Still, migrants have higher education than residents.

The MAFE survey is uniquely suited for the present analysis because it records information on past migration attempts. For each migration attempt, respondents report the intended destination, the year(s) during which the attempt took place, the steps undertook, the failure or the success, and the reason of the failure, when applicable. In the baseline estimation, a migration attempt is defined as any self-reported attempt, irrespective of the stage at which the attempt stopped. In a robustness analysis, a stricter definition of a migration attempt is implemented (see Section E.4).

At this point, it is worth discussing two limitations of the model. First, a migration attempt is usually observed after education completion, while the model presents the two decisions as simultaneous. This anachronism implies that some individuals might have changed their mind

[10]The MAFE Survey data divides the curricula into four levels:(i) primary education: 1 to 7 years in DR Congo and Ghana, 1 to 6 years in Senegal, (ii) lower secondary education: 8 to 11 years in DR Congo and Ghana, 7 to 10 years in Senegal, (iii) Upper Secondary education: 12 to 14 years in DR Congo and Ghana, 11 to 13 years in Senegal, (iv) and tertiary education.

[11]In DR Congo, free and compulsory education between age 6 to 12 (primary school) is stipulated in the constitution. In Ghana, free and compulsory primary school has been introduced in 1961 and extended to cover all children between 6 to 14 years of age in 1981. Only recently in Senegal (2004) have tuition fees for primary education been waved and compulsory education introduced for children aged between 6 to 16 years of age.

between the time they made the schooling investment decision and the time when the attempt decision is observed. Unfortunately, to the best of my knowledge, there exists no comparable data source that is more precise on the attempt decision during years of schooling.

Second, attempts involve different levels of investments that are not captured by the binary structure of the variable M^* in the model.[12] Thus, one might see the attempt variable as a continuous variable. Nevertheless, the model can be adapted by assuming that, first, households decide attempting migration or not; then, they choose the level of effort to invest in the attempt. The latter choice will determine the subjective probability of success. As long as no household invests in an attempt when the success probability is zero, the main prediction of the model is valid.

In the following, a migrant is defined as someone who was born in one of the African countries (DR Congo, Ghana, Senegal) and had emigrated out of Africa at age 21 or later, for a stay of at least one year in one of the main European destinations. This restriction is dictated by data constraints since comparable information on respondents' households are only available for residents and migrants to the main destinations (for example the father's occupation at age 15 or the household-with-migrant status).

Figure 2 shows the proportion of the population who attempted migration to Europe and the proportion of those who actually migrated, by country of origin, gender and schooling. The upper panel is for men, the lower panel for women. The probability to attempt migration varies substantially across countries. DR Congo has the lowest rate followed by Ghana and finally Senegal where one out of three men, and one out of six women attempted migration. Moreover, respondents with more schooling are clearly more likely to attempt migration, and to migrate.

5.3 Household Characteristics

The estimation strategy differentiates the selection, incentive and net effect by the following characteristics: gender, father's occupation when the potential migrant is aged 15, and the existence of a previous migrant member when the potential migrant is aged 15 (household-with-migrant status). For each of these subgroups, estimation of the bounds is conducted separately.[13] The father's occupation is divided into four categories: high-level occupation or employer, skilled employee, unskilled employee, and self-employed (without employee) or unemployed. Father's occupation proxies household's wealth. Thus, it allows understanding, which of the poor or rich households are most likely to experience strong selection or incentive effects.

Tables 3, 4, and 5 in Appendix B compare the observed characteristics of those who attempt migration (treated), to the characteristics of those who never attempt migration (the non-treated). Further characteristics used to match the two groups are: network size at age 15, age at survey, household size (not presented), religion (not presented), and ethnicity (when available, not presented).

Overall, those who attempt are more likely to have fathers with high-level or skilled occupations. They are also more likely to have at least one household member living abroad when they are 15 years old. Thus, their migrant network is on average larger. There is no obvious difference of

[12] For example, some respondents "failed" because they did not receive a visa, while some other did not initiate any administrative procedure.

[13] Father's education is also available but highly correlated with occupation.

household size between the two groups. However, the distribution of religious and ethnic groups differ substantially between the two groups, suggesting the importance of religious and ethnic networks.

5.4 Assessment of Assumptions

Matching, local instrument variables and the restricted version of the bounds rest on different sets of assumptions that can be assessed to a certain extent.

5.4.1 Matching: Overlap and Selection-on-Observable

A lack of overlap (SI-1) can be assessed in the data. For any given characteristic, a difference of means between treated and non-treated groups larger than a quarter of a standard deviation is symptomatic of a lack of overlap (Imbens, 2015). Tables 3, 4, and 5 in Appendix B, indeed show that for several characteristics, the normalized difference is larger than 0.25. Therefore, to ensure overlap, I drop observations outside the common support of the propensity score for both groups. This has little effect for the estimation on DR Congo and Ghana, and on women in Senegal. However, a quarter of the treated respondents (88 observations) are dropped among men in Senegal.

Selection-on-Observable (SI-2) is untestable; however, finding no treatment effect on pre-treatment variables strengthen the claim for the validity of SI-2. The treatment is the decision to attempt migration; since its implementation occurs later in life, it is reasonable to think that the decision is not taken very early in life. Hence, early schooling decisions should not be affected by the decision to attempt migration. Considering the decision to enroll in secondary education, the matching procedure finds a zero effect in DR Congo and Ghana.[14] However, it suggests a negative, statistically significant effect on Senegalese men. Therefore, one cannot be confident that, for men in Senegal, selection-on-observable holds based on this analysis.

5.4.2 LIV: Exogeneity and Relevance

The construction of the instrumental variable is detailed in Appendix C. For the LIV methodology to identify the incentive effect, the main assumption is that labor demand shocks at destination have no effect on the schooling decision when the individual does not attempt migration. This assumption is plausible, but ultimately untestable.

The second requirement is that the instrument is a "strong" predictor of the decision to attempt migration. I conduct the traditional F-test on the first-stage equation to ascertain the strength of the instrument. Only for men in Senegal is the F-stat above the usual threshold of 10 (F-stat=17.35). Otherwise, the F-stats range from up to 9.63 (women in Senegal) to as low as 0.27. Hence, the presence of weak instrument might lead to biased estimates. In line with this concern, I find that the LIV estimates sometimes lie outside the worst-case bounds estimates (with disjoint confidence intervals). Therefore, one cannot be confident that the LIV estimates are unbiased.

[14] This is the result of enrollment rates close to 100%.

5.4.3 Restricted Bounds: Selection and Sorting

The restricted bounds assume positive selection and positive sorting (RB-1 and RB-2). Since the potential outcome $S(0)$ is unobserved, these two assumptions are untestable. However, they jointly imply that:

$$P(S = h|M^* = 0, X = x) \leq P(S = h|M^* = 1, X = x) \text{ for all } x. \tag{15}$$

This condition is very demanding since it must hold for all subgroups defined by X. In the present set-up, the subgroups are characterized by the gender, father's-occupation, and household-with migrant status. this amounts to 16 subgroups. In each country, for each x, the test is $H_0 : P(S = h|M^* = 0, X = x) \leq P(S = h|M^* = 1, X = x)$ for all x, against $H_1 : P(S = h|M^* = 0, X = x) > P(S = h|M^* = 1, X = x)$, for the variable S defined successively as obtaining either secondary education, upper secondary education, or tertiary education and S as number of years of schooling; that is 64 tests times three country. The null hypothesis is rejected twice at the 10% level and never rejected at the 5% level .[15]

In the empirical analysis to follow, Assumption RB-2 applies only to those who migrate legally. Legal migrants are defined as those who report arriving in Europe with a proper residence permit. While, the MAFE survey data allows observing residence status only after successful migration, it does not allow observing whether an attempt is made through exclusively legal ways. No restriction is imposed on the counterfactual schooling investment of the rest of the population.

Furhermore, the MAFE survey data contains some information on wages.[16] Using standard Mincer regressions, the unexplained productivity in the origin country can be compared for those who attempt migration, and those who do not attempt migration. In all three countries, the distribution for those who attempt migration stochastically dominates the distribution for those who do not attempt migration (results not reported), strengthening the claim of positive selection into migration attempt. The conclusion is that the assumptions of positive selection and positive sorting are plausible.

The technical details of each estimation procedure are described in Appendix D. The main results are presented in the next section.

6 Results

The estimation results are presented for the selection, the incentive and the net effect respectively. An assumption of the baseline estimation is that migration investments are *not* decided very early in life. For this reason, the focus is first on individuals with some secondary education. For each country, the effects on completion of some upper secondary and tertiary education, as well as the number of years of schooling are discussed separately, by gender, father's occupation, and households-with-migrant status. The main focus is on households without a migrant. A detailed description of the results for this group is provided in Sections 6.1 and 6.2. A short summary is offered in Section 6.3. To keep the main exposition concise, additional results are reportted in

[15] For each country, the joint test is not rejected at the 5% level. The converse test, permuting H_0 and H_1, leads to 89 rejections of the null at the 5% level and 68 rejections at the 5% level.

[16] Respondents provided retrospective information on their employment history. Wage is recorded for the end period of each employment spell.

Appendix E. Hence, households with a migrant are considered in Appendix E.1. Return migration is considered in Section E.2. *Absolute measures* of selection, incentive and net effects are considered in Section E.3. Finally, alternative specifications (for example, including individuals with primary education) are considered in Section E.4.

6.1 Selection Effect (Households without a migrant)

The *selection effect*, Δ_{sel}, measures the gap between the average schooling of non-migrants and the average schooling of the whole population (non-migrants and migrants). This quantity is directly identified from the data (with some sampling error), without any further assumption.

Starting with upper secondary education, Figure 3 shows, by country, by gender, and father's-occupation subgroup, the point estimates for the selection effect (orange dot on the left in each father-occupation group) and corresponding 90% confidence intervals (thin gray line in the background). It also shows the average effect for all occupation groups (first from the left).

First, consider men (left panels in Figure 3). In DR Congo, the selection effect at the upper secondary level is virtually zero in all gender and father's-occupation subgroups. Hence, the proportion of men with some upper secondary education does not decrease in this country because of selection into migration. The picture is similar in Ghana. By contrast, in Senegal, the selection effect is negative for all men taken together (-4.3 percentage points (pp)). This negative effect is mainly observed among the richest households, that is households with a father who has a higher-level occupation or is an employer (-10.1 pp). These estimates are statistically different from zero at the 10% level.

Figure 4 (left panels) shows the equivalent point estimates and confidence intervals for tertiary education. The selection effect displays a similar pattern in all three countries; it is strictly negative for the richest households, and close to zero among the poor households. However, its magnitude varies considerably across countries. Overall, the effect is -1.2 p.p in DR Congo, -3.1 pp in Ghana, -2.3 pp in Senegal. Among the richest households, the effect is -1.6 pp in DR Congo, -7.0 pp in Ghana, and -8.5 pp in Senegal. Thus, the proportion of men with tertiary education decreases in all three countries as a consequence of selection into migration, mainly by the richest households.

Finally, Figure 5 (left panels) shows the equivalent point estimates and confidence intervals for the number of years of schooling. The effect is negative in all three countries, yet is smaller in DR Congo (-0.1 years), compared to Ghana (-0.29 years) and Senegal (-0.30 years). The gap between rich and poor is particularly pronounced between the richest (-0.71 years) and the poorest group in Senegal (-0.08 years).

Second, consider women. There is hardly evidence of selection, except when focusing on years of schooling (Figures 3, 4 and 5, right panels). The decrease in women average years of schooling is -0.06 year in DR Congo, -0.13 in Ghana. In Senegal, the overall effect is zero, but could be negative for the richest households or the households with an unskilled employed father (the point estimates are -0.22 year and -0.40 respectively, not statistically fiferent from zero at the 10% level).[17]

[17]The subgroup "unskilled employee" stands out as a peculiar case, possibly due to its small size. It is the smallest group in the data 34 observations. See below.

6.2 Incentive and Net Effects (Households without a migrant)

The *incentive effect*, Δ_{inc}, measures the difference in the average schooling of the whole population, between the factual scenario and the counterfactual scenario in a *closed economy*. The range of values permitted by the model for Δ_{inc} is described by the worst-case bounds described in Section 4.2. Δ_{inc} is point identified or partially identified under the additional assumptions SI, LIV, or RB from Section 4. The *net effect* is the sum of the selection and the net effect.

The worst case bounds are presented first followed by the results from the matching and restricted bounds procedure. The discussion starts with the incentive effect and turns to the net effect. Because the LIV estimator appears to be biased, and has in most instances large standard errors, it is not discussed further.

6.2.1 Worst-case bounds

Starting with the incentive effect in upper secondary education, Figure 3 shows the estimates for the worst-case bounds on the incentive effect (yellow bold lines on the right in each father-occupation group) and corresponding 90% confidence interval (thin gray lines in the background). The average effect for all occupation groups is also displayed (first from the left).

First, consider men. In DR Congo, the incentive effect at the upper secondary is bounded below by zero in all subgroups. This implies that, in a scenario of a *closed economy*, all households would have invested equally or less in upper secondary education. When considering all households together, the increase in schooling due to migration prospects would not be larger than 7 pp in DR Congo. The picture is similar in Ghana, with a maximum increase of 11 pp.

In Senegal, the picture is substantially different. Considering all households together, the bounds range from -15.5 p.p to 30.0 pp The range varies according to the father's occupation. For example, the incentive effect for households with an unskilled employed father ranges between -17.8 pp to 14.5 pp By contrast, the incentive effect for households where the father has a higher-level occupation ranges between -1.3 p.p and 18.4 pp Hence, under the most pessimistic case (B-1), that is all those who attempted migration would have obtained some upper secondary education had they *not* attempted migration, poor households' investment in upper secondary education would have been between 17 and 21 pp higher in the closed economy than in the factual economy.

The estimation for men' tertiary education conveys a similar message (Figure 4, left panels). In DR Congo, the incentive effect is essentially positive. In Ghana, it ranges between -3.8 p.p and 8.0 p.p for the whole population with a marked difference between the richest (-2.5 pp to 21.5 p.p) and the poorest household (-3.5 pp to 6.1 pp). In Senegal, the bounds estimates are tilted toward the negative region, ranging for the whole population between -26.0 pp and 12.5 pp.

Finally, for men' years of schooling, the bounds are almost everywhere centered around zero. Their amplitude varies importantly from one country to another, and from rich to poor households (Figure 5, left panels). For example, considering the whole population, the magnitude of the average effect is at the maximum about 0.9 year in DR Congo, 1.3 years in Ghana, and 3.2 year in Senegal.

Second, consider women. The patterns remain very similar, except for somewhat smaller magnitudes. In particular, in Senegal, incentive effects are essentially close to zero or negative. For example, the average incentive effect at the tertiary level ranges between -15.8 pp and 0.5 pp

when the father is a skilled employee, or between -17 pp and zero when the father is an unskilled employee.

By construction, the bounds on the *net effect* are a translation of the bounds on the incentive effect. Where the selection effect is close to zero, for example, for less well-off households, both bounds are very similar. Where the selection effect is strictly negative, for example, for the richest households, the bounds on the incentive effect are translated downward to give the bounds the net effect. In Figures 3, 4 and 5, the bounds estimates are represented as the blue thick lines in the center of each father's-occupation subgroup, and the corresponding 90% confidence interval as the thin gray lines in the background.

In DR Congo, the net effect at the upper secondary level or the tertiary level is zero or positive in all subgroups. For the years of schooling, the bounds are centered around zero, with similar magnitude as the incentive effect.

In Ghana, the net effect at the upper secondary level is zero or positive in almost all subgroups. At the tertiary level, the bounds are centered around zero with a magnitude ranging between -6.2 p.p and 5.5 p.p, for all men, and -3.4 pp and 1.7 pp for all women. For the years of schooling, the bounds estimate are predominantly in the negative region, suggesting a zero or negative effect, that can be as low as -0.6 years for men and -0.25 years for women.

Finally in Senegal, for poor households, the net effect mirrors pretty closely the incentive effect. For the richest household, the bounds estimates are centered around zero at the upper secondary and tertiary levels. They are in the negative region for the years of schooling, suggesting a net loss.

The worst-case bounds give a somewhat commensurate perception of the incentive and net effects. However, they still include most extreme cases. The additional estimation strategies select a subset from these bounds, excluding cases that violate some plausible set of assumptions. The results from the matching procedure and the restricted bounds are presented in the following section.

6.2.2 Matching

Starting with the incentive effect, first consider men. In DR Congo, there is some positive incentive effect for rich households, and a zero effect for poor households (unskilled employed or self-employed without employee). The effect is modest at the upper secondary education level because of the already high enrollment rate (Figure 6, left panels). The effect is more sizable at the tertiary level: a 3 pp increase. This implies that compared to the *closed economy* scenario, investment in tertiary education is higher by 3 pp in DR Congo. This is due to the large incentive effect among the richest households: 9 pp (Figure 7, left panels). For the years of schooling, the richest households invest in 0.8 additional year of schooling.

For Ghana, the findings are very similar to those for DR Congo. In particular, the measured incentive effects for the richest households imply a 5 pp increase on average at the upper secondary level, a 9 pp increase at the tertiary level. For the number of years of schooling, the richest households invest in 0.8 additional year of schooling.and one year increase as for the years of schooling.

In Senegal, the picture is more contrasted (Figures 6, 7 and 8). Although, not statistically significant, the point estimates suggest positive incentive effects on the richest households at

the upper secondary (7 pp) and the tertiary education levels (9 pp), and for years of schooling (0.8 years). By contrast, in the remaining households, the estimates suggest the possibility of zero or even negative incentives at the upper secondary level. The possibility of negative incentive effects corresponds to a pattern uncovered at the secondary level (see Section 5.4; more details in Section E.4), and is recurrent finding of the empirical analysis on Senegal.

There is hardly any sizable effect for women, except when looking at the number of years of schooling of richest households in DR Congo (0.33 year), Ghana (0.19 year) and Senegal (0.26 year). If anything, the estimates suggest that rich households would invest more in schooling because of migration prospects.

In general, the estimates suggest zero net effects. Therefore, incentives compensate for the selection observed among the richest households. There are two exceptions. First, in DR Congo and in Ghana, incentives might even dominate selection, producing a positive net effect on the years of schooling (0.6 years). Second, in Senegal, where incentives are negative, net effects are also negative (for example, Figure 6).

6.2.3 Restricted bounds

Starting with the incentive effect, first consider men. In DR Congo, the restricted bounds suggest small effects on upper secondary and tertiary education, with similar effects across different household types (Figures 9 and 10). Bounds on the effects on years of schooling are still centered around zero, with larger amplitude for the richest households (Figure 11). The magnitude of the average effect is at the maximum about 0.2 year.

In Ghana, the bounds suggest small effects on upper secondary education (Figure 9). At the tertiary level (Figure 10), incentive effects typically range between -5 p.p and 3 p.p, except for the richest households (-2.5 p.p and 7.5 pp). Bounds on the effects on years of schooling are centered around zero (Figure 11). The magnitude of the average effect is at the maximum about 0.4 year.

In Senegal, the richest households have zero or positive incentives, while the remaining households have zero or negative incentive effects (Figures 9, 10 and 11). For example, at the high school level, the bounds are bounded below by zero for the richest households, while the bounds estimates are negative in all remaining households. Moreover, large negative incentives are not excluded: pooling together all father's-occupation groups, except the "higher-level occupation", the lower bound estimate is -14.5 pp at the upper secondary level, -21.3 pp at the tertiary level, and -0.8 years as for the average years of schooling.

Second, consider women. In DR Congo and Ghana, there is hardly any sizable effect, except for the years of schooling. In this case, the magnitudes are smaller than for men. Again, the size of the bounds differ between the richest households and the rest of the population.

For women in Senegal, bounds estimates also point to negative incentives at upper secondary and tertiary levels, except in the richest households, and households with an unskilled employee father. The subgroup "unskilled employee" stands out as a peculiar case. It is the smallest group in the data 34 observations. When pooling together all father's-occupation subgroups, except the "higher-level occupation", the bound estimate range between -7.4 pp and -1 pp at the upper secondary level, -9.8 pp and -2.4 pp and at the tertiary level, and -0.29 and -0.15 year as for the average years of schooling. For the richest group, the bounds are centered around zero. Thus, the pattern for women is similar to the pattern uncovered for men.

How does it affect the net effect? In DR Congo, the restricted bounds estimates imply zero net effect on upper secondary and tertiary education, for all gender and father-occupation subgroups (Figures 9 and 10). In other words, the average proportion of non-migrants with upper secondary or tertiary education is the same as it would have been in the *closed economy*. However, the net effect on years of schooling might be negative, in particular for the richest households (bounds estimates give -0.31 to 0.04 year, for men, and -0.21 and 0.07 year, for women - Figure 11).

In Ghana, the restricted bounds estimates imply zero net effect on upper secondary education (Figure 9). With respect to tertiary education, the net effect is essentially zero or negative for most gender and father-occupation subgroups. For all men, it lies between -5.3 pp and -0.4 pp, and for all women, between -2.8 pp and 0.2 pp (Figure 10). With respect to the years of schooling, the net effect is also zero or negative for most gender and father-occupation subgroups (-0.47 to 0.03 year for men, and -0.17 to 0.02 for women - Figure 11).

Finally, in Senegal, the bounds suggest zero or (possibly strong) negative net effects, in almost all subgroups, at the upper secondary and tertiary education, and for the years of schooling (Figures 9, 10 and 11). Note that this is the combined result of strong selection among the richest households, and of negative incentives among the remaining households.

6.3 Short Summary

To summarize, in DR Congo, there is hardly any selection or incentive effect, when one looks at upper secondary and tertiary education. Still, there is the possibility of a slight net decrease or net increase in the average years of schooling of the richest households. However, the two estimation procedures do not agree.

In Ghana, selection and incentive effects are sizable at the tertiary level, essentially for rich households. Their respective magnitude lie essentially between -5 pp and 5 pp for men (8.9% of the average enrollment rate), and -2 pp and 2 pp for women (6.1% of the average enrollment rate), with strong effects on the richest households, and close to zero effects on the remaining households. Matching suggest zero or modest positive net effects; the restricted bounds suggest zero or modest negative net effects, with strongest effects on the richest households. The results are similar with respect to the years of schooling.

In Senegal, incentives to invest in schooling might be negative for men and women from less well-off households. These effects translate into negative net effects that could be quite large. For example, if the net effect for men in upper secondary education would be at the lower bound of the bounds estimates (-15 pp), this would represent a 31% decrease compared to the *closed economy* scenario. The matching estimate concludes to a 16% decrease. The corresponding estimate would be 25% and 19% decrease for women.

These findings are further discussed in the next section.

7 Discussion

So, what makes brain drain more likely? The empirical analysis suggests that country-specific context matters greatly. While DR Congo is not greatly affected by "Brain Drain", Ghana might experience relatively small effects on his average human capital, and Senegal relatively large effects. In particular, low structural levels of education in Senegal makes the Brain Drain issue

salient already at secondary and upper secondary education levels. Return migration does little to mitigate these effects (see Appendix E.2).

With respect to the microeconomic determinants, household economic status is an important determinant of selection and incentive, and hence, net effects. Rich households tend to experience a strong selection into migration. However, they also have the highest incentives to invest in schooling. By contrast, poor households experience little effects, except in Senegal, where the incentive effect appears negative. Besides, there is a substantial gender difference, with smaller migration effects on women's schooling. Finally, households with a migrant tend to experience a stronger selection, and in some cases, have also stronger incentives to invest in schooling (see Appendix E.1).

It is important to examine the possible mechanisms that could explain the observed results, as well as the limits of the present study. The gender gap in selection and incentive could be explained by the gender difference in migration motives. While men are more likely to report that they migrate for work related reasons or to improve their living conditions, the large majority of women report migrating because of household reasons.

Why would emigration prospects cause lower incentives to invest in schooling in Senegal? The literature evokes three potential reasons, possibly complementary but not exclusive. First, Senegalese migrants to Europe might be constrained to jobs with lower skill requirements than the one they would have obtained in Senegal. Therefore, they would need lower schooling in Europe than in Senegal.[18] The MAFE survey data contains information on respondents' job history, both for migrants and stayers. Considering migration as a treatment and occupation characteristics at age 35 and 40 as an outcome, an Average Treatment Effect on the Treated is measured using matching and difference-in-difference matching.[19] From this analysis, there is little support for the hypothesis that, after migration, Senegalese would obtain jobs with lower skills requirements than the one they would have obtained, had they stayed in Senegal.

Second, wage returns to education could be lower in Europe than in Senegal. It is important to note that the distribution of schooling level by household economic status suggests that, at the margin, returns to upper secondary education matter most to poor households, while returns to tertiary education matter most to rich households.[20] The MAFE survey data contains some information about wages, that allow approximating wage returns to schooling for Senegalese migrants to Europe and for non-migrants. Table 2 shows the results of this calculation.[21]

On average, a Senegalese with at most some secondary education earns a monthly wage in

[18] One example from Sub-Saharan Africa is given by Girsberger (2014), where migration from Burkina Faso to Côte d'Ivoire is mainly due to migration to work in cocoa plantations. As a consequence of low skill requirements for these jobs, children are more likely to drop out of school to emigrate as farm workers.

[19] In the dataset, occupations are coded and ranked using ISCO and ISEI codes.

[20] The proportion of children with exactly some upper secondary education is 13% for households with a father that is in the "higher-level occupation" subgroup, against 14% in the remaining households. At tertiary level, the same proportions are respectively 30% and 10%.

[21] Monthly wages are self-reported at the end of a job spell. The sample is restricted to job spells after 1980. Reported wages are converted into 2008 $US and further corrected for different Purchase Power Parities between Senegal and Europe using implied PPP from the IMF World Economic Outlook database. The sample is trimmed for the 5% lowest and highest values. The displayed results are for a Mincer regression, controlling for gender, age, age squared, time fixed-effects, and quadratic term for years since migration (for regressions on migrants). Unfortunately, for DR Congo and Ghana, similar analyses are unreliable. Multiple reported currencies and large exchange rate fluctuations overtime make it difficult to provide comparable wages after that correct for differences in inflation rates and PPP. Nyarko (2011) estimates the net present value of tertiary education for migrants and non-migrants in Ghana, and finds large income gains from high skill migration.

		Aver. Wage Prim./Sec. $US	Returns Upper Sec. (%)	$US	Returns Tertiary (%)	$US
All	Senegal	574	0.40	229	0.64	367
	(s.e.)	(55)	(0.04)		(0.13)	
	Europe	1200	0.12	138	0.56	672
	(s.e.)	(33)	(0.02)		(0.04)	
Poor households	Senegal	507	0.25	127	0.84	427
	(s.e.)	(33)	(0.08)		(0.32)	
	Europe	1175	0.07	80	0.50	583
	(s.e.)	(35)	(0.03)		(0.07)	
Rich households	Senegal	715	0.17	123	0.31	223
	(s.e.)	(35)	(0.20)		(0.29)	
	Europe	1308	0.14	179	0.54	700
	(s.e.)	(35)	(0.07)		(0.08)	

Table 2: Monthly wage returns to schooling for Senegalese in Senegal and in Europe

Europe (1200$US) that is twice as large as his potential wage in Senegal (574$US). Relative returns to upper secondary and tertiary education are higher in Senegal than in Europe (respectively 0.40 and 0.64, against 0.12 and 0.56). Moreover, absolute wage returns to education are higher in Senegal (229$US) than in Europe (138$US). However, absolute wage returns to tertiary education in Europe (672$US) are almost double those in Senegal (367$US).

Splitting the sample by father's-occupation subgroups shows that returns to schooling in Europe dominate those in Senegal for rich households.[22] Absolute returns to tertiary education in Europe are even three times those in Senegal. For poor households, relative returns are clearly larger in Senegal, while absolute returns are fairly comparable. Furthermore, for poor households, returns to schooling represents a smaller fraction of the wage gain from migration (at the upper secondary level, 11%), compared to the same fraction for rich households (30%).

Therefore, it seems plausible that high wage returns to (tertiary) education in Europe influence rich households at the margin, and triggers more investment in education. By contrast, returns to upper secondary schooling seem too low to give additional incentives to poor households. In fact, if poor households compare relative returns to education in Senegal and in Europe that could explain the negative incentive effects, at both upper secondary and tertiary level.[23]

Third, the discrepancy between rich and poor households could come from liquidity constraints for poor households with an emigration option. Appendix A shows that when the budget constraint is binding, migration prospects do not provide additional incentives to obtain further education. Moreover, the household might *substitute* the migration investment to the schooling investment:

[22]There are relatively few "higher-level occupation" households and from these, few have a child with secondary education or below. Hence, estimates of the returns to education are very imprecise, when splitting the sample as done before. Here, rich households are those from the "skilled employee" and "higher-level occupation" groups.

[23]The Borjas-Roy model suggests that households compare relative wage returns rather than absolute wage returns (Borjas, 1987). In Equation (7), this would mean that Π has a log-linear approximation. It is well known that relative wage returns to education are higher in Sub-Saharan African countries than in OECD countries (Kuepie, Nordman, and Roubaud, 2009). The hypothesis of a log-linear utility function is disputed in the context of migration from developing to developed countries. Grogger and Hanson (2011) note that "given the vast income differences that exist between countries [...] [the] linear utility appears to abuse reality less than the strong curvature of the log-linear utility". Batista, Lacuesta, and Vicente (2012) also favor the idea that absolute returns to education matter more than relative returns, in the case of Cape Verde.

candidates for migration drop out of school earlier in order to enter the labor market, and accumulate capital that they will invest in a migration investment. The household Survey "1-2-3" conducted in Dakar in 2001-2002, reveals that along with school failures, financial constraints and preference for labor market opportunities account for the most common reason for secondary school drop out (32%) (De Vreyer and Roubaud, 2013).

As for the limits of the present study, it is important to recall that the exposed generalized Roy model is intended as an approximation of households' schooling decisions. The present exposition does not deal with uncertainty in schooling outcomes, although it could be extended to encompass some of it. Besides, it does not distinguish between attempts to migrate legally or illegally. This is mainly due to data restrictions: it is not possible to observe whether an attempt is through legal or illegal ways.[24]

More importantly, the characterization of the counterfactual schooling investment relies on the assumption that there exists no peer effect or social interaction. This assumption is also known as the Stable Unit Treatment Value Assumption (SUTVA). If in the *closed economy*, there is overall less schooling investment (as suggested for DR Congo and Ghana), and hence, less social pressure for schooling investment, the incentive effect created by migration is actually larger than the one measured in this study. Conversely, if in the *closed economy*, there is overall more schooling investment (as suggested for Senegal), and hence, more social pressure for schooling investment, the incentive effect is actually smaller than the one measured in this study. Thus, relaxing the SUTVA might well reinforce the above conclusions.

Furthermore, the model takes the attempt decision as observable at the time of the schooling investment decision. As discussed earlier, because there does not exist better data sources, the migration attempt observed after schooling completion is used as an approximation. In the case of men in Senegal, the magnitude of our estimate (not their sign) proves sensitive to our definition of a migration attempt, even if the main conclusions remain unchanged (see Appendix E.4). Hence, there is scope for improvement, by obtaining better information on attempt decisions at the time of the schooling investment.

Finally, although very rich in information, the MAFE survey data is limited to a sub-sample of countries in Europe. As long as migrants to other European destinations do not differ too much in their schooling decision from migrants to the observed major destinations, the present results would remain valid.[25]

8 Conclusion

Education and migration are two important (and often irreversible) human capital investments. The main insight of this study is that the relationship between the two varies importantly with institutional context and household characteristics. The two investments can be fairly independent as for households in DR Congo and for poor households in Ghana, complementary as for rich households in Ghana and Senegal, or even substitute, as for poor households in Senegal.

[24] The MAFE survey allows observing residence status only after successful completion of the migration. From the sample, estimation suggest that 28% of Congolese and 31% of the Senegalese migrants had no proper visa or resident permit at the time of their first stay abroad, against 4% in Ghanaian. On illegal migration from Senegal, see Mbaye (2014).

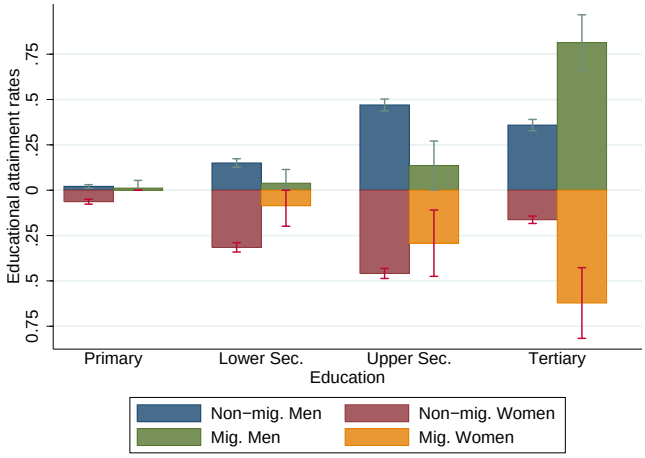
[25] Missing some destination countries is problematic, even if the schooling investments are similar, when the focus is on *absolute measures* (see Appendix E.3).

Compared to the existing literature, this study enriches the analysis of the effects of migration prospects on origin countries' human capital by differentiating the *selection, incentive* and *net effects* at the household level. Appealing to a generalized Roy model of schooling investment with a migration option, it provides simple characterizations of these effects, so that existing econometric tools can be fruitfully applied for their estimation (matching, LIV, bounds).

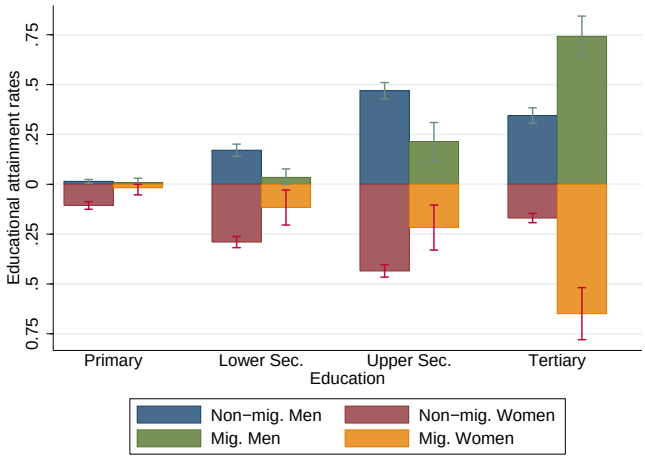
An important insight gained from this study is that reliable information on migration plans and attempts are as valuable to answer interrogations around brain drain as are information about actual migration spells and skills of migrants.

With respect to the debate on the brain drain in Sub-Saharan Africa, the empirical analysis provides interesting insights. It suggests that DR Congo is mainly unaffected, the effect on Ghana is fairly limited, and Senegal might be losing an important part of its high skill population. The finding on Senegal is a combination of the selection of high skilled into migrants, and of negative incentives to invest in schooling among the poor households. Return migration does little to mitigate these effects.

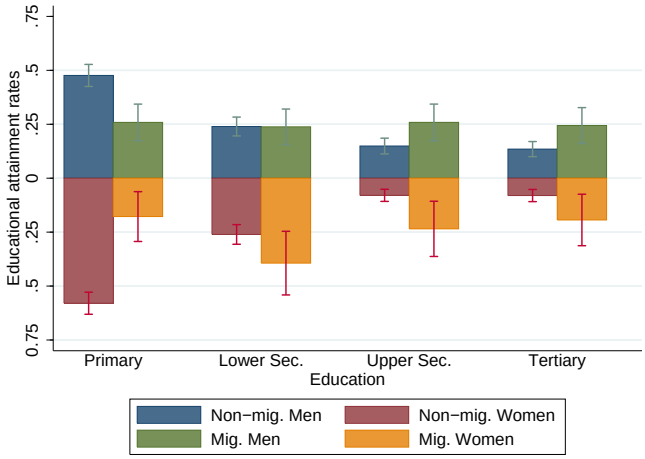
These findings would be most useful if seen as directions for further scientific inquiry, and least if seen as support for strong policy recommendations. In Senegal, which is the most "worrying" case, the motivation for poor households to invest less in education when the migration option exist should be clarified. It could be an economically rational and optimal choice, or the result of credit market imperfections. Each mechanism implies a different policy response.



(a) DR Congo



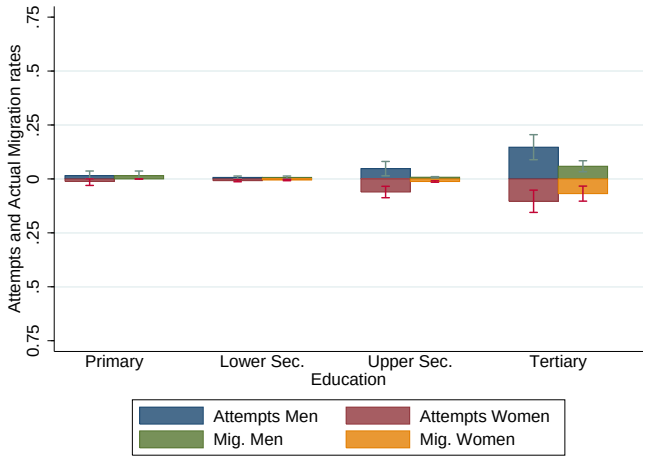
(b) Ghana



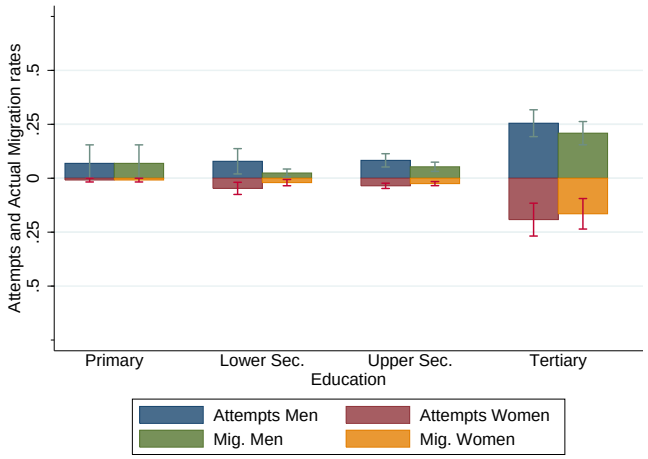
(c) Senegal

Note: Estimated distribution of education, conditional on migration status, and gender. Education: 'Primary' - Primary or Less than Primary, 'Lower Sec.' some Lower Secondary - 'Upper Sec.': some Upper Secondary, and 'Tertiary' - some Tertiary. The bars represent the 90% Confidence Interval.

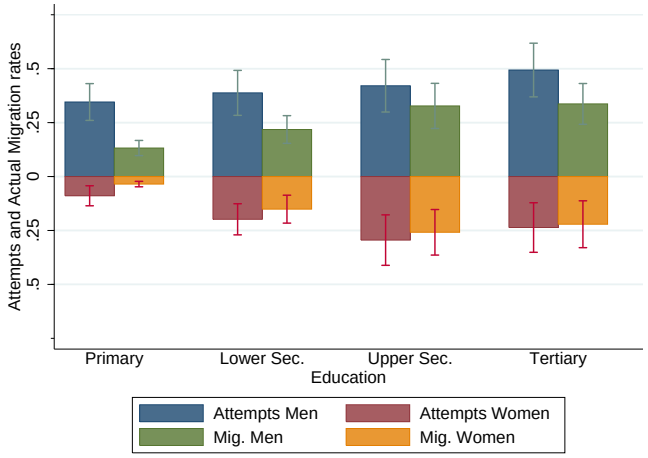
Figure 1: Education by country, gender, and migration status



(a) DR Congo



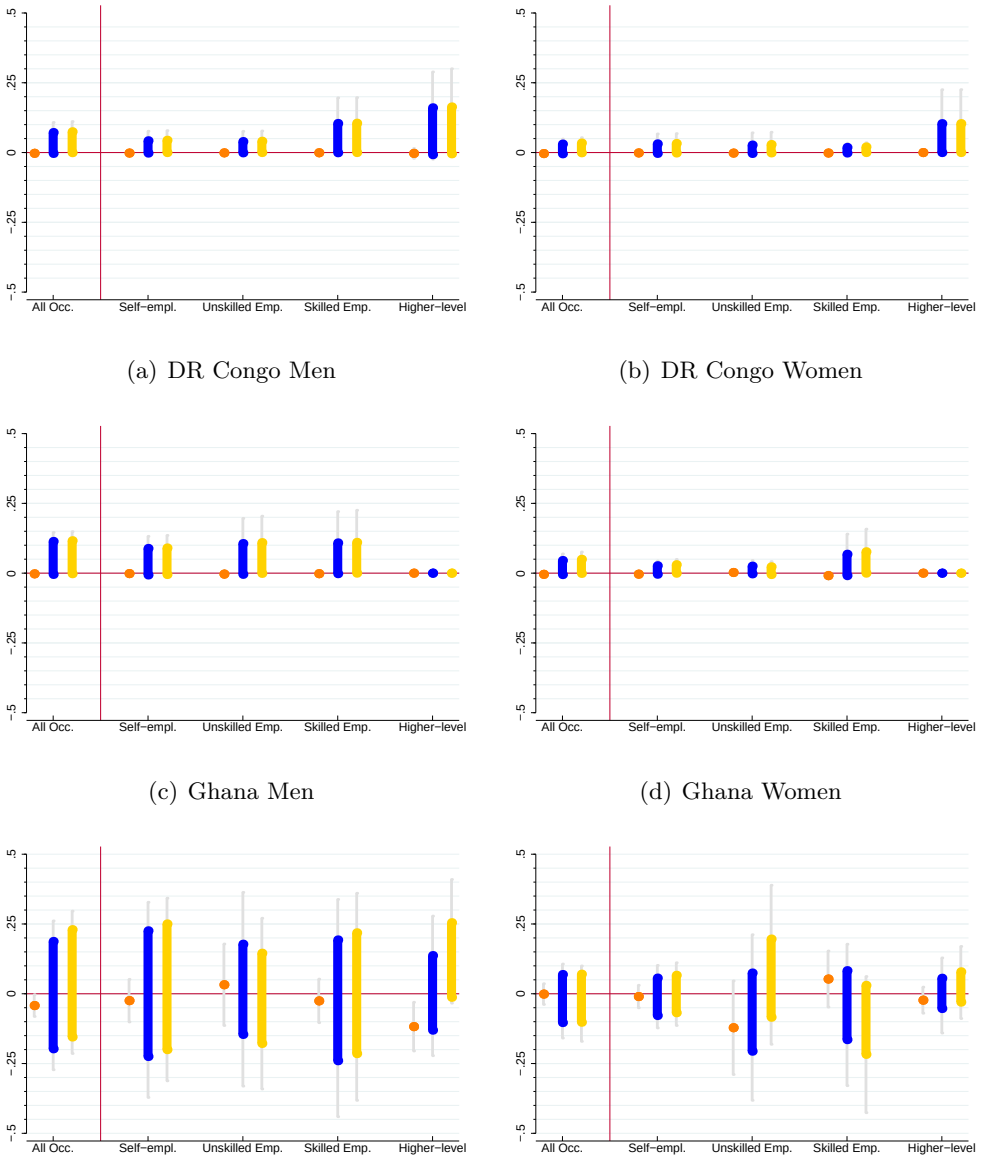
(b) Ghana



(c) Senegal

Note: Estimated migration attempt and actual migration rates, conditional on education, and gender. Education: 'Primary': Primary or Less than Primary, 'Lower Sec.' some Lower Secondary, 'Upper Sec.': some Upper Secondary, and 'Tertiary': some Tertiary education. The bars represent the 90% Confidence Interval.

Figure 2: Migration attempt and actual migration rates by country, gender, and education



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

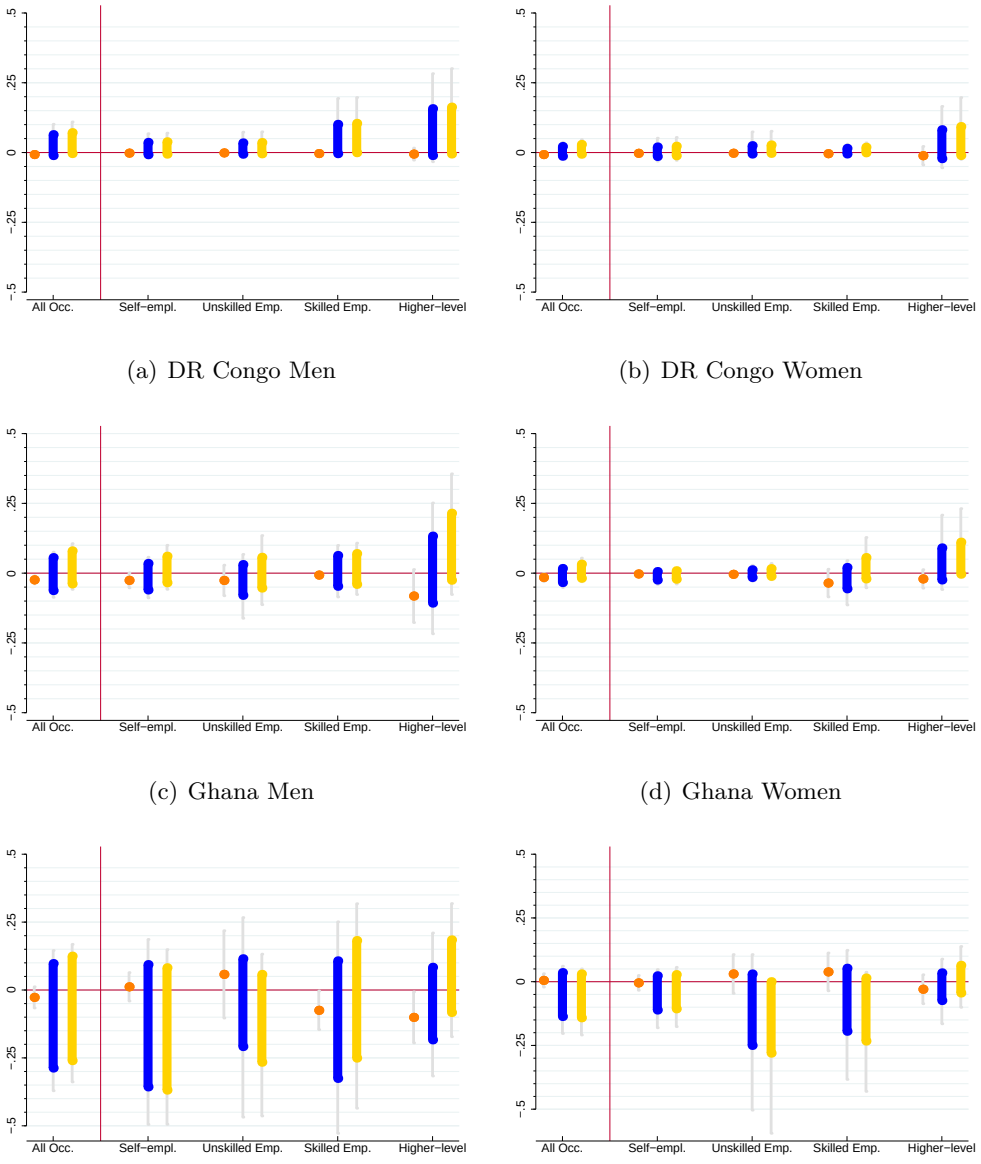
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 3: Estimated effects at upper secondary level with worst-case bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

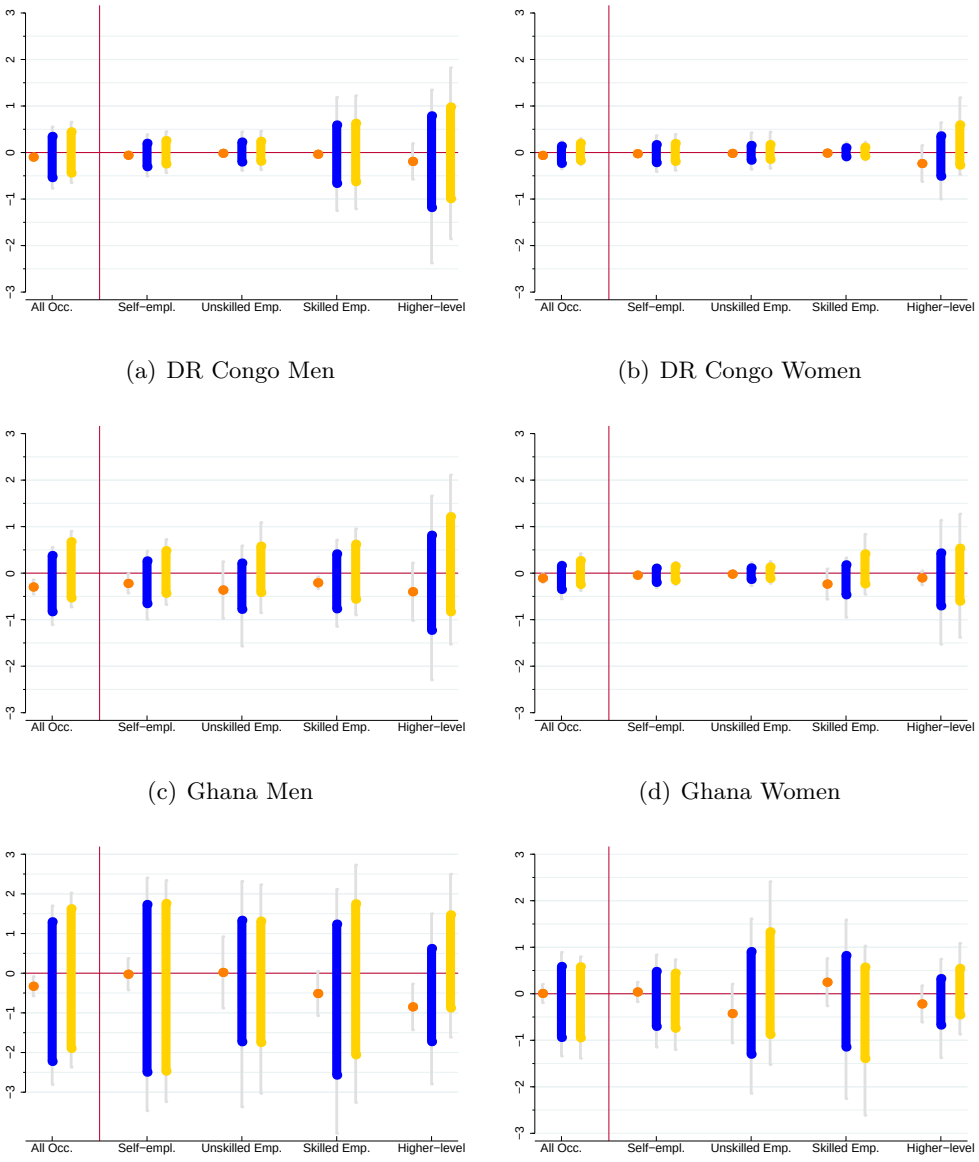
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 4: Estimated effects at tertiary level with worst-case bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

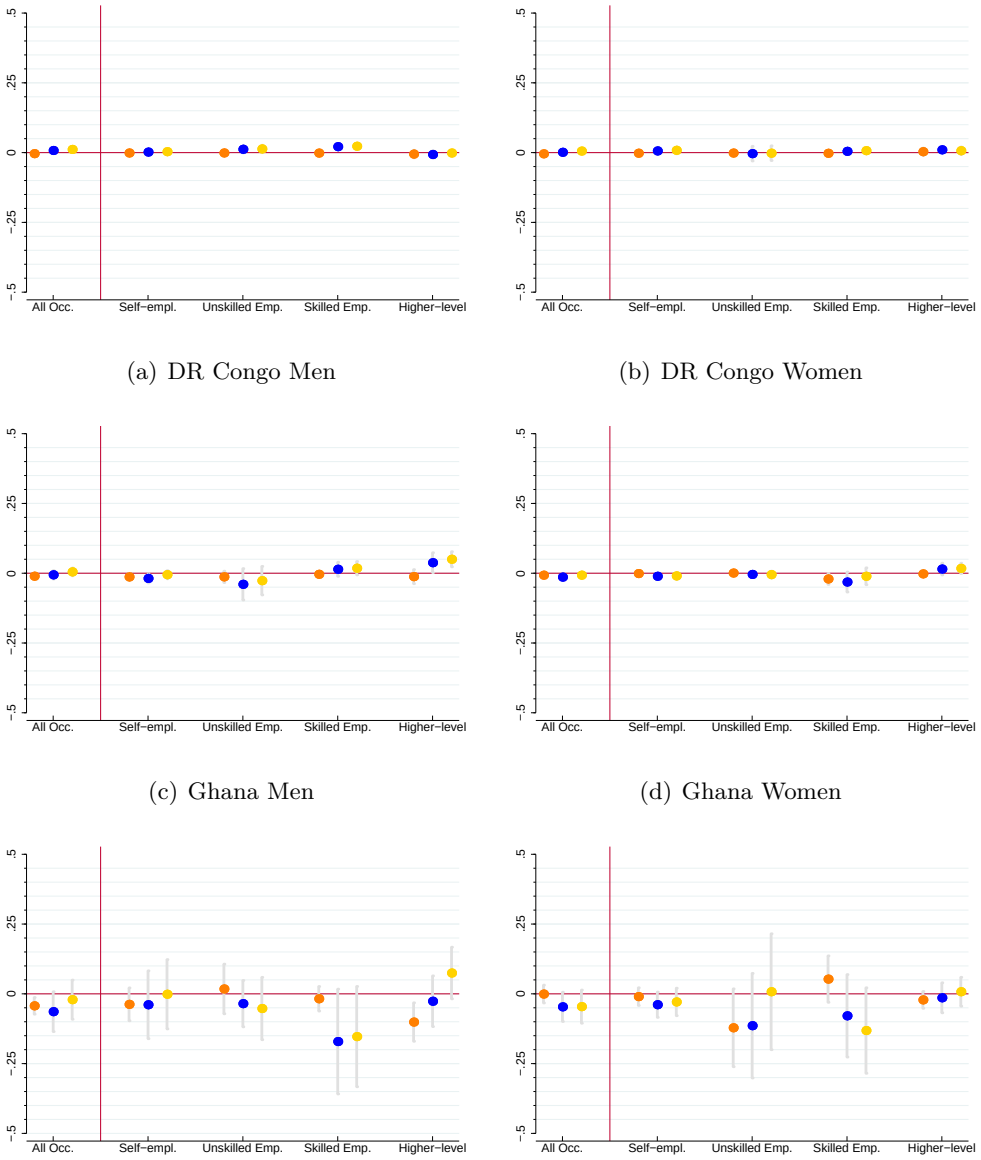
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 5: Estimated effects for years of schooling with worst-case bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

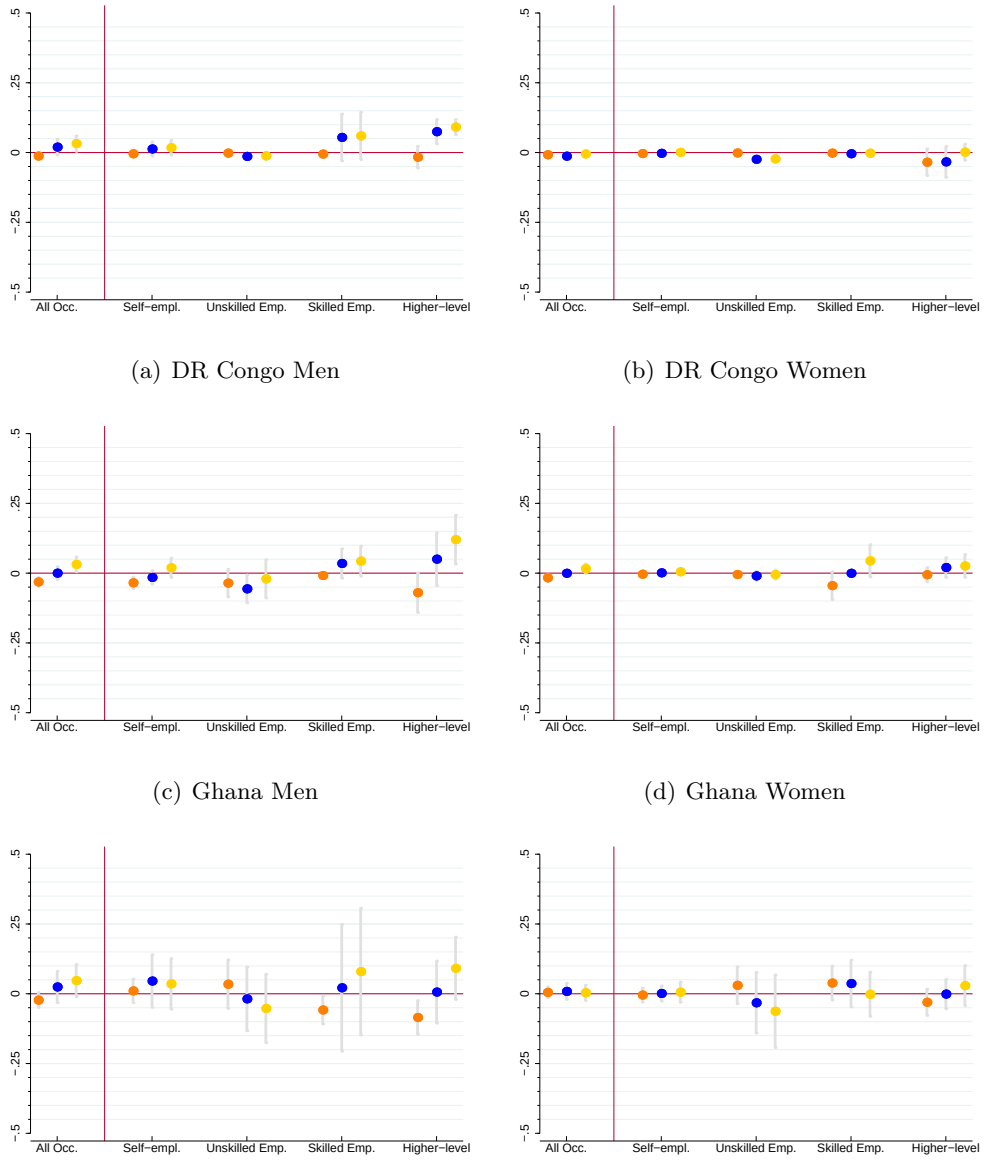
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 6: Estimated effects at upper secondary level with matching. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

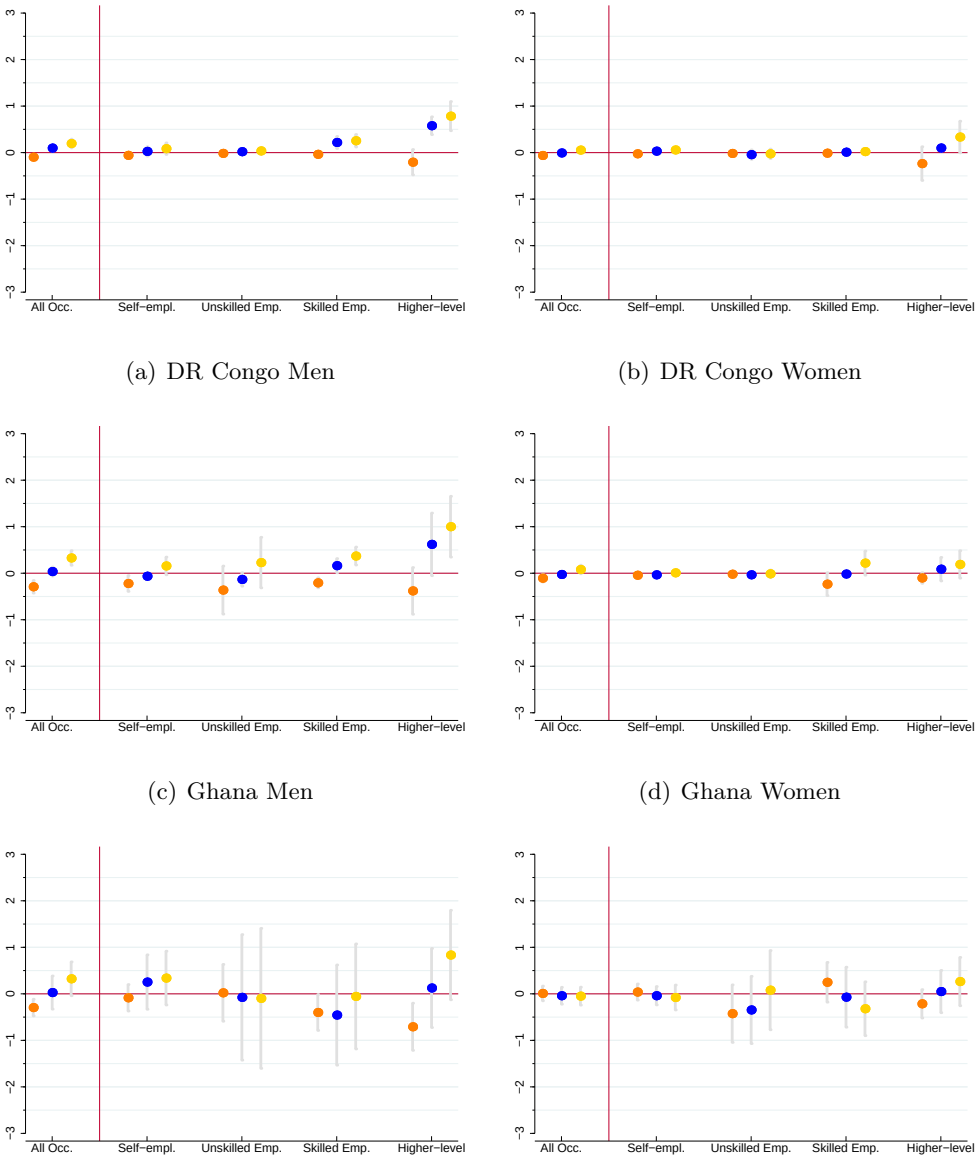
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 7: Estimated effects at tertiary education level with matching. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

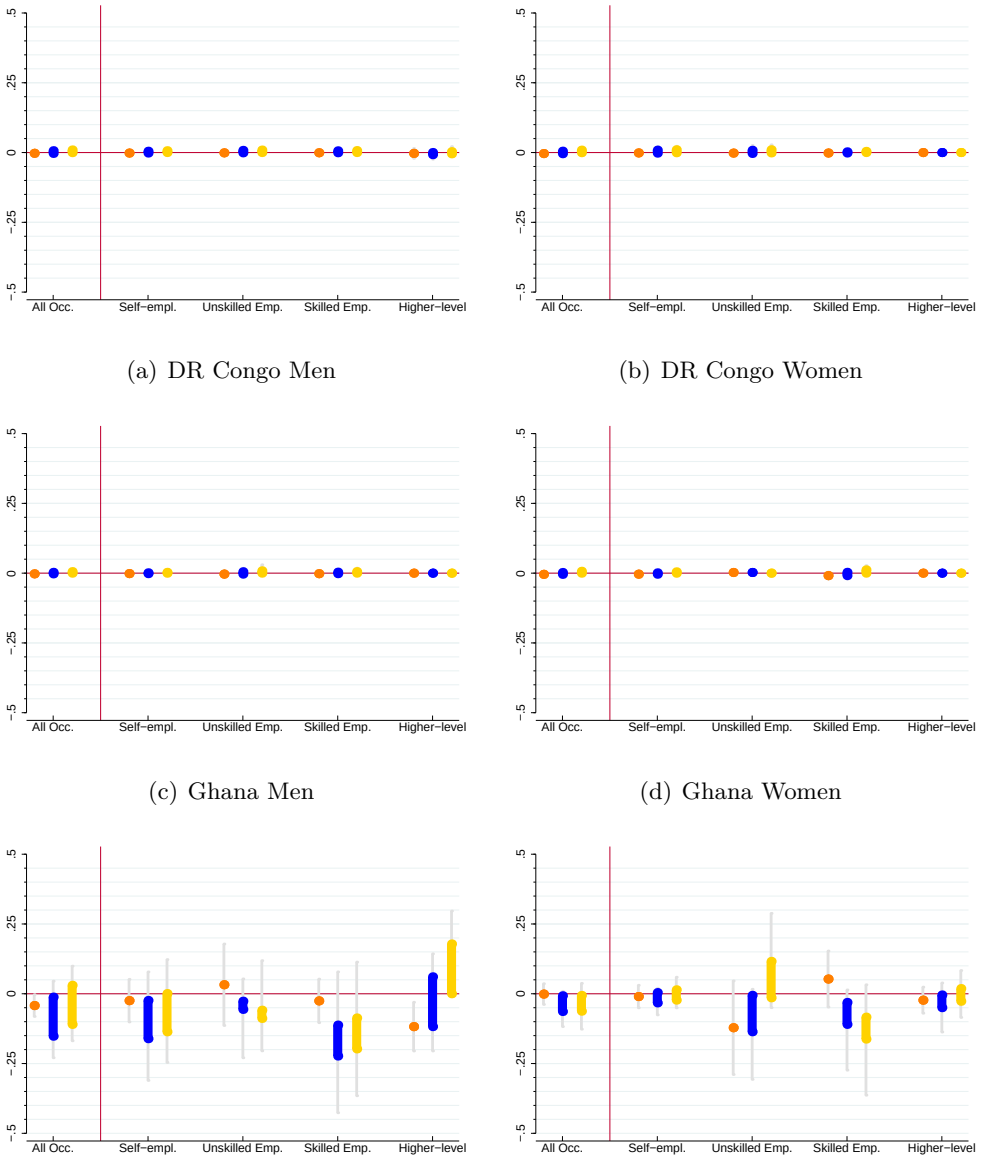
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 8: Estimated effects for years of schooling with matching. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

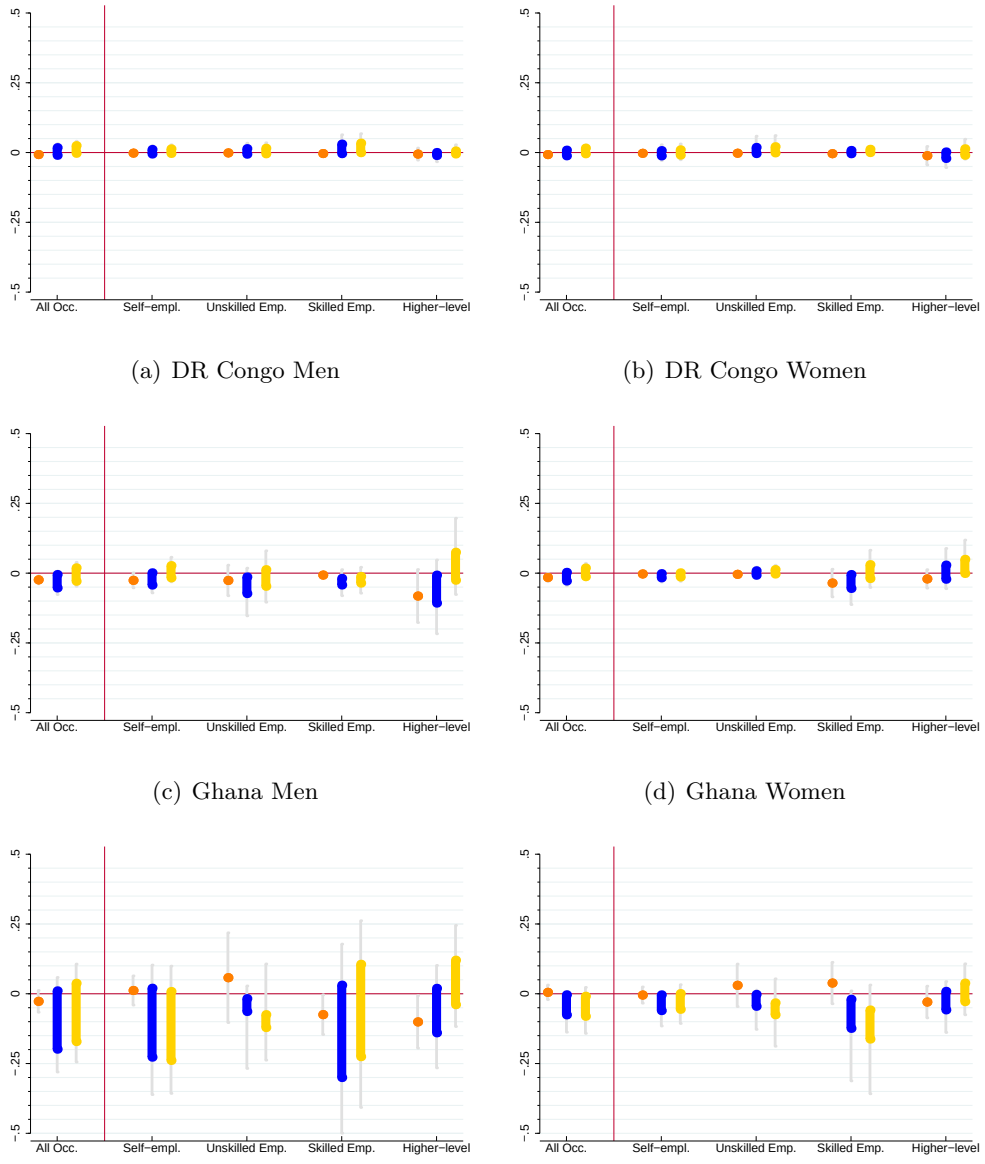
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 9: Estimated effects at upper secondary level with restricted bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

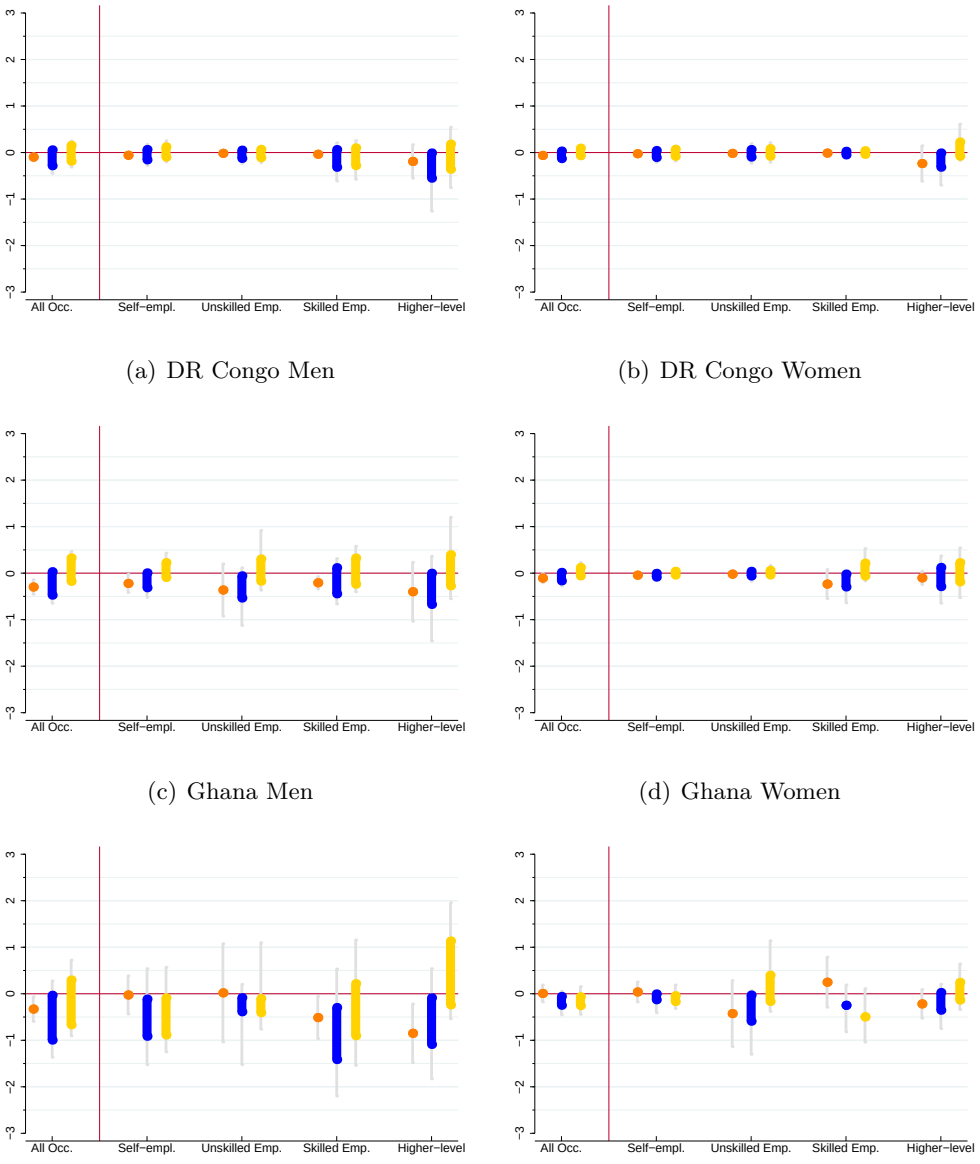
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 10: Estimated effects at tertiary level with restricted bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

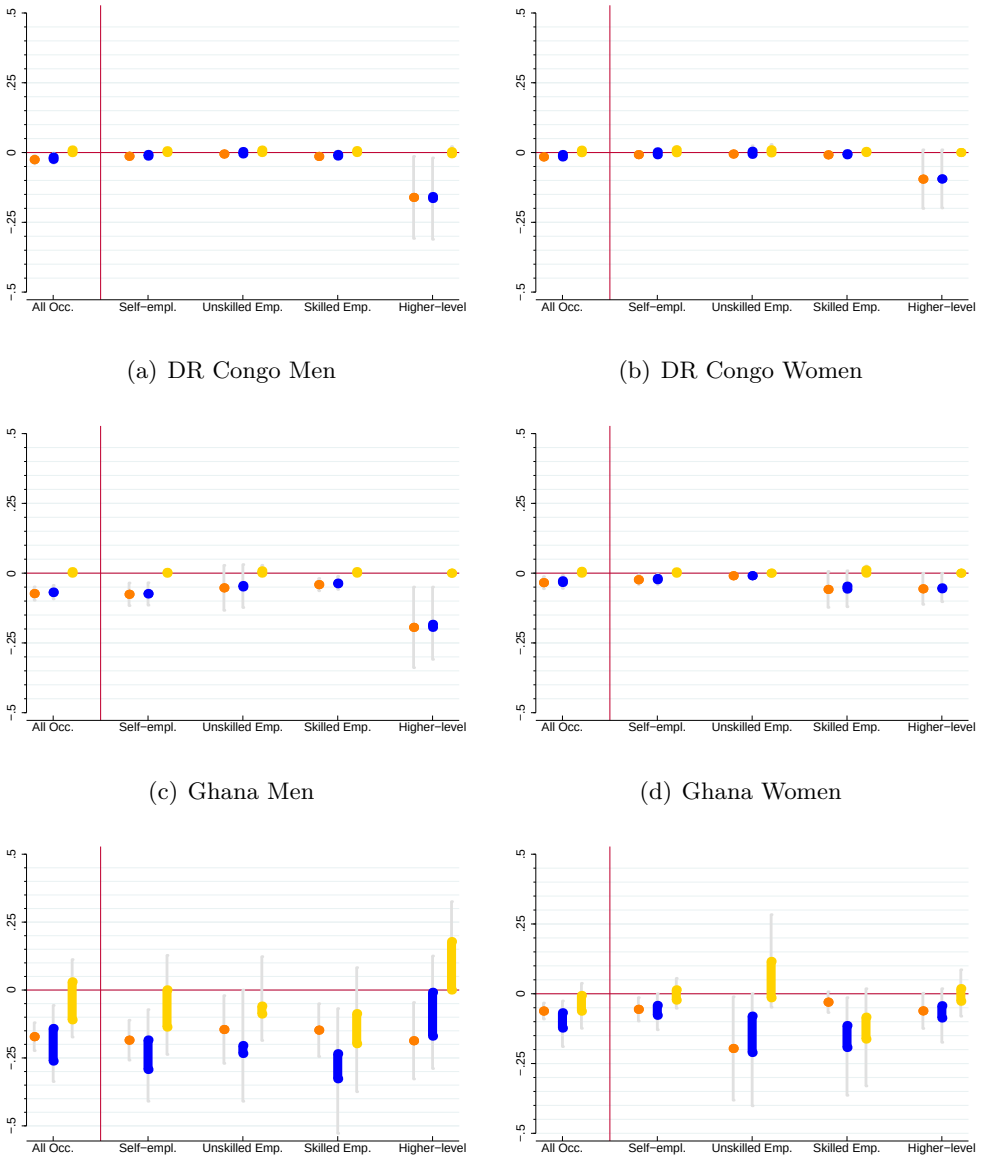
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 11: Estimated effects for years of schooling with restricted bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

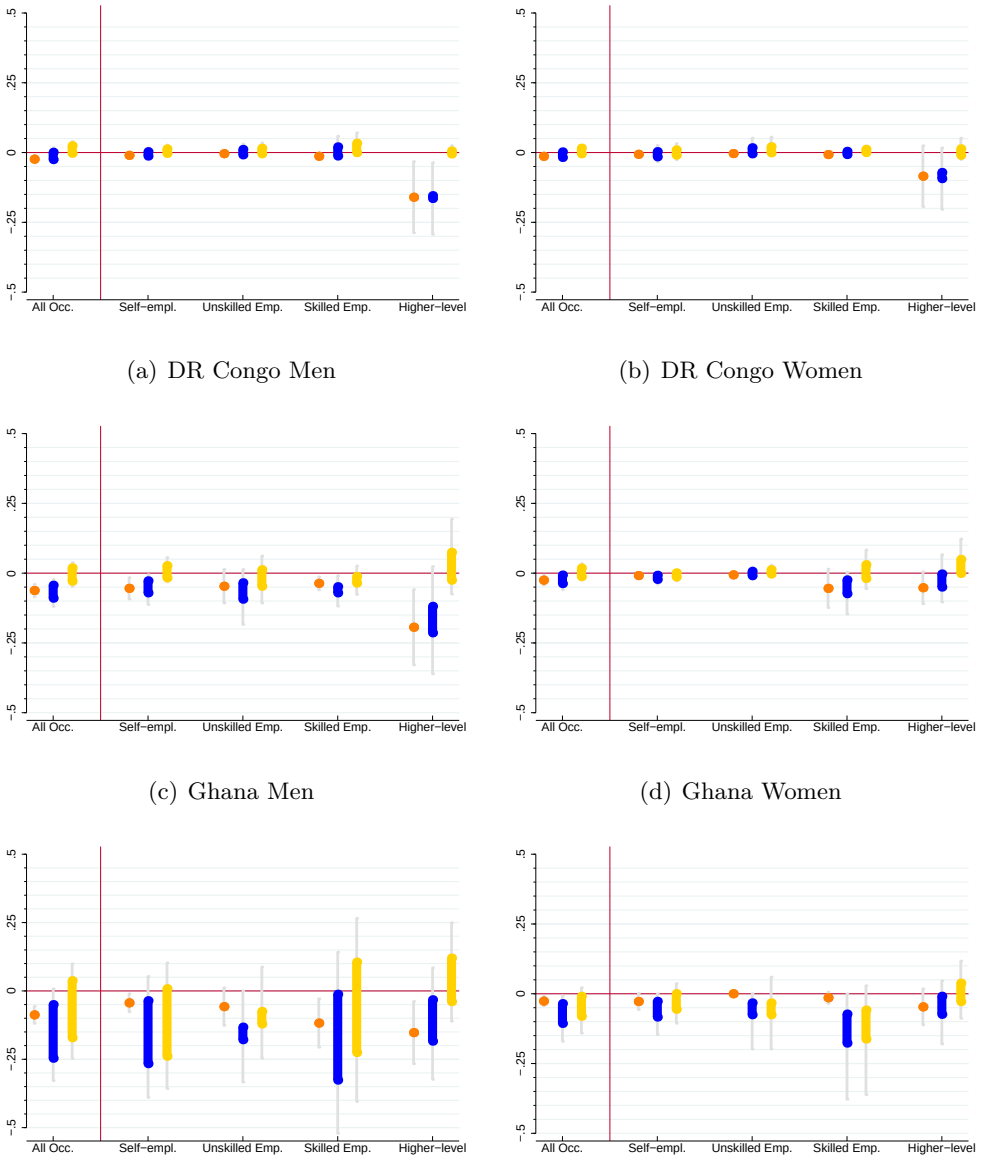
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ^{abs}_{sel} (orange), Δ^{abs}_{net} (blue), and Δ^{abs}_{inc} (gold). Households without a migrant.

Figure 12: Absolute estimated effects at upper secondary level with restricted bounds. Households without a migrant.



(a) DR Congo Men

(b) DR Congo Women

(c) Ghana Men

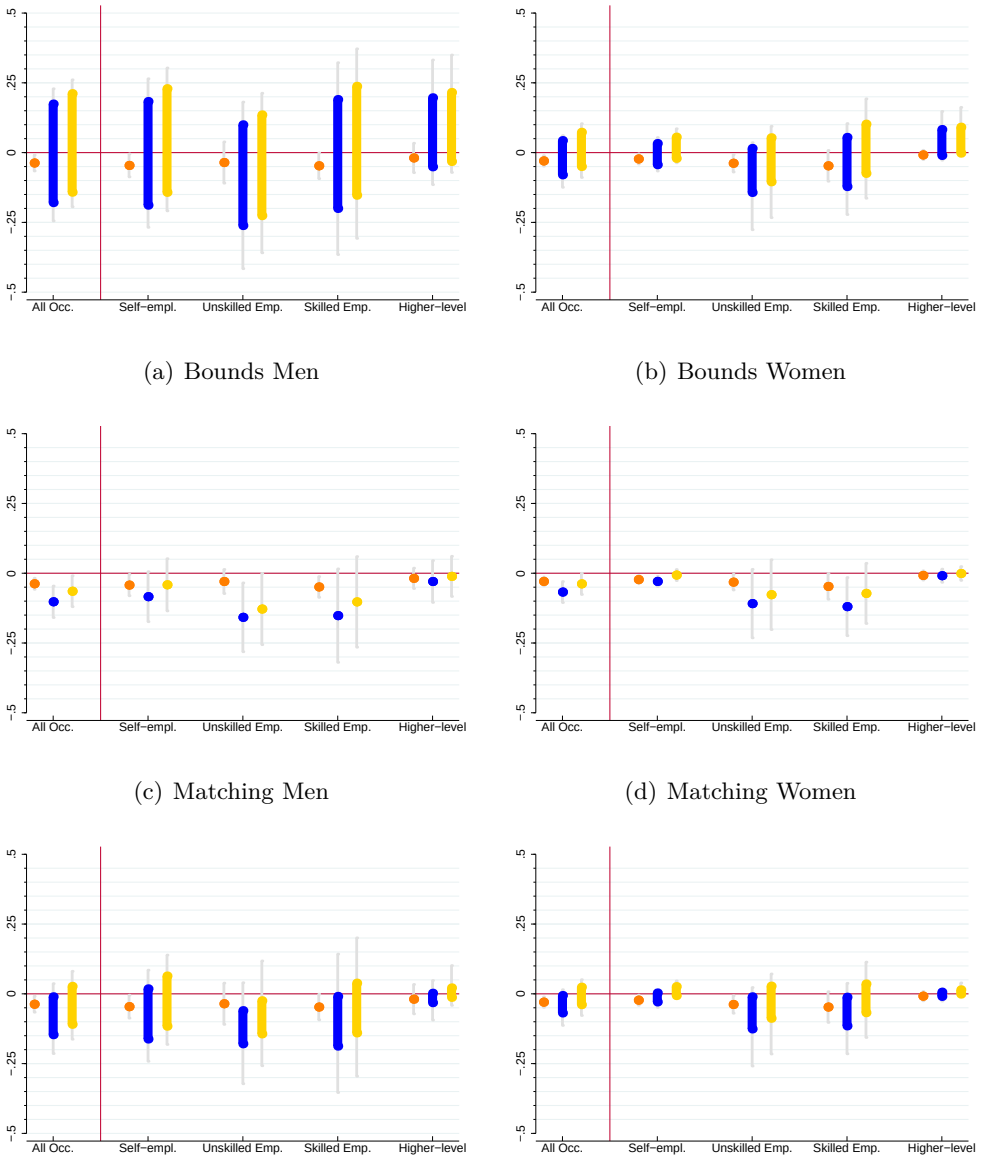
(d) Ghana Women

(e) Senegal Men

(f) Senegal Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ^{abs}_{sel} (orange), Δ^{abs}_{net} (blue), and Δ^{abs}_{inc} (gold). Households without a migrant.

Figure 13: Absolute estimated effects at tertiary level with restricted bounds. Households without a migrant.



(a) Bounds Men

(b) Bounds Women

(c) Matching Men

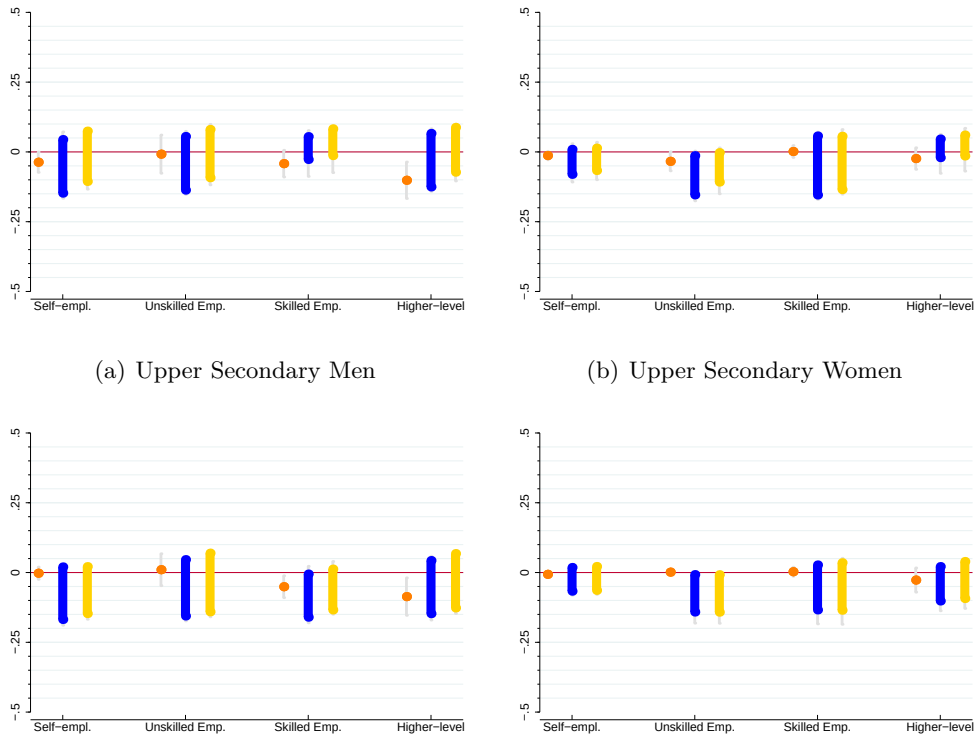
(d) Matching Women

(e) Restricted bounds Men

(f) Restricted bounds Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 14: Estimated effects at secondary level. Households without a migrant.



(a) Upper Secondary Men

(b) Upper Secondary Women

(c) Tertiary Men

(d) Tertiary Women

Note: Father's occupation, from left to right: 'All Occ.': All occupation groups; 'self-empl.': self-employed without employee or unemployed; 'unskilled empl.': unskilled employee or laborer, 'Skilled empl.': skilled employee or laborer, 'High-level': Intellectual/Higher-level occupation or employer. In each father's-occupation group, from left to right: Δ_{sel} (orange), Δ_{net} (blue), and Δ_{inc} (gold). Households without a migrant.

Figure 15: Estimated effects with Instrument in Senegal. Households without a migrant.

References

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996): "Identification of Causal Effects using Instrumental Variables," *Journal of the American statistical Association*, 91(434), 444–455.

Baizán, P., C. Beauchemin, and A. González-Ferrer (2013): "Determinants of Migration between Senegal and France, Italy and Spain," MAFE Project Working Paper 25.

Batista, C., A. Lacuesta, and P. C. Vicente (2012): "Testing the 'Brain Gain' Hypothesis: Micro Evidence from Cape Verde," *Journal of Development Economics*, 97(1), 32–45.

Beauchemin, C. (2012): "Migrations between Africa and Europe: Rationale for a Survey Design," MAFE Methodological Note 5 accesssed at: http://mafeproject.site.ined.fr/en/methodo/methodological_notes/.

Beauchemin, C., P. Sakho, B. Schoumaker, and M.-L. Flahaux (2014): "Changing Pattern of Senegalese Migration," MAFE Project Working Paper 21.

Beine, M., F. Docquier, and H. Rapoport (2001): "Brain Drain and Economic Growth: Theory and Evidence," *Journal of Development Economics*, 64(1), 275–289.

——— (2008): "Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers," *The Economic Journal*, 118(528), 631–652.

Bhagwati, J., and K. Hamada (1974): "The Brain Drain, International Integration of Markets for Professionals and Unemployment: a Theoretical Analysis," *Journal of Development Economics*, 1(1), 19–42.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 77(4), 531–53.

Chand, S., and M. A. Clemens (2008): "Skilled Emigration and Skill Creation: a Quasi-experiment," Center for Global Development Working Paper 152.

Chernozhukov, V., W. Kim, S. Lee, and A. Rosen (2015): "Implementing Intersection Bounds in Stata," *Stata Journal*, 15(1), 21–44.

Chernozhukov, V., S. Lee, and A. Rosen (2013): "Intersection Bounds: Estimation and Inference," *Econometrica*, 81(2), 667–737.

Clemens, M. A. (2007): "Do Visas Kill? Health Effects of African Health Professional Emigration," Center for Global Development Working Paper 114.

——— (2011): *The Financial Consequences of High-Skill Emigration: Lessons from African Doctors Abroad.* in (Sonia Plaza and Dilip Ratha, eds.), *Diaspora for development in Africa*, pp. 165-182, World Bank Publications.

De Vreyer, P., and F. Roubaud (2013): *Urban Labor Markets in Sub-Saharan Africa.* World Bank Publications.

Dinkelman, T., and M. Mariotti (2015): "What Are the Long Run Effects of Labor Migration on Human Capital? Evidence from Malawi," unpublished manuscript.

Easterly, W., and Y. Nyarko (2009): *Is the Brain Drain Good for Africa?* in (Jagdish Bhagwati and Gordon Hanson, eds.), *Skilled Migration today: Prospect, Problems and Policies,*,Oxford University Press, USA.

Gibson, J., and D. McKenzie (2011): "Eight Questions about Brain Drain," *The Journal of Economic Perspectives*, 25(3), 107–128.

Girsberger, E. M. (2014): "Essays on Migration, Education and Work Opportunities," Doctoral dissertation.

Grogger, J., and G. H. Hanson (2011): "Income Maximization and the Selection and Sorting of International Migrants," *Journal of Development Economics*, 95(1), 42–57.

Heckman, J. J., and E. J. Vytlacil (2007): "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in new Environments," *Handbook of econometrics*, 6, 4875–5143.

Imbens, G. W. (2015): "Matching Methods in Practice: Three Examples," *Journal of Human Resources*, 50(2), 373–419.

Kuepie, M., C. J. Nordman, and F. Roubaud (2009): "Education and Earnings in Urban West Africa," *Journal of Comparative Economics*, 37(3), 491–515.

Manski, C. (1993): "Identification of Endogenous Social Effects: the Reflection Problem," *Review of Economic Studies*, 60, 531–542.

Mbaye, L. M. (2014): ""Barcelona or Die": Understanding Illegal Migration from Senegal," *IZA Journal of Migration*, 3(1), 1–19.

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

——— (2011): "Can Migration Reduce Educational Attainment? Evidence from Mexico," *Journal of Population Economics*, 24(4), 1331–1358.

Mountford, A. (1997): "Can a Brain Drain be Good for Growth in the Source Economy?," *Journal of Development Economics*, 53(2), 287–303.

Murard, E. (2016): "The Impact of Migration on the Family Left Behind : Estimation in Presence of Intra-Household Selection of the Migrants," Unpublished manuscript.

Nyarko, Y. (2011): "The Returns to the Brain Drain and Brain Circulation in Sub-Saharan Africa: Some Computations Using Data from Ghana," NBER Working Paper 16813.

Özden, Ç., and D. Phillips (2014): "What Really Is Brain Drain? Location of Birth, Education and Migration Dynamics of African Doctors," in *Workshop on 'The Drivers and Dynamics of High-Skilled Migration'*. Oxford Martin School, University of Oxford.

Rosenzweig, M. R. (2008): "Higher Education and International Migration in Asia: Brain Circulation," in *Annual World Bank Conference on Development Economics*, pp. 59–100.

Schans, D., M. Valentina, B. Schoumaker, and M.-L. Flahaux (2013): "Changing Pattern of Ghanaian Migration," MAFE Project Working Paper 20.

Schoumaker, B., M.-L. Flahaux, and A. M. Mobhe (2013): "Changing Pattern of Congolese Migration," MAFE Project Working Paper 19.

Shrestha, S. A. (Forthcoming): "No Man Left Behind: Effects of Emigration Prospects on Educational and Labour Outcomes of Non-migrants," *The Economic Journal*.

Stark, O., C. Helmenstein, and A. Prskawetz (1997): "A Brain Gain with a Brain Drain," *Economics letters*, 55(2), 227–234.

Steinmayr, A. (2014): "When a Random Sample Is Not Random: Bounds on the Effect of Migration on Household Members Left Behind," Kiel Working Paper 1975.

Tankwanchi, A. B. S., Ç. Özden, and S. H. Vermund (2013): "Physician Emigration from Sub-Saharan Africa to the United States: Analysis of the 2011 AMA Physician Masterfile," *PLoS Med*, 10(9), e1001513.

Theoharides, C. (2014): "Three Essays on the Economics of International Migration," Doctoral dissertation.

Vidal, J.-P. (1998): "The Effect of Emigration on Human Capital Formation," *Journal of Population Economics*, 11(4), 589–600.

Willis, R. J., and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5, Part 2), 7–36.

Appendix

A Extension of the Model: Binding Budget Constraint

Imperfect credit markets are a common feature of developing economies. Therefore, it is important to account for the possibility of a binding budget constraint. Two cases are possible: (1) the budget constraint is binding for the schooling investment irrespective of the migration decision, or (2) the budget constraint is binding for the schooling investment only when the household decides to attempt migration. Equation (7) and (8) already account for the first case. In the second case, the maximization problem of the household includes an additional term, $\lambda(x,u)M^* \leq 0$, that reflects the liquidity constraint. That is, the household chooses $S(1) = h$ over $S(1) = l$, if and only if :

$$\begin{aligned}(\Pi_h^0(X) &+ u_h^0) - (\Pi_l^0(x) + u_l^0) \\ &+ p_h \times (\Pi_h^1(X) + u_h^1 - (\Pi_h^0(X) + u_h^0)) \\ &- p_l \times (\Pi_l^1(X) + u_l^1 - (\Pi_l^0(X) + u_l^0)) + \lambda(x,u) > 0\end{aligned}$$

$\lambda(x,u)$ should increase with the wealth of the household, and be zero if the budget constraint is not binding. Conversely, if no borrowing opportunity exists, $\lambda(x,u) = -\infty$, and $S(1) = l$.

In the case of a constrained maximization, $S_{cf} = S(0)$ if no one attempts migration for $p_1 = p_0 = 0$. The latter will be true under three plausible conditions: (1) given a schooling choice, individuals maximize there expected returns to migration, (2) any migration attempt is costly, and (3) migration yields a positive return only in the case of a successful emigration.

B Observed Characteristics

Tables 3, 4, and 5 compare the observed characteristics of those who attempt migration (Treated), to the characteristics of those who never attempt migration (the Non-Treated).

	Treated[a]		Non-Treated[a]		Norm.
	Mean	s.d.	Mean	s.d.	diff.[a]
Men (N=713)	(N=218)		(N=495)		
Secondary	1.00	0.07	0.98	0.16	0.22
Upper Secondary	0.98	0.14	0.82	0.41	0.76
Tertiary	0.70	0.48	0.34	0.50	1.03
Years of schooling	17.04	2.77	15.06	3.44	0.89
Father's occupation					
Higher-level	0.23	0.44	0.13	0.36	0.36
Skilled employee	0.41	0.52	0.31	0.49	0.28
unskilled employee	0.09	0.30	0.20	0.42	-0.43
Self-employed/unemployed	0.27	0.47	0.36	0.51	-0.26
Migrant Household[b]	0.26	0.47	0.21	0.43	0.16
Age at survey	42.50	9.71	38.60	10.25	0.55
Network size at age 15	0.68	2.06	0.52	1.27	0.13
Women (N=973)	(N=185)		(N=788)		
Secondary	0.99	0.13	0.94	0.26	0.35
Upper Secondary	0.93	0.26	0.61	0.52	1.11
Tertiary	0.36	0.51	0.16	0.39	0.63
Years of schooling	15.81	2.88	12.79	4.06	1.22
Father's occupation					
Higher-level	0.30	0.49	0.10	0.31	0.71
Skilled employee	0.25	0.46	0.35	0.51	-0.29
unskilled employee	0.10	0.32	0.20	0.43	-0.39
Self-employed/unemployed	0.35	0.50	0.35	0.50	0.00
Household with a migrant[b]	0.46	0.53	0.20	0.42	0.78
Age at survey	38.04	9.85	37.74	9.70	0.04
Network size at age 15	1.08	2.21	0.34	0.87	0.63

[a] Treated: respondents who report a migration attempt. Non-Treated: respondents who do not report a migration attempt. Normalized difference: difference between mean of treated and Non-Treated, divided by the standard deviation.

[b] Dummy for having a migrant a migrant in the household at age 15.

Table 3: Observed Characteristics: DR Congo

	Treated[a]		Non-Treated[a]		Norm.
	Mean	s.d.	Mean	s.d.	diff.[a]
Men (N =580)	(N=267)		(N=313)		
Secondary	0.99	0.08	0.98	0.13	0.12
Upper Secondary	0.91	0.29	0.81	0.40	0.41
Tertiary	0.67	0.48	0.32	0.48	1.03
Years of schooling	17.19	3.74	14.69	3.61	0.96
Father's occupation					
Higher-level	0.28	0.45	0.09	0.29	0.70
Skilled employee	0.32	0.47	0.33	0.48	-0.03
unskilled employee	0.11	0.32	0.16	0.38	-0.18
Self-employed/unemployed	0.29	0.46	0.42	0.50	-0.39
Household with a migrant[b]	0.32	0.47	0.11	0.32	0.73
Age at survey	41.65	9.57	37.12	9.41	0.68
Network size at age 15	1.16	2.23	0.23	0.78	0.79
Women (N=784)	(N=194)		(N=590)		
Secondary	0.99	0.13	0.89	0.35	0.52
Upper Secondary	0.78	0.46	0.60	0.55	0.51
Tertiary	0.55	0.56	0.16	0.41	1.14
Years of schooling	15.90	3.97	12.20	4.83	1.19
Father's occupation					
Higher-level	0.28	0.50	0.08	0.30	0.69
Skilled employee	0.39	0.55	0.28	0.50	0.27
unskilled employee	0.07	0.28	0.18	0.42	-0.42
Self-employed/unemployed	0.27	0.50	0.46	0.56	-0.52
Household with a migrant[b]	0.41	0.55	0.14	0.38	0.81
Age at survey	38.45	9.42	37.90	10.19	0.08
Network size at age 15	0.80	1.38	0.29	0.90	0.62

[a] Treated: respondents who report a migration attempt. Non-Treated: respondents who do not report a migration attempt. Normalized difference: difference between mean of treated and Non-Treated, divided by the standard deviation.

[b] Dummy for having a migrant a migrant in the household at age 15.

Table 4: Observed characteristics: Ghana

	Treated[a]		Non-Treated[a]		Norm.
	Mean	s.d.	Mean	s.d.	diff.[a]
Men (N=522)	(N=302)		(N=220)		
Secondary	0.62	0.46	0.54	0.48	0.25
Upper Secondary	0.39	0.47	0.30	0.44	0.28
Tertiary	0.20	0.38	0.13	0.32	0.27
Years of schooling	10.50	4.39	9.70	3.70	0.28
Father's occupation					
Higher-level	0.12	0.32	0.19	0.37	-0.25
Skilled employee	0.26	0.42	0.21	0.39	0.17
unskilled employee	0.18	0.37	0.17	0.36	0.04
Self-employed/unemployed	0.43	0.48	0.43	0.47	0.01
Household with a migrant[b]	0.33	0.45	0.21	0.39	0.41
Age at survey	38.66	8.45	36.05	7.09	0.47
Network size at age 15	0.75	1.43	0.46	1.03	0.33
Women (N=527)	(N=218)		(N=309)		
Secondary	0.68	0.41	0.42	0.43	0.88
Upper Secondary	0.33	0.41	0.16	0.33	0.62
Tertiary	0.14	0.31	0.08	0.24	0.29
Years of schooling	10.43	3.30	8.57	3.08	0.83
Father's occupation					
Higher-level	0.15	0.31	0.13	0.29	0.09
Skilled employee	0.35	0.42	0.27	0.39	0.27
unskilled employee	0.19	0.34	0.16	0.32	0.11
Self-employed/unemployed	0.31	0.41	0.44	0.44	-0.42
Household with a migrant [b]	0.37	0.43	0.19	0.35	0.66
Age at survey	40.11	7.31	36.69	7.40	0.66
Network size at age 15	0.81	1.18	0.48	1.07	0.41

[a] Treated: respondents who report a migration attempt. Non-Treated: respondents who do not report a migration attempt. Normalized difference: difference between mean of treated and Non-Treated, divided by the standard deviation.

[b] Dummy for having a migrant a migrant in the household at age 15.

Table 5: Observed characteristics: Senegal

C Instrument: index measure interacting the network and the unemployment rate

To take advantage of exogenous variations that affect the returns to education abroad but not at home, the index measure interacts, at a given age, the strength of a network in one of the three destinations and the unemployment rate in this location.

Let $\mathcal{C}$ be the set of destination countries. For example, for Senegal, $\mathcal{C}$ consists of France, Italy, Spain and other OECD countries pooled together. Let $N_{c,t,i}$ be the size of i's network in country c when he is aged t. Let $U_{c,t,i}$ be the unemployment rate by gender for individual between 25 to 34, in country $c \in \mathcal{C}$ when i is aged t. For the group "other OECD countries", the unemployment rate is the OECD unemployment rate. The index takes the following form:

$$I_i = \frac{1}{17-15} \sum_{15 \leq t \leq 17} \frac{\sum_{c \in \mathcal{C}} U_{c,t,i}(1 + \ln(N_{c,t,i}+1))}{\sum_{c \in \{\mathcal{C}\}}(|\mathcal{C}| + \ln(N_{c,t,i}+1))}$$

The data on unemployment rate are from the OECD database, starting in year 1972.

D Technical Details of estimation

Choices for estimation are detailed here. Unless stated otherwise, estimation is conducted separately by gender, father's-occupation and household-with-a-migrant status. Means of conditional expectations are computed, using provided weights. For all procedures, standard errors are computed with the canonical bootstrap, using 199 replications. The 90% confidence interval are computed using quantiles from a normal distribution.

D.1 Matching

Observed characteristics used to match individuals are network size at age 15, age at survey, household size, religion, and ethnicity (when available). The first step ensures that the common support assumption holds. The propensity score estimation is done by gender, using a logit regression of the migration attempt variable on the observed characteristics, the squared age at time of the survey, and interaction terms between the father's occupation and household-with-a-migrant status, father's occupation and kinship size. The results are robust to alternative specifications of the propensity score estimator. Observations outside the common support of treated and non-treated are dropped: DR Congo, 12 men and 4 women, Ghana, 20 men and 3 women, Senegal, 88 men and 6 women.

The second step matches treated to non-treated with replacement, using the nearest-neighbor with respect to the Mahalanobis distance. Treated units receive the number of years of schooling of their matched unit (an average in case of ties). A bias-correction adjustment is performed following Imbens (2015). Then, the bias-adjusted years of schooling of the matched non-treated are assigned to the treated. The dichotomous variables for completion of some secondary, upper secondary and tertiary education are constructed according to the number of years of schooling corresponding to each level.

D.2 Local Instrumental Variable

The construction of the instrument is detailed in Appendix C. In the first step, a linear model controlling for cohort indicators (25-34, 35-44, 45 or more) and the instrument predicts the probability to attempt migration by gender, father's-occupation and household-with-a-migrant status. To mitigate the weak instrument problem, the instrument is dichotomized into two categories: above or below median.

In a second step, the outcome of interest (completion of some secondary, upper secondary, tertiary education, or years of schooling) is regressed on the predicted attempts probabilities and the cohort indicators, without a constant term. Ideally, the predicted attempts probabilities should enter through a flexible functional form (example, polynomial function). Here, the predicted probability is dichotomized, into two categories: below or above median. The regression is :

$$S = X\beta_0 + K(\hat{p}(Z))$$

where X is the vector of cohort dummies, $\hat{p}(Z)$ the estimation propensity score, Z the instrument, and K is a step function with different values above and below the median. The predicted average level of schooling in the counterfactual, $\mathbb{E}S(0)$ in each subgroup is then:

$$\hat{\mathbb{E}}S(0) = X\hat{\beta}_0$$

E Additional Empricial Analysis

E.1 Households with a Migrant

Households with a migrant are defined as households where at least one regular member of the household lived in Europe in the calendar year where the potential migrant turned 15 years old. A migrant can influence schooling investments in several ways. For example, remittances relax the household's budget constraint and, thus, allow larger investments in schooling. This is the *income effect* of migration (Dinkelman and Mariotti, 2015; Theoharides, 2014). Besides, migrant networks decrease migration costs and lead to looser selection of migrants (McKenzie and Rapoport, 2010). Finally, migrants influence the perceived probability of success in migration, and hence, the decision to attempt migration. The last panels of Tables 3, 4, and 5 in Appendix B show that those who attempt migration are clearly more likely to belong to a household with a migrant.

With respect to the selection effect, there is not much change in DR Congo, compared to the previous results. In Ghana, households with a migrant tend to experience stronger (negative) selection effect than their counterpart without a migrant, in particular within the richest households. In Senegal, the selection effect appears stronger for female candidates when they have a migrant in the household; for example, at the upper secondary level, considering all father's-occupation subgroups, the selection effect is -12.2 pp for households with a migrant compared to -0.1 pp for households without migrant.

Because the subgroups have sometimes small sizes, discussions of incentive and net effects focus mainly on estimates for all father's-occupation subgroups. With respect to the incentive effect, in DR Congo, except for their magnitude, the estimates do not vary much from the previous

section.

In Ghana, both worst-case and restricted bounds suggest zero or positive incentive effects at the tertiary level, and for the years of schooling, for men and for women. These effects are comparable to previous estimated magnitude. Matching concludes strong positive incentive effects for men (11.9 pp) and modest positive effects for women (2.7 pp).

In Senegal, worst-case bounds on incentives resemble pretty much the previous bounds. Matching essentially concludes to some negative incentives at the upper secondary level (-2.1 pp for men, -3.6 pp for women). At the tertiary level, matching estimates suggest some positive incentive for men (4 pp), and zero effect for women. The constrained bounds mainly suggest zero or negative effects for men: at the upper secondary level, a lower bound of -13.1 pp. For women, the corresponding effect ranges between -5.7 p.p and 13.0 pp

Estimated net effects remain mainly unchanged, except in two cases: matching concludes to larger positive net effects for men in Ghana, while both procedures hint to more negative net effects for women in Senegal.

From this analysis, we can conclude that, on one hand, households with a migrant tend to experience stronger selection effects; this tends to worsen the net effect, in particular in Senegal. On the other hand, in Ghana, incentives are somewhat higher for households with a migrant, so that they mitigate the selection effect.

E.2 Return Migration

This section addresses the contribution of returned migrants to the average human capital in the sending countries. Return migration should help mitigating the selection effect, if high-skilled migrants return more often than than their low-skilled counterparts. After observing that students are over-represented among returned migrants in Senegal, Baizán, Beauchemin, and González-Ferrer (2013) conclude that the "brain drain appears to be a limited issue in the context of Senegalese migration".

The change in the net effect when accounting for returners, $\Delta^r_{net} - \Delta_{net}$, is identified from the data as the difference:

$$E(S|\{\mathcal{R}\}, X = x) - E(S|M = 0, X = x) \tag{16}$$

where $\{\mathcal{R}\}$ is the pool of "never-migrants" and "returnees".

Table 6 presents this difference (16) by country of origin and gender, and distinguishes the richest households (father with higher-level occupation) from the rest of the households.

In DR Congo, there is hardly any change in the net effect when one also accounts for the human capital brought back by returned migrants. In Ghana, there is a slight increase in the net effect for men at the tertiary level (1.1 pp) and for the years of schooling (0.16 year). The effect is mainly due to rich households. In Senegal, there is hardly any significant change observed; if any, it is in the direction of a reduction in the average human capital in Senegal. Note again the difference between the richest households and the remaining households.

Overall, these findings qualify the hypothesis that return migration might be a stronger channel for "Brain Gain" than the prospect of emigration.

		All occupations except higher-level		Higher-level Occupations		All Occupations	
		$\Delta^r_{net} - \Delta_{net}$	p^{a}	$\Delta^r_{net} - \Delta_{net}$	p^{a}	$\Delta^r_{net} - \Delta_{net}$	p^{a}
Men							
DR Congo	Upper Sec.	0.000	*0.03*	0.002	*0.41*	0.001	*0.03*
	Tertiary	0.000	*0.64*	0.003	*0.45*	0.003	*0.09*
	Years	0.012	*0.15*	0.109	*0.19*	0.049	*0.06*
Ghana	Upper Sec.	0.003	*0.03*	0.000	*0.21*	0.003	*0.04*
	Tertiary	0.010	*0.07*	0.029	*0.18*	0.011	*0.01*
	Years	0.152	*0.01*	0.158	*0.26*	0.158	*0.00*
Senegal	Upper Sec.	0.011	*0.41*	0.038	*0.12*	0.018	*0.08*
	Tertiary	-0.012	*0.14*	0.042	*0.15*	0.001	*0.92*
	Years	-0.011	*0.87*	0.370	*0.10*	0.081	*0.33*
Women							
DR Congo	Upper Sec.	0.004	*0.36*	0.000	*0.48*	0.001	*0.24*
	Tertiary	0.000	*0.30*	0.006	*0.59*	0.002	*0.33*
	Years	0.001	*0.29*	0.122	*0.43*	0.024	*0.28*
Ghana	Upper Sec.	0.003	*0.06*	0.000	*0.03*	0.003	*0.09*
	Tertiary	0.008	*0.22*	0.002	*0.38*	0.007	*0.28*
	Years	0.055	*0.20*	0.029	*0.33*	0.051	*0.19*
Senegal	Upper Sec.	-0.016	*0.24*	0.012	*0.39*	-0.010	*0.33*
	Tertiary	-0.011	*0.15*	0.018	*0.30*	-0.004	*0.54*
	Years	-0.073	*0.24*	0.142	*0.37*	-0.028	*0.60*

[a] p: Bootstrapped p-value of the difference $\Delta^r_{net} - \Delta_{net}$.

Table 6: Difference in net effect measures when accounting for return migration

E.3 "Absolute" Brain Drain

As defined so far, selection, incentive and net effects are relative measures (i.e. relative to the proportion of residents). However, much of the public discussion focuses on *absolute measures* of the brain drain, that is, the number of high-skilled that are "lost", rather than the proportion of the resident population. The results for these absolute measures are presented below.

Absolute measures of selection, incentive and net effects are easy to compute, given our previous characterization of the counterfactual. Denote : $\Delta_{sel}^{abs}, \Delta_{inc}^{abs}$, and Δ_{net}^{abs} and note that: $\Delta_{inc}^{abs} = \Delta_{inc}$, since in the counterfactual scenario, all the population stays in the origin country. We have:

$$\Delta_{sel}^{abs}(X) = \mathbb{E}(S; M = 0|X) - \mathbb{E}(S|X) \quad (17)$$

$$\Delta_{sel}^{abs}(X) = \mathbb{E}(S; M = 0|X) - \mathbb{E}(S(0)|X) \quad (18)$$

The results presented here are for households without a migrant, using the restricted bounds to estimate incentive and net effects.

In DR Congo, there is a net decrease in the number of residents with at least some upper secondary education compared to the closed economy: about 2.5 pp for men, and about 1.5 pp for women. That is, in a case where 1000 men would have obtained upper secondary education in the *closed economy*, 975 ($= 1000 \times (1 - 0.025)$) would stay in DR Congo after incentives and selection have operated. This decrease is mainly driven by the richest households where the decrease is about 16 pp for men and 9 pp for women (Figure 12). The effect is very similar at the tertiary level (Figure 13).

In Ghana, at the upper secondary level, absolute selection measures (Equation 17) are strictly negative for almost all households, and even more negative for the richest households (Figure 12). This is a sign that migration is mostly prevalent among the richest households. It translates into negative absolute net effects: for men, as low as -18 pp for the richest households, and between -7 pp and -3 pp for the other households. In other words, out of 1000 men with upper secondary education in the *closed economy*, there is about 930 stayers. For women, -5 pp for the richest households, and between -5 pp and -1 pp for the other households. The picture is very similar at the tertiary level (Figure 13).

Finally, in Senegal, the absolute selection effect is very similar across all households at the upper secondary level. For men, it ranges between -20 pp and -15 pp. For all women taken together, it amounts to a 6 pp decrease (Figure 12). Hence, the absolute net effect is relatively large: between -26 pp and -14 pp for men, between -9 pp and 3 pp for women. Out of 1000 men with upper secondary education in the *closed economy*, there is about between 740 and 860 stayers. At the tertiary level, the absolute effects are less pronounced -8 pp for all men and -3 pp for all women, but still leads to negative net effects (Figure 13).

Hence, compared to the *closed economy* scenario, all countries experience a decrease in the *number* of their high skilled population. The incentives to invest more in schooling do not fully compensate for these absolute losses. In DR Congo and Ghana, the effects are mainly concentrated on the richest households, while in Senegal, the effects are spread across all households. The findings are very similar, and sometimes even more pronounced when one focuses on households with a migrant.

E.4 Robustness Analyses

This section assesses the robustness of the main findings under alternative specifications: including individuals with primary education, using a stricter definition of a migration attempt, restricting the sample to the younger cohort, using father's education instead of father's occupation, and, finally, using an instrument to refine the worst-case bounds.

E.4.1 Including Individuals with Primary Education

An assumption of the baseline estimation is that migration decisions are not made very early in life, and so do not affect early schooling decisions. However, estimates of the selection effect at the secondary education level (estimated as was done at the upper secondary and tertiary level previously) give negative values for some subgroups in Senegal. Moreover, matching estimates of the incentive effects at this level also give negative values in some subgroups in Senegal.

Including individuals with only primary education makes no difference in the estimates for DR Congo and Ghana, since this is only a small proportion of the population. However, it changes the estimates for Senegal. More precisely, at the upper secondary level, the lower bounds of the worst-case bounds and the restricted bounds are lower in all households groups except the richest, yet matching estimates still suggest some negative incentives for men at the upper secondary level. Estimates at the tertiary level are mainly unchanged.

Including individuals with primary education allows analyzing the effects of migration prospects on secondary education investments in Senegal. First, consider men. Figure 14 (left panels) shows the selection, incentive and net effects as estimated using the different available methods. There is a negative selection effect in all households, leading to an overall decrease of 3.8 pp.

As for the incentive effect, the unrestricted bounds are wide and centered around zero, except for the richest households. Matching estimates suggest negative incentives of order -6.5 pp. The restricted bounds suggest either small positive effects or strong negative effects, except for the richest where the effect is close to zero.

As for the net effect, both matching and restricted bounds estimates conclude negative net effects. The matching estimate concludes to a 10.2 pp decrease in secondary schooling, compared to the *closed economy* (15% change from the average). The bounds estimates suggest that the net effect lie between -14.7 p.p and -1.1 pp. Note that there seems to be no substantial effect on the richest households.

For women, the results are very similar. Matching estimates conclude to a 6.8 pp decrease in secondary schooling, compared to the *closed economy* (17% change from the average). The bounds estimates suggest that the net effect lies between -6.8 p.p and -0.5 pp. Note again, that there seems to be no substantial effect on the richest households.

Thus, it seems plausible that migration prospects affect early schooling investments in Senegal.

E.4.2 A Stricter Definition of Migration Attempt

So far, a migration attempt has been defined as any self-reported attempt at migration. In this Section, *serious* attempts of migration are distinguished from *non-serious* ones. An attempt is labeled as *non-serious* in three cases: first, the respondent has undertaken none of the following steps: applying for papers, saving money or applying for enrollment at a European university.

Second, the respondent reports being unsuccessful because he changed his mind, changed target destination (for a destination outside Europe) or because he found it easier to stay. Third, the reason why the individual was unsuccessful was either missing or reported as "other". This classifies as "having not attempted migration" 23 attempts in DR Congo (5.7% of attempts in the sample), 34 in Ghana (7.4%), 41 in Senegal (7.9%).

This change in the definition of migration attempts has no effect on the estimates for DR Congo and Ghana. However, the estimates of incentive and net effects for men in Senegal become less negative. This is driven by individuals with skilled employed parents. For them, matching and bound estimates of the incentive effect tend to be less negative. As a result, matching concludes to zero or small positive incentive effects for all men in Senegal, except at the secondary level, where the effect is still negative (-5 pp). As for the bounds estimates, they suggest zero or negative incentive effects for all households, except the richest. Pooling all households together except the richest, the lower bound is -7 pp at the upper secondary level (compared to -14 pp with the previous definition), -14 p.p at the tertiary level (compared to -28 pp), and -0.57 year of schooling (compared to -1.1 years).

Thus, the results of estimation are sensitive to the definition of migration attempts for men in Senegal. A stricter definition of a migration attempt yields less pessimistic results.

E.4.3 Additional Robustness Analyses

In addition to the previous analyses, father's occupation when the respondent is age 15 is replaced with father's education at the same time period; the results remain fairly consistent. The estimation is also restricted to the young cohort (below age 45) to avoid the confounding effect of the expansion of the education system in the early post-colonial years. Again, the results remain consistent.

Finally, since the instrument used in the LIV procedure shows some predictive power on migration behavior in Senegal, it is applied to tighten the worst-case bounds for this country (Manski, 1993).[26] Figure 15 show the results for this procedure. For women, the results imply negative incentive and net effects for the poor households, consistent with the previous discussion. For men, the bounds are similar across group and, typically, do not exclude zero, though they imply similar magnitudes as those discussed previously.

[26]The bounds with an instrument are intersection bounds (Chernozhukov, Lee, and Rosen, 2013). Estimation is implemented using the Stata Package *clrbound* provided by (Chernozhukov, Kim, Lee, and Rosen, 2015).