Heuer, Nina

Working Paper

The effect of occupation-specific brain drain on human capital

University of Tübingen working papers in economics and finance, No. 7

Provided in cooperation with:
Eberhard Karls Universität Tübingen

The Effect of Occupation-Specific Brain Drain
on Human Capital

by

Nina Heuer

The Effect of Occupation-Specific Brain Drain on Human Capital[*]

Nina Heuer[**]

April 2011

Abstract

This paper tests the hypothesis of a beneficial brain drain using occupation-specific data on migration from developing countries to OECD countries around 2000. Distinguishing between several types of human capital allows to assess whether the impact of high-skilled south-north migration on human capital in the sending economies differed across occupational groups requiring tertiary education. We find a robust negative effect of the incidence of high-skilled emigration on the level of human capital in the sending countries, thereby rejecting the hypothesis of a beneficial brain drain. The negative effect was significantly stronger for *professionals* – the occupational category with the largest incidence of south-north migration and the highest educational requirements – than for *technicians and associate professionals*.

[**] Eberhard Karls University Tübingen, Faculty of Economics and Social Sciences, Chair in International Economics, Mohlstraße 36, 72074 Tübingen, Germany. Telephone: +49 70 71/ 29 721 21, Fax: +49 70 71/ 29 55 74, E-mail: nina.heuer@uni-tuebingen.de.

1   Introduction

Emigration of high-skilled individuals from developing countries to developed countries –
or "brain drain" (Docquier and Rapoport, 2008) – leaves the migrant-sending economies
initially with a reduced supply of skilled labor in production, research, public services,
and political institutions. However, the migration literature has identified several positive
feedback effects of the brain drain on the source countries. These include remittances,
network effects, and return migration of individuals with enhanced skills. Stark et al.
(1997, 1998) and Mountford (1997) were the first to argue that there might even be a
(net) "brain gain", i.e. an increase in the human capital stock of the sending economies,
from the emigration of high-skilled workers. The reason is that the prospect of emigration
to countries with higher wages, through increasing expected returns to education, might
incentivize people in developing countries to invest more in education. If the brain gain
exceeds the brain drain, this is called a "beneficial brain drain" (cf. Beine et al. 2001,
2008).

   This paper contributes to the empirical literature on the impact of the brain drain
on human capital in the sending economies in various respects. Most importantly, we
improve upon existing studies in that we reassess the brain gain argument distinguishing
between several types of human capital and brain drain. This is possible due to the use of
recently constructed datasets providing occupation-specific south-north migration rates at
two distinct levels of disaggregation for, respectively, a large and a small set of developing
countries around 2000 (Heuer 2010). We are thus in a position to test the hypothesis of a
beneficial brain drain for different high-skilled occupational categories. Furthermore, we
can exploit the cluster-sample structure of the data to extract observed and unobserved
heterogeneity at the country level via the use of panel data estimation techniques. Using
data at the major level of the International Standard Classification of Occupations 1988
(ISCO-88), we estimate the impact of occupation-specific brain drain on the corresponding
occupation-specific *ex post* employment share in the migrant-sending countries in order
to assess the net effect of the brain drain on human capital in the source economies.
The empirical work builds on the theoretical model by Mountford (1997) and includes the
brain drain variable measured in the same period as the dependent human capital variable
in order to model anticipatory expectation-building. In addition, we control for possible
convergence effects by inserting the lag of the dependent variable as a regressor.

   Employing occupation-specific brain drain rates and employment shares at the ISCO-
88 sub-major level, we estimate a similar specification in order to allow for further het-
erogeneity distinguishing between eight high-skilled occupational categories. However,
since these more disaggregated data are only available for a few developing countries of
emigration, this analysis is rather considered as a robustness check.

   In order to address endogeneity concerns regarding the occupation-specific emigration

rates, we use bilateral information on immigrants in OECD countries as well as geographic variables to predict bilateral migrant stocks. These can be aggregated to an instrument for unilateral emigration. This procedure is in analogy to Frankel and Romer (1999), who construct the geographical component of trade as an instrument for trade, and to Felbermayr et al. (2010), who apply this method in the context of migration.

The remainder of this paper is organized as follows: Section 2 reviews the existing empirical literature. Section 3 presents stylized facts on the occupation-specific brain drain. Section 4 revisits the hypothesis of a beneficial brain drain in a model with probabilistic migration and heterogeneous agents. Section 5 presents the empirical specifications and estimation results. Section 6 summarizes the main results and concludes.

2 Review of the Empirical Literature

Recently created macro-level datasets of south-north migration rates based on information on immigrants in OECD countries by country of origin and – partly imputed – educational attainment (Carrington and Detragiache 1998; Adams 2003; Docquier and Marfouk 2006; Beine et al. 2007; Docquier et al. 2008; Defoort 2008) have rendered possible to quantify the effects of high-skilled emigration on human capital for many developing countries, and thus to empirically test the hypotheses of brain gain and beneficial brain drain at the aggregate level. Whereas the relevant empirical studies commonly use the proportion of tertiary educated natives (residents plus migrants) from developing countries living in OECD countries to account for the incentive effect, they differ with respect to the empirical counterparts for human capital investments and the human capital stock in the sending economies. The results from these estimation analyses – mainly stemming from cross-sectional data – are mixed: Studies proxying investments in human capital by the growth rate of the proportion of tertiary educated natives find that the brain drain rate (the proportion of migrants among the tertiary educated native population) measured in the base period exerts a positive effect on the rate of change of the *ex ante* stock of human capital in cross-sectional analyses (Beine et al., 2003, 2008; Docquier et al., 2008), and in a panel data analysis if countries are poor (Beine et al., 2011). This positive impact is interpreted in favor of the suggested brain gain. Yet, there is also evidence for a "disincentive effect", since the same brain drain measure has a negative impact on investments in human capital accounted for by tertiary school enrollment rates measured in the same period as the brain drain in cross-sectional analyses (Groizard and Llull, 2006, 2007a), and in a panel data analysis (Checchi et al., 2007).[1] Although appearing contradictory, these findings could derive from a time lag with which individuals acquire tertiary education for

---

[1]   Estimating a random-effects model of a similar specification, Faini (2004) does not find any significant effect of either the tertiary or the secondary emigration rate on tertiary and secondary school enrollment, respectively.

given observed emigration. A high emigration rate in some base period would thus coincide with a decreased tertiary school enrollment rate (many emigrants from poor countries leave their home countries to study in the OECD) – reflecting the brain drain effect, but also with a potential increase in the proportion of the tertiary educated native population over time – reflecting a dynamic brain gain effect. By contrast, the findings cannot be reconciled under the assumption of anticipatory expectation-building, which is a standard assumption in the relevant theoretical literature: Under this assumption, the negative relation between the brain drain and tertiary school enrollment would reflect a "disincentive effect", whereas the analysis of the impact of the brain drain on the growth rate of the proportion of tertiary educated natives would not have any sensible interpretation.

Di Maria and Lazarova (2009) find evidence that the possibility of high-skilled emigration decreases the contemporaneous enrollment in science and technology specialties compared to a situation in which emigration is inhibited for countries with a low level of development, whereas the opposite is the case for relatively more developed countries.

Ha et al. (2009) study the effects of permanent and temporary emigration on contemporaneous human capital formation and economic growth in Chinese provinces between 1980 and 2005. They find that permanent emigration improves the enrollment in both middle and high schools, whereas temporary emigration has a significantly positive effect only on middle school enrollment. Their results furthermore suggest a positive relation between the educational level of migrants and school enrollments.

Relying on micro-level data from a household survey conducted in Cap Verde, the results by Batista et al. (2011) suggest an important impact of the brain gain channel on educational attainment: Estimation and counterfactual analyses reveal that the probability of completing intermediate secondary schooling is increasing in the probability of own future migration.

The hypothesis of a beneficial brain drain (or net brain gain) is assessed empirically in three types of macro-level studies for several developing countries: Beine et al. (2008) and Docquier et al. (2008) conduct counterfactual experiments in which they compare observed proportions/numbers of skilled residents to hypothetical ones which they calculate using the predictions of their human capital estimations and the emigration rates of low-skilled workers. They find that a beneficial brain drain is most likely if the probability of emigration is not too high and if the level of human capital was previously low. Based on parameter estimates obtained from the regression of the growth rate of the *ex ante* stock of human capital on the high-skilled emigration rate, Beine et al. (2011) simulate the impact of high-skilled emigration on the steady state level of *ex post* human capital. From this numerical exercise they specify a concrete threshold range for the brain drain rate (20% to 30%) below which countries experience a beneficial brain drain. By contrast,

Groizard and Llull (2006, 2007b) use an estimation approach to test for a beneficial brain drain and find evidence for a negative impact of the brain drain rate on the *ex post* level of human capital, proxied by the proportion of the population with more than 13 years of school (excluding emigrants).[2] Whereas in the former paper, Groizard and Llull measure human capital in the same period as the brain drain, it is measured with a five-year lag in the latter paper. Following the reasoning presented above, the former approach is thus based on anticipatory expectation-building, whereas the latter approach as well as the counterfactual experiments rely on the assumption that individuals base their education decision on observed emigration (retrospective expectation-building).

Concerning the effects of occupation-specific brain drain on human capital in the sending countries, Clemens (2007) and Bhargava et al. (2011) are the first to empirically assess the effect of physician emigration on the supply of physicians in the sending economies. Bhargava et al. (2011) estimate a similar dynamic model as Beine et al. (2011) using a panel dataset on the number of physicians in the sending countries as well as on physician immigrants in 18 OECD receiving countries. Whereas their estimation results indeed point to the existence of a 'physician brain gain' effect, inferences on the number of physicians remaining in the sending countries suggest that the latter effect is too small to generate a 'beneficial physician brain drain', implying thus a net brain drain. Clemens (2007) studies the latter hypothesis with a different dataset for African sending countries around 2000 in a cross-sectional estimation analysis. His results are less pessimistic than those by Bhargava et al. (2011), because they do not reveal a significant impact of per capita physician emigration on the per capita number of physicians in the sending countries.

Evaluating survey data of overseas doctors in the UK, Kangasniemi et al. (2007) only find weak support for the hypothesis of a 'physician brain gain' effect.

Further evidence on occupation-specific brain drain comprises several case studies analyzing one or a few specific occupations or sectors in one or at most a few countries of emigration or immigration (e.g. Watanabe 1969; Meyer et al. 2000; Bhorat et al. 2002; Thomas-Hope 2002; Alburo and Abella 2002; Pellegrino 2002; Commander et al. 2004). The most comprehensive data are generally available for the medical sector (cf. Hagopian et al. 2004; Bhargava and Docquier 2008; Clemens and Pettersson 2008; OECD 2008).

---

[2]   In addition to using the *ex post* level of human capital, the empirical model in Groizard and Llull (2006, 2007b) differs from the one in Beine et al. (2003, 2008, 2011), and Docquier et al. (2008) in that they do not include human capital in the baseline period as a regressor.

3 Descriptive Evidence on the Occupation-Specific Brain Drain

This section uses south-north migration rates for 'high-skilled' occupational categories[3] from Heuer (2010) to highlight the large extent of heterogeneity inherent in the phenomenon brain drain. The migration rates give the number of workers from a specific developing country employed in a certain occupation in one of the OECD countries around the year 2000 over the total number of workers native of that developing country in the considered occupation. In what follows we argue that the strength of the incentive effect that is potentially triggered by the brain drain is likely to differ across high-skilled professions, because they exhibit different incidences of emigration.

Aggregating information at the ISCO-88 major level over the two high-skilled major categories *professionals* and *technicians and associate professionals*, we find that the mean brain drain rate amounts to 12.6% for the available developing countries of emigration[4] around 2000. This percentage is by construction lower than the one obtained from conventional emigration rates, which rest upon educational attainment. The reason is that some employees that worked in high-skilled occupational categories in their origin countries are employed in occupations with lower education requirements in the OECD or even unemployed due to the imperfect international transferability of (formal) skills. Looking at employment rather than population data, however, allows for a disaggregation of the brain drain: At the ISCO-88 major level, the average incidence of high-skilled south-north migration is 15.1% for *professionals*, the most highly educated employees. It exceeds the percentage of emigrated *technicians and associate professionals* (11.4%).[5] Figure 1 shows that this trend applies to all world regions, albeit on a different level.

Figure 2 illustrates variation in the brain drain at the further disaggregated ISCO-88 sub-major level. Abstracting from emigrated professionals experiencing occupational downgrading in the OECD, *life science and health professionals* as well as *physical, mathematical and engineering science professionals* seem to be the most mobile professionals – both when focusing on the 16 developing countries for which data are available and when concentrating on Eastern European and Central Asian countries.

The differences in these probabilities across different high-skilled occupations have two

---

[3] Occupational categories are defined as high-skilled if they generally require tertiary education. This applies to *professionals* (ISCO-88 major 2), who are associated with ISCED-76 levels 6 and 7, and to *technicians and associate professionals* (major 3), who mostly require education at ISCED-76 level 5 (ILO, 1990, 3-4).

[4] At the ISCO-88 major level, emigration rates for 74 developing countries have been considered. Compared to the data summarized in Heuer (2010), emigration rates including either data originally classified at ISCO-1968 or ambiguous ISCO-88 codings have been disregarded.

[5] Note that the descriptive analyses in Heuer (2010, 10-12) suggest that this difference in magnitude cannot be simply attributed to a better transferability of skills of *professionals* compared to *technicians and associate professionals*. Taking into account the similar structure of occupations included in majors 2 and 3, these analyses rather point to significant differences in the transferability of skills across different sub-major categories that are similarly represented within either of these majors.

major determinants: On the one hand, the probability of a perfect job match is higher for those high-skilled individuals with occupations requiring skills that are more easily transferable across borders (such as engineers) compared to those with rather country-specific skills (such as lawyers), *ceteris paribus*. On the other hand, this difference is partly reinforced by the migration legislations of many OECD countries, which try to attract specific types of immigrant professionals (such as doctors, engineers and other scientists) by easing their work and resident conditions.

Yet does the high probability of emigration of doctors suggest that developing countries are likely to end up with relatively more or rather less doctors compared to other professionals? On the one hand, it seems plausible that this observed large probability of emigration of doctors provokes an incentive to study medicine that is quite high relative to other subjects. On the other hand, however, the large emigration in turn curbs the supposedly large brain gain. Thus, the answer to this question is unclear *a priori*. It depends on whether a higher brain drain, reflecting a larger probability of emigration, is indeed accompanied by a higher brain gain, and on whether the latter effect outweighs the former.

Figure 1: Average south-north migration rates for 74 developing countries around 2000, by ISCO-88 majors requiring tertiary education and regions (%)
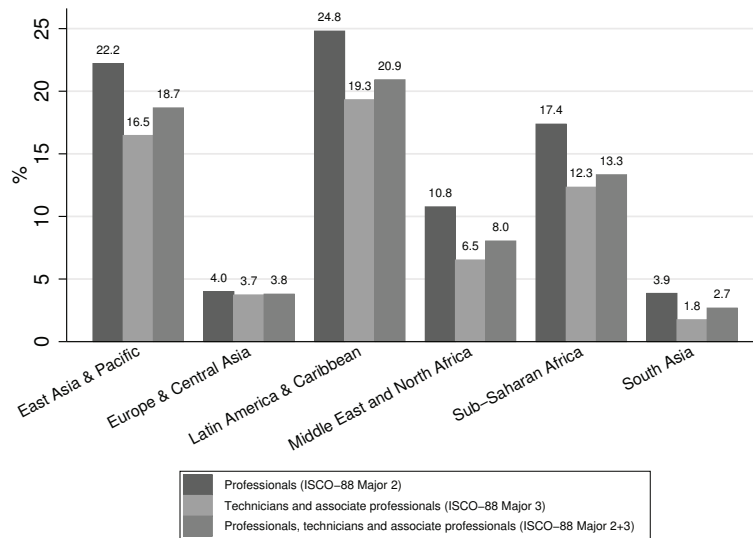*Source: Data by Heuer (2010)*

6

Figure 2: Average south-north migration rates for 16 (9) developing countries around
2000, by ISCO-88 sub-majors requiring tertiary education (%)
Source: Data by Heuer (2010)
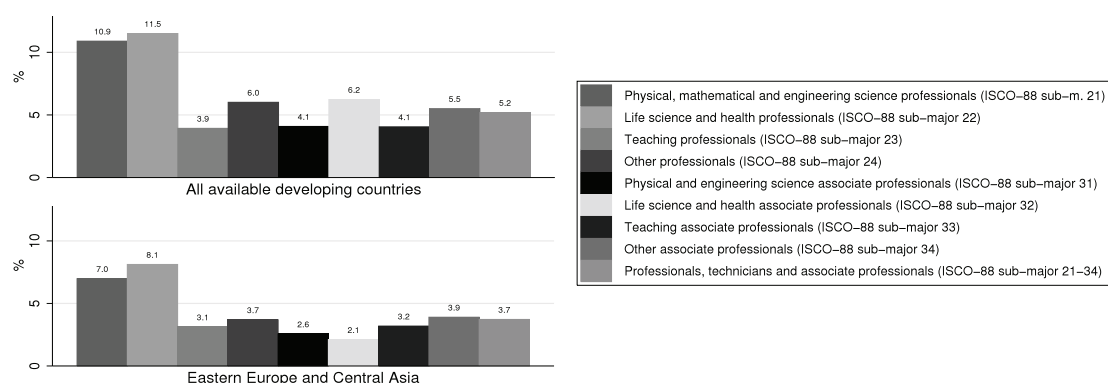
4 The Hypotheses of Brain Gain and Beneficial Brain Drain Revisited

Since the late 1990s, the brain drain literature argues that the emigration of the most
highly educated individuals from developing countries to developed countries might moti-
vate a positive effect on the formation of human capital in the migrant-sending countries
(e.g. Stark et al. 1997, 1998; Mountford 1997). The models commonly study the brain
drain in a context of high inter-country wage differences, probabilistic migration, and per-
fect transferability of skills across countries. Most of the models consider some type of
positive externality to human capital. The main argument is that the prospect of em-
igration, through increasing expected returns to education, might incentivize people in
developing countries to invest more in education (brain gain).[6]

In the following, we revisit the theoretical model by Mountford (1997) and derive the
hypothesis of a beneficial brain drain (BBD) along with the conditions under which it
arises. The model has been chosen for two reasons: First, the economy-wide aggregate
level of human capital is modeled as the share of educated individuals. This allows us to
use aggregate employment data by sending country and occupation to construct adequate
empirical counterparts of occupation-specific human capital in the empirical analysis. Sec-
ond, the dynamic structure of the growth externality, which is a function of human capital,
implies that human capital in some period is a positive function of human capital in the
previous period. The relevance of this modeling can be tested empirically, as done in
convergence-like models in the panel context by Beine et al. (2011), and in the cross-
sectional context e.g. by Beine et al. (2008).

---

[6] In models with homogeneous individuals, this effect takes the form of an increase in an individual's
investment in education due to the migration perspective. In models with heterogeneous individuals,
the brain gain is modeled as an increase in the share of individuals who choose to become educated.

Consider a small open economy in a world with one consumption good, free capital mobility, and limited mobility of labor. Production requires input factors capital (K) and labor (L) measured in efficiency units, and is characterized by constant returns to scale:

$$Y_t = F(K_t, \lambda_t L_t) = f(k_t)\lambda_t L_t \quad \text{with} \quad k_t = \frac{K_t}{\lambda_t L_t} \tag{1}$$

λ_t denotes the productivity of labor or, alternatively, the state of technology in period t. $f(k_t)$ is positive, concave in k_t, and satisfies the Inada conditions[7]. With factors being paid their marginal product, the wage rate per efficiency unit of labor is given by $w_t = \lambda_t[f(k) - kf'(k)] \equiv \lambda_t w(k)$. In a steady state equilibrium, the world interest rate r^* is constant. It follows that $r_t = r^*$, and $k_t = k \ \forall t$. The labor force is recruited from overlapping generations, whereby the continuum of heterogeneous agents in each generation is normalized to 1. The model abstracts from population growth. There are two types of agents, the educated and the uneducated, implying that the education decision is a simple discrete choice. An individual i differs from other individuals only with respect to her level of latent ability e^i, which is independent of her parents' abilities and distributed over the interval $[0, E]$ according to the (positive) density function $g(e^i)$, whereby $\int_0^E g(e^i)de^i = 1$. All agents live for three periods. In their first period, they can acquire education at a constant cost of c units of output by borrowing on the world capital markets. In the second period agents work, repay their possible debt from the first period, and save for consumption during their retirement in the third period. Individuals who invest in education are rewarded with an amount of efficiency units of labor equal to their level of latent ability e^i when working in the second period, while uneducated workers have only one efficiency unit of labor. From the condition that individual i will only invest in education if this increases her level of consumption in the third period,

$$\lambda_t w(k)e^i - c(1 + r^*) > \lambda_t w(k), \tag{2}$$

one can determine the threshold latent ability e^{*NM},[8] that separates individuals that acquire education from those who do not in the absence of migration possibilities:

$$e^{*NM} = \frac{\lambda_t w(k) + c(1 + r^*)}{\lambda_t w(k)} \tag{3}$$

The economy-wide amount of human capital is given by the proportion of educated workers:

$$s_t^{NM} = \int_{e_t^{*NM}}^{E} g(e^i)de^i \tag{4}$$

---

[7] $\lim_{k \to 0} f(k) = 0, \ \lim_{k \to 0} f'(k) = \infty, \ \lim_{k \to \infty} f'(k) = 0.$

[8]   In Mountford (1997), it is assumed that $e^{*NM} \in [0 + \epsilon, E - \epsilon]$, where $0 < \epsilon < E/2$.

This proportion is decisive for growth through an intergenerational externality which relates productivity in one period to the level of human capital in the previous period: $\lambda_t = \lambda(s_{t-1})$, with $\lambda'_t > 0$. The implication of this assumption for the dynamics of the human capital stock can be derived as follows:

$$\frac{ds_t^{NM}}{ds_{t-1}^{NM}} = \frac{ds_t^{NM}}{de_t^{*NM}} \cdot \frac{de_t^{*NM}}{d\lambda_t} \cdot \frac{d\lambda_t}{ds_{t-1}^{NM}} = [-g(e_t^*)] \cdot \left[-\frac{c(1-r^*)}{\lambda^2(s_{t-1}^{NM})w(k)} \right] \cdot \lambda'(s_{t-1}^{NM}) > 0 \qquad (5)$$

Thus, in the benchmark case without the possibility of emigration, human capital in t is a positive function of human capital in t-1.[9]

Mountford (1997) models the case of a brain drain assuming that only educated agents successfully emigrate with probability π, motivated by a higher wage per efficiency unit of labor in the world economy, denoted w^F, compared to the wage in the home economy: $w^F > \lambda_t w^H$. This emigration probability is meant to reflect immigration quotas imposed by the receiving countries. It transforms the agent's decision problem into an expected utility problem. Anticipating the opportunity to migrate in their second period of life, individuals will opt for education if:

$$[\pi w^F + (1 - \pi)\lambda_t w^H]e^i - c(1 + r^*) > \lambda_t w^H \qquad (6)$$

Since agents are assumed to be risk-neutral, they do not attribute any discount factor to the uncertain option of finding a job in a foreign country. The threshold latent ability in the presence of migration possibilities for the educated is then given by:

$$e^* = \frac{\lambda_t w^H + c(1 + r^*)}{\pi w^F + (1 - \pi)\lambda_t w^H} < e^{*NM} \qquad (7)$$

Thus, in the presence of a positive probability of emigration for the educated, more individuals will opt for education. The economy's *ex post* level of human capital is:

$$s_t = \frac{(1 - \pi) \int_{e_t^*}^{E} g(e^i)de^i}{1 - \pi \left(\int_{e_t^*}^{E} g(e^i)de^i \right)} \qquad (8)$$

s_t is decreasing in education costs c and in the domestic wage rate w^H, yet increasing in the foreign wage rate w^F. Contrasting the baseline situation without the possibility of emigration, the dynamics of the human capital stock in the presence of the brain drain

---

[9]   Depending on the functional form of $\lambda(s_{t-1})$, there will exist either a single or multiple steady state equilibria for the economy's level of human capital.

are less clear. This can be easily seen from the following derivative:

$$\frac{ds_t}{ds_{t-1}} = \frac{\partial s_t}{\partial e_t^*} \cdot \frac{\partial e_t^*}{\partial \lambda_t} \cdot \frac{\partial \lambda_t}{\partial s_{t-1}} = \frac{-(1-\pi)g(e^i)}{\left[1 - \pi \int_{e_t^*}^{E} g(e^i)de^i\right]^2} \cdot \frac{w^H[\pi w^F - (1-\pi)c(1+r^*)]}{[\pi w^F + (1-\pi)\lambda w^H]^2} \cdot \lambda'(s_{t-1})$$

(9)

With the first fraction being unambiguously negative and the last term by assumption positive, it depends on the sign of the second fraction whether the derivative (9) is positive as in the baseline case, or not. If π is lower than $\frac{c(1+r^*)}{c(1+r^*)+w^F}$, the *ex post* level of human capital will be increasing in the human capital of the previous period.

In order to derive the condition for a brain drain to be beneficial for the economy's *ex post* level of human capital (and ultimately for growth), it is straightforward to compare the share of educated individuals in the case when the latter are allowed to migrate with probability π to the share of the educated when no such emigration is possible. In terms of comparative statics, the condition for a BBD is:

$$\left.\frac{ds_t}{d\pi}\right|_{\pi=0} > 0 \quad \text{where} \quad \frac{ds_t}{d\pi} = \frac{\partial s_t}{\partial \pi} + \frac{\partial s_t}{\partial e^*} \cdot \frac{\partial e^*}{\partial \pi}$$

(10)

The first component of (10) gives the negative brain drain effect:

$$\frac{\partial s_t}{\partial \pi} = -\frac{\int_{e_t^*}^{E} g(e^i)de^i[1 - \int_{e_t^*}^{E} g(e^i)de^i]}{[1 - \pi \int_{e_t^*}^{E} g(e^i)de^i]^2} < 0$$

(11)

The second component of (10) captures the positive brain gain effect: By reducing the threshold ability level e^*, any increase in the emigration probability of the educated π is accompanied by a positive impact on the proportion of educated individuals s_t:

$$\frac{\partial e^*}{\partial \pi} = -\frac{[\lambda_t w^H + c(1+r^*)](w^F - \lambda_t w^H)}{[\pi w^F + (1-\pi)\lambda_t w^H]^2} < 0$$

(12)

and

$$\frac{\partial s_t}{\partial e^*} = -\frac{(1-\pi)g(e^*)}{[1 - \pi \int_{e_t^*}^{E} g(e^i)de^i]^2} < 0$$

(13)

Evaluating these countervailing effects at $\pi = 0$ and noting that the numerator of (11) is at most $\frac{1}{4}$ (cf. footnote 8) yields the following condition for a BBD:

$$\frac{g(e^{*NM})[\lambda_t w^H + c(1+r^*)](w^F - \lambda_t w^H)}{(\lambda_t w^H)^2} > \frac{1}{4}$$

(14)

Thus, if inequality (14) is satisfied, there exists a positive optimal emigration probability for the educated such that the brain gain effect dominates the brain drain.

The additional assumption of uniformly distributed abilities allows to illustrate the countervailing effects in a simple diagram: In figure 3, the dark area represents the brain

drain and the light area characterizes the brain gain. A BBD arises if the latter area is of larger size than the former.

Figure 3: Brain drain vs. brain gain effect, cf. Mountford (1997, 295)
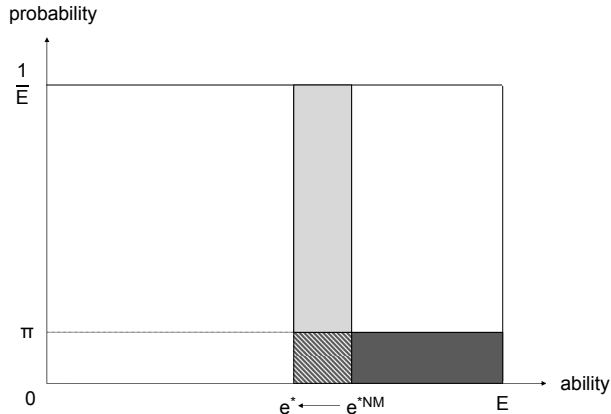
With uniformly distributed abilities, the condition for a BBD simplifies to:

$$1 - \frac{e^*}{E} < (1 - \pi)\frac{w^F - \lambda_t w^H}{\pi w^F + (1 - \pi)\lambda_t w^H}, \tag{15}$$

where the level of human capital has become $\left(1 - \frac{e^*}{E}\right)$. The circumstances under which this condition is likely to hold can be described as follows: If the probability of emigration for the educated is low, the level of human capital was previously low, and the foreign wage is very high relative to the home wage, a brain drain will benefit the human capital in the home economy (BBD).

5 Econometric Implementation

This section describes the econometric model and the data, and proceeds with a discussion of the estimation results. As a benchmark, we present estimation results using the panel data on the brain drain by Defoort (2008). Contrasting this benchmark analysis as well as the reviewed empirical analyses that rely on aggregate data about high-skilled south-north migration, we then uncover substantial heterogeneity in the net effect of the brain drain on human capital for different types of human capital.

From Theory to Empirics

Let us turn from the considered theoretical model and the aggregated level with one type of human capital (the educated or high-skilled) to disaggregated data distinguishing between several types of human capital (or professionals).

The core theoretical prediction is that due to an anticipated opportunity of migration to a high-wage economy for the highly educated, more individuals in developing countries will opt for higher education compared to the hypothetical situation in which no migration is possible. If this incentive effect exceeds the pure outflow of human capital, the sending

economies will experience a net gain in human capital with enhanced growth perspectives.

The descriptive evidence presented in section 3 suggests that the probability of successful emigration (in the sense of a perfect job match) from developing countries to developed countries greatly varies across different high-skilled occupational categories.

This observed heterogeneity poses the additional question whether developing economies are more likely to experience a BBD in terms of professionals with internationally transferable skills than in terms of professionals with rather country-specific skills. If we assume that the potential brain gain effect is increasing in the probability of successful emigration, the answer to this question crucially hinges on the exact characteristics of this relationship. In other words, it depends on whether the brain gain is disproportionately high in occupations with a large brain drain compared to occupations with a low incidence of brain drain. The theoretical model revisited in section 4 predicts that the higher the probability of emigration, the less likely is a BBD, *ceteris paribus* (cf. equation 15). Since the empirical model that is estimated in the following quantifies the elasticity of human capital with respect to brain drain, we might thus expect more conservative estimates of this elasticity for occupational categories with a high probability of emigration.

Against this background, the subsequent empirical analysis does not only test the hypothesis of a BBD, but also sheds light on these additional considerations.

5.1 Econometric Model and Data

Analysis with Data at the ISCO-88 Major Level

We propose the following log-linear econometric model in order to assess the net effect of the brain drain on human capital with employment data at the ISCO-88 major level:

$$\ln(h_{ij,2000}) = \alpha + \beta_1 \cdot \ln(h_{ij,1995}) + \beta_2 \cdot \ln(m_{ij,2000}) + \gamma_i + \delta_j + \epsilon_{ij} \tag{16}$$

The empirical counterpart of human capital is $h_{ij,2000}$, which measures the employment share of residents working in country i and occupational category j around 2000. These information are taken from LABORSTA, the main ILO database on labor statistics. $m_{ij,2000}$ denotes the emigration rate for country i and occupational category j and comes from Heuer (2010). It is defined as the share of native individuals from country i (residents plus migrants) employed in occupational category j that worked in the OECD around 2000:

$$m_{ij,2000} = \frac{M_{ij,2000}}{R_{ij,2000} + M_{ij,2000}} \tag{17}$$

$m_{ij,2000}$ thus gives the probability of working in the OECD around 2000 for a randomly chosen individual from country i with occupation j. In order to account for heterogeneity in the occupation-specific impact of emigration on human capital, we interact the brain drain variable with the occupation fixed effect δ_j in an alternative specification.

The empirical setup closely follows the theoretical reference model by Mountford (1997)

and is dynamic in the sense that the occupation-specific employment share in the base period 1995, $h_{ij,1995}$, is included as a regressor. This allows to account for "β-convergence"[10] in the accumulation process of human capital.

γ_i stands for country-specific effects that do not vary across occupational categories j (e.g. legislation on job protection). δ_j captures effects specific to the occupational group of *professionals*, which are common to all countries in the sample (e.g. honor associated with this occupation). ϵ_{ij} is an error term.

Subscript i refers to developing countries of emigration[11], and j distinguishes the two high-skilled occupational categories *professionals* and *technicians and associate professionals* (ISCO-88 majors 2 and 3, respectively), which generally require tertiary education. Thus, whereas the nature of the dataset is in principle cross-sectional, we dispose of two observations for each country, or cluster. The panel data estimation techniques fixed effects (FE) and random effects (RE) can be generally applied to cluster-sample data. Since it is very likely that the outcomes within a cluster are correlated, one should allow for an unobserved cluster effect (Wooldridge, 2009, 495). In this context, an unobserved effect at the country level could e.g. be the reputation enjoyed by employees in high-skilled jobs in general. This effect, however, is likely to be correlated also with the (natural logarithm of the) employment share in the base period, which is included as regressor in specification (16). Therefore, FE seem preferable to RE. FE estimation is furthermore appropriate, because the available sample cannot be considered as a random sample from a much larger universe of countries (Wooldridge 2009, 493).

Despite exhibiting the additional occupational dimension j, equation (16) differs from the empirical models estimated by Beine et al. (2003, 2008), Docquier et al. (2008), and Beine et al. (2011) with respect to three important aspects: First, the human capital variable in equation (16) measures the *ex post* level of human capital, thus excluding emigrants, whereas the human capital measure in the aforementioned studies includes the high-skilled emigrants, thus accounting for *ex ante* human capital. Therefore, specification (16) assesses the *net* effect of the brain drain on human capital in the sending economies, while the empirical models in the mentioned literature intend to capture the gross brain gain effect. Second, equation (16) reflects anticipatory expectation-building, given that the occupation-specific emigration rate is measured in the same period as the dependent variable. By contrast, the aforementioned empirical studies model the incentive effect from a retrospective point of view. Third, whereas the aforementioned cross-sectional studies

---

[10]   In cross-country growth regressions, introducing the initial income level (in some baseline period) is standard. If the estimated coefficient for this variable is negative, this is called "β-convergence" (Durlauf et al., 2005, 585).

[11]   All countries classified as low- or middle-income countries in 2000 by the World Bank are considered as 'developing' countries. These are countries with a GNI per capita ≤ 755 US\$ (low-income countries), and with a GNI per capita between 756 and 9,265 US\$ (middle-income countries).

assess the effect of high-skilled emigration on the change in human capital between 1990 and 2000 with the data by Docquier and Marfouk (2006), we consider the shorter time span from 1995 to 2000 in order to maximize observations on occupation-specific human capital. Thus, the estimation results presented in section 5.2 cannot be directly compared to those of earlier studies. Therefore, we use the panel dataset on the aggregate brain drain employed in Beine et al. (2011) to estimate a model that is similar to equation (16) in order to obtain relevant benchmark results.[12] Whereas the panel data by Defoort (2008) equally entail the possibility to extract unobserved heterogeneity at the country level as do the cluster data, this comes along with a severe econometric problem: If FE (least squares dummy variable, LSDV) estimation is applied to a dynamic panel data model, the correlation between the lagged dependent variable (LDV) and the error term biases the estimates (Nickell, 1981). Similarly as Beine et al. (2011), we therefore estimate the benchmark panel data model additionally with the linear dynamic panel-data estimator by Arellano and Bond (1991). It is important to note that the so-called Nickell bias does not materialize in the cases in which model (16) is estimated with FE using the cluster-sample data. The reason is that the demeaning of the variables in the latter case is not performed considering the time dimension but the occupational dimension, preventing thus the relevant correlation between the LDV and the error term.

The estimation results for equation (16) can be more easily compared to those in Groizard and Llull (2006, 2007b), who also use the *ex post* level of human capital to test for a BBD, accounting for both anticipatory and retrospective expectation-building with the brain drain rates for 1990 from Docquier and Marfouk (2006), respectively.

The occupation-specific emigration rates $m_{ij,2000}$ exhibit an important advantage when compared to the purely education-based emigration rates employed in existing empirical analyses: By construction, the former account for the fact that skills are only imperfectly transferable internationally, because they exclude emigrated professionals who did not manage to find a job as a professional in the OECD. Therefore, the former emigration rates are lower than the latter (Heuer, 2010). Whereas this implies that the occupation-specific emigration rates are more conservative empirical measures of the migration prospect than the purely education-based counterparts, they capture exactly the emigration potential that is relevant for the decision to enroll in some type of tertiary education: The incentive mechanism is likely to operate in the case of observed south-north migration with a perfect job match for professionals, it is however unlikely to be at work in the case of observed emigration of professionals from developing countries who work as a taxi driver or caretaker

---

[12]   Contrasting equation (16), the benchmark model estimated with the panel data by Defoort (2008) has time t as a second dimension, with $t \in [1975, 1980, 1985, 1990, 1995, 2000]$. The human capital variable in this case, $\ln(h_{it})$, is defined as the share of country i's resident labor force that is high-skilled in t. The brain drain variable, $\ln(m_{it})$, measures the share of the high-skilled native labor force from developing country i living in one of the six main OECD receiving countries in t.

in the receiving OECD countries. According to Beine et al. (2008, 632), the incentive effect is not determined solely by a higher probability of emigration when educated, but it is importantly linked to the possibility of accessing legal, high-skilled jobs. At the same time, the occupation-based measures of the probability of emigration to the OECD are less conservative estimates of the brain drain effect compared to the education-based counterparts. This tendency is weakened to some extent, however, because the occupation-specific emigration rates also include individuals who obtained their university degree in one of the receiving OECD countries.[13] [14]

In order to allow for non-linearities in the relation between *ex post* human capital and the brain drain, we additionally include the square of the latter variable.[15]

Whereas the technique of FE estimation by construction impedes the inclusion of time- (in this case occupation-)invariant regressors, it admits the interaction of time-varying regressors with time-invariant control variables. To allow for inter-regional heterogeneity, we interact the brain drain measure with a set of dummies for the different world regions.[16]

An important econometric concern is possible endogeneity of the brain drain variable in equation (16). On the one hand, this might be due to omitted variable bias. E.g., the occupation-specific wage rate which varies both over countries and occupations might be correlated also with $m_{ij,2000}$.[17] On the other hand, reverse causality might also be a source of inconsistency. In order to address this concern, we instrument the brain drain variable with the constructed geographic component of emigration, relying on a similar procedure as initially proposed by Frankel and Romer (1999) in the empirical trade literature, and applied to the context of international migration by Felbermayr et al. (2010).

$$
\begin{aligned}
M_{hij} = & \zeta_1 + \zeta_2 \cdot \ln(dist_{hi}) + \zeta_3 \cdot combord_{hi} + \zeta_4 \cdot comlang_{hi} + \zeta_5 \cdot \ln(pop_h) \\
& + \zeta_6 \cdot \ln(pop_i) + \zeta_7 \cdot \ln(area_h) + \zeta_8 \cdot \ln(area_i) + \zeta_9 \cdot land_h + \zeta_{10} \cdot land_i \\
& + \zeta_{11} \cdot prof_j + \text{interaction terms} + \epsilon_{hij}
\end{aligned}
\tag{18}
$$

In equation (18), we regress the bilateral occupation-specific (j) migrant stocks of *professionals* and *technicians and associate professionals*[18] on the natural logarithm of the distance between sending (i) and receiving (h) countries, on dummy variables indicating

---

[13]   Whereas the same holds true for the widely used data by Docquier and Marfouk (2006), Beine et al. (2007) explicitly take into account immigrants' age of entry as a proxy for the country where they acquired their education.

[14]   It is thus implicitly assumed that the emigrants who went to university in the OECD would have pursued the same studies and acquired the same skills in the home countries if they had not emigrated.

[15]   The results from this robustness test are reported in table A.5 in the appendix.

[16]   These are the regional country groups defined by the World Bank. The results from the estimations including these regional interaction terms are reported in table A.5 in the appendix.

[17]   Ideally, we would like to control for the effect of the occupation-specific wage rate in equation (16). Whereas occupation-specific wage data is in principle available from Freeman and Oostendorp (2003), these data by detailed occupations cannot be adequately aggregated to the ISCO-88 major level.

[18]   Note that all generally time-varying variables in equation (18) refer to the year 2000.

a common border and a common language, on the natural logarithms of population size and area, as well as on dummies for landlockedness of the sending and receiving country, respectively. In the absence of geographical variables that vary over occupational categories, we add a dummy for *professionals* as well as interaction terms of all variables with this dummy. Since we also want to account for zero migrant stocks, we rely on the Poisson Pseudo Maximum Likelihood estimator by Santos Silva and Tenreyro (2006). Using the predicted bilateral migrant stocks aggregated over all receiving OECD countries, we construct occupation-specific emigration rates reflecting emigration that can be solely explained by geographical factors:

$$\widehat{m_{ij,2000}} = \frac{\sum_h \widehat{M_{hij,2000}}}{R_{ij,2000} + M_{ij,2000}} \tag{19}$$

Using $\widehat{m_{ij,2000}}$ as an instrument in fixed effects instrumental variables estimations (FE IV) allows us to check the exogeneity assumption for the brain drain variable. The relevance of this instrument can be analyzed relying on the first stage F test and the Kleibergen-Paap LM statistic, whereas the Kleibergen-Paap Wald F test provides some information on the strength of the considered instrument. We use the square of the constructed emigration rate, $(\widehat{m_{ij,2000}})^2$, as a second instrument in order to assess the validity of the constructed instrument via tests on overidentifying restrictions.[19]

Analysis with Data at the ISCO-88 Sub-Major Level

Due to restricted employment data availability, we cannot re-assess the dynamic model given by equation (16) with data on several sending countries at the ISCO-88 sub-major level. Occupation-specific employment shares (and consequently emigration rates) at the ISCO-88 sub-major level can only be calculated for 16 developing countries around 2000 using data from ILO and OECD (cf. Heuer 2010). Therefore, we propose the following modified log-linear model for the empirical assessment of the hypothesis of a BBD using the more disaggregated data at the ISCO-88 sub-major level:

$$\ln(h_{ij,2000}) = \alpha + \beta_1 \cdot \ln(w_{ij,2000}) + \beta_2 \cdot \ln(m_{ij,2000}) + \gamma_i + \delta_j + \epsilon_{ij} \tag{20}$$

As before, subscript i identifies the clusters (countries). Subscript j now refers to the eight sub-major occupational categories which generally require tertiary education.[20]

Specification (20) essentially differs from equation (16) in that it does not control for convergence forces via the inclusion of the level of human capital in some baseline period

---

[19] Potential instruments must be correlated with the endogenous regressor $\ln(m_{ij})$, be uncorrelated with the error term ϵ_{ij}, and vary both over countries and occupations. This latter requirement is essentially the reason why no further instrument (besides the constructed emigrant share) is available.

[20] These are: *Physical, mathematical and engineering science professionals*, *life science and health professionals*, *teaching professionals*, and *other professionals* (sub-majors 21-24), as well as *physical and engineering science associate professionals*, *life science and health associate professionals*, *teaching associate professionals*, and *other associate professionals* (sub-majors 31-34).

(due to data unavailability). On the one hand, the theoretical benchmark model suggests that this dynamic component is an important explanatory factor of the evolution of an economy's human capital stock over time. Furthermore, in cross-sectional analyses in the empirical growth literature the inclusion of the income level in the baseline period as a regressor accounts for possible convergence of countries to their own (or a common) steady-state growth path (Durlauf et al., 2005, 578). If the LDV is indeed an important explanatory factor of human capital that is missing in specification (20), the estimation results will suffer from omitted-variable bias.[21] On the other hand, however, it will turn out that the estimation results from equation (16) weaken these concerns somewhat, since the coefficient on the LDV is found insignificant in most specifications.

Specification (16) furthermore differs from equation (20) with respect to the inclusion of the natural logarithm of the average monthly wage of male workers, $\ln(w_{ij,2000})$. At the more disaggregated ISCO-88 sub-major level, concerns about aggregating wage information by detailed occupations without accounting for empirical employment weights seem less severe. We therefore assign equal weights to the wage rates reported by detailed occupations (available from Freeman and Oostendorp 2003) in order to calculate average wages for ISCO-88 sub-majors 21-24 and 31-34.[22] This is possible for 11 out of the 16 countries for which employment data are available.

In order to account for heterogeneity in the effect of the brain drain on human capital, we interact the brain drain variable with a set of dummy variables for the different high-skilled sub-major categories (δ_j). Sub-major category 23 (*teaching professionals*) serves as the reference category. As before, γ_i denotes country-specific effects that will be extracted via the use of panel data estimation techniques, and ϵ_{ij} is the error term.

As in the case with specification (16), we instrument the brain drain variable with the constructed geographic component of emigration,[23] as well as with the squared value thereof. As a further robustness check, we include the square of the brain drain variable, and interact the brain drain variable with a dummy for Eastern European and Central Asian countries.[24]

Summary statistics of all variables, listings of the considered countries, and the estimation results from the robustness tests are provided in the appendix.

---

[21]   More specifically, if the potentially omitted variable is positively related to the dependent variable and negatively correlated with the brain drain variable, the coefficient on the latter variable, $\widehat{\beta_1}$, will be underestimated. The intuition on the relation between the variables has been derived from the estimation results using the more aggregated data at the ISCO-88 major level.

[22]   The recoding of the detailed occupations to the ISCO-88 sub-major occupational categories is based on the table of translation of the ILO October Inquiry.

[23]   This instrument variable has been constructed with the more disaggregated information on bilateral migrant stocks relying on specification (18).

[24]   The results are reported in table A.5 in the appendix.

5.2 Estimation Results

Results from Benchmark Analysis with Education-Based Data

Table 1 reports the estimation results from the benchmark model, relying on the panel dataset on south-north migration by educational categories from Defoort (2008). Considering two different (nested) samples, the results suggest a statistically significant negative relation between the share of the labor force with tertiary education (13 years or more of education) and the contemporaneous share of the tertiary-educated native labor force that live in one of the six main immigrant-receiving OECD countries. This finding is robust to estimation by FE[25], GMM[26] (to account for the endogeneity of the LDV), and to GMM estimation instrumenting the brain drain variable with lagged values. Thus, whereas Beine et al. (2011), using the same data source, find a positive relation between the share of the *native* labor force with tertiary education and the *lagged* high-skilled emigration rate which they interpret in favor of the incentive effect, the results in table 1 suggest that the proposed brain gain effect is not strong enough to compensate for the brain drain. The reported evidence thus rejects the hypothesis of a general BBD.

The estimation results furthermore report a robust positive coefficient for the LDV. Considering the impact of the lagged level of human capital on the growth rate of human capital between $t-1$ and t as proposed in Beine et al. (2011)[27], the estimated coefficient varies between -0.731 and -0.596, thus indicating convergence in the accumulation process of human capital.

Table 1: Estimation results from benchmark analysis with panel data by Defoort (2008). Dependent variable: $\ln(h_t)$.

	Full Sample			Reduced Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(h_{t-1})$	0.269***	0.347***	0.386***	0.350***	0.319***	0.404***
	(0.0422)	(0.0985)	(0.0949)	(0.0970)	(0.0737)	(0.0994)
$\ln(m_t)$	-0.333***	-0.395***	-0.239***	-0.473***	-0.576***	-0.503**
	(0.0632)	(0.0722)	(0.0902)	(0.151)	(0.206)	(0.231)
Constant	-3.581***	-	-	-2.444***	-	-
	(0.300)	-	-	(0.482)	-	-
Observations	662	512	512	95	69	69
No. of countries	150	129	129	26	18	18
R-squared (within)	0.803	-	-	0.877	-	-

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, adjusted for clustering on countries, in parentheses. In columns (1)-(3), the full sample of developing countries (low- and middle-income countries according to the World Bank classification in 2000) available from Defoort (2008) is considered. In columns (4)-(6), only developing countries considered in the main analyses with data at the ISCO-88 major level are included. Columns (1) and (4) report results from FE estimation. Columns (2) and (5) report results from GMM estimation, and columns (3) and (6) those from GMM in which $\ln(m_t)$ is instrumented with lagged values. Time fixed effects are not reported.

---

[25] Throughout the paper, FE refers to fixed effects estimation relying on the within transformation.

[26] GMM here refers to the linear dynamic panel-data estimator proposed by Arellano and Bond (1991).

[27] This is tantamount to subtracting $\ln(h_{t-1})$ from both sides of the model, implying that the coefficient of the LDV is reduced by 1, while the estimated coefficients of the other regressors are unaffected.

Results from Analysis with Occupation-Specific Data (ISCO-88 Major)

The estimation results for different variants of model (16) are reported in table 2. Columns (1) and (2) contain results from pooled ordinary least squares estimation (POLS), and the remaining columns those from FE and FE IV estimation. In the FE IV estimations, the brain drain variable is instrumented with the geographic share of emigration, which is constructed relying exclusively on the (respectively) relevant sample countries and the high-skilled major occupational categories 2 and 3 of ISCO-88. All standard errors are adjusted for clustering on countries. In columns (1)-(9), a sample of 27 developing countries for which all relevant variables are available is considered. Excluding the LDV, an enlarged sample of 54 developing countries of emigration is considered in columns (10) and (11).

Once the interaction term of the brain drain variable with the occupation fixed effect is included (columns 6-11), the hypothesis of the F test that all country fixed effects are equal to zero (not reported) can be rejected (at the 10- or 1-% level for columns (6) and (8), (10), respectively). This implies that POLS is inappropriate. The conducted Hausman tests (not reported) that compare the FE estimation results from columns (6), (8) and (10) to those of RE estimation suggest that FE is preferred to RE (the hypothesis that the difference in coefficients is not systematic can always be rejected at the 1-% level).

The estimated coefficient on the brain drain variable is negative and statistically significant at the 1-% level in the FE and FE IV estimations, and at the 5-% level in the POLS estimation. It is smallest in absolute terms in the latter estimation (2), and largest in the FE IV estimation reported in column (5): The elasticity of the employment share of the two high-skilled occupational categories with respect to the relevant share of the occupation-specific native emigrant population in 2000 varies between -0.112 and -2.862. The elasticity estimated in the benchmark analysis with panel data on 26 of the 27 developing countries amounts to approximately -0.5, figuring below the results from FE and FE IV estimation on the cluster dataset. Furthermore, the reported estimates on the impact of the brain drain on the *ex post* level of human capital are larger in absolute terms than those obtained from the cross-sectional analyses in Groizard and Llull (2006, 2007b).[28] This might be due to the use of the more realistic (conservative) measures of the incentive effect, or to the extraction of unobserved heterogeneity at the country level, which potentially biases the estimates in cross-sectional analyses relying on OLS.

Columns (6)-(11) suggest that the negative impact of the brain drain on human capital was larger for *professionals*: A 1% higher emigration rate of *professionals* is accompanied by an approximately 1% lower employment share of the latter in the sending countries, whereas a 1% higher emigration rate of *technicians and associate professionals* translates only into a 0.8% lower employment share *ceteris paribus* (cf. column 6). This finding implicitly suggests a (larger) brain gain effect for the latter occupational category. The coefficients are larger in absolute terms when the brain drain variable is instrumented.

---

[28] Modeling either anticipatory or retrospective expectation-building, in Groizard and Llull (2006, 2007b) the estimated coefficients for the high-skilled emigration rate in 1990 vary between -0.256 and -0.635.

Table 2: Estimation results using data at the ISCO-88 major level. Dependent variable: $\ln(h_{2000})$.

	(1) POLS	(2) POLS	(3) FE	(4) FE	(5) FE IV	(6) FE	(7) FE IV	(8) FE	(9) FE IV	(10) FE	(11) FE IV
$\ln(h_{1995})$	0.635***	0.535**	0.473**	0.212	-0.431	0.203	-0.120				
	(0.191)	(0.199)	(0.222)	(0.214)	(0.283)	(0.167)	(0.178)				
δ	0.067	0.063	0.008	0.150*	0.503**	-0.403**	-0.244*	-0.413*	-0.252*	-0.421**	0.108
	(0.069)	(0.073)	(0.091)	(0.087)	(0.213)	(0.188)	(0.148)	(0.207)	(0.132)	(0.167)	(0.198)
$\ln(m_{2000})$		-0.112**		-0.824***	-2.862***	-0.844***	-1.866***	-1.086***	-1.672***	-0.763***	-1.675***
		(0.048)		(0.260)	(0.691)	(0.242)	(0.328)	(0.248)	(0.100)	(0.102)	(0.287)
$\ln(m_{2000}) \times \delta$						-0.196***	-0.202***	-0.198***	-0.201***	-0.176***	-0.100*
						(0.043)	(0.033)	(0.046)	(0.031)	(0.039)	(0.053)
Constant	-0.860*	-1.444**	-1.242**	-4.429***		-4.512***		-5.746***		-5.163***	
	(0.458)	(0.606)	(0.548)	(1.222)		(1.030)		(0.783)		(0.351)	
Observations	54	54	54	54	54	54	54	54	54	108	108
No. of countries	27	27	27	27	27	27	27	27	27	54	54
R-squared (within)	0.537	0.626	0.416	0.594	-0.493	0.760	0.486	0.719	0.575	0.728	0.091
First stage F test					13.58		10.32		13.05		7.059
Kleibergen-Paap LM test					3.464		4.380		5.426		5.976
Kleibergen-Paap LM p-value					0.063		0.036		0.020		0.015
Kleibergen-Paap Wald F					13.58		10.32		13.05		7.059
Endog. test					3.747		5.507		6.993		11.54
Endog. test p-value					0.053		0.019		0.008		0.001

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, adjusted for clustering on countries, in parentheses.

Major category 3 (*technicians and associate professionals*) serves as reference category for the occupation fixed effect and the interacted brain drain variable.

In columns (5), (7), (9) and (11), $\ln(m_{2000})$ is instrumented with the constructed geographic component of emigration ($\widehat{m_{2000}}$), considering only the relevant sample countries and occupational categories.

The two samples include developing countries (low- and middle-income countries according to the World Bank classification in 2000).

In columns (10) and (11), a larger sample of 54 developing countries can be considered due to the exclusion of $\ln(h_{1995})$.

The coefficient on the LDV is positive and below 1 in the POLS and FE estimations, but it loses its statistical significance once country-level fixed effects are extracted and the brain drain variable is introduced as a further regressor. This result contrasts the robust and statistically significant positive effect of the LDV in the benchmark analysis with the panel data, and might be explained by the different model framework: Relying on cluster-sample data, the FE and FE IV estimation results in table 2 explain variation across two high-skilled occupational categories at one point in time, whereas the FE and GMM estimation results in table 1 focus on variation over time. Thus, the benchmark analysis considers a real dynamic panel data model, and the inclusion of the LDV in the latter context is more closely related to the dynamic modeling of human capital as proposed by Mountford (1997) than it is in the context of the cluster-sample model. The result that the estimated coefficient on the LDV is not significantly different from zero suggests the absence of occupation-specific convergence forces in human capital accumulation. Therefore, the exclusion of the LDV in the subsequent analysis with data at the ISCO-88 sub-major level due to scarce data availability might not be as problematic as one might expect, because it might not introduce an omitted variable bias at all.

Excluding the LDV in the estimations reported in columns (8)-(11), the impact of the brain drain variables remains robust also when considering an enlarged sample of 56 developing sending economies (columns 10 and 11).

Both sign and significance of the fixed effect for *professionals* vary across the different estimations. Adding the interacted brain drain variable in columns (6)-(9) yields a negative and statistically significant effect at the 5- or 10-% level, suggesting a smaller employment share for *professionals* than for *technicians and associate professionals*, *ceteris paribus*.

The hypothesis of exogeneity of the brain drain variable can be rejected in the different FE IV estimations, albeit on different significance levels. The reported test statistics indicate that the constructed instrument is relevant: The first stage F statistic exceeds the critical value of 10 (Staiger and Stock, 1997) except in column (11), and the null hypothesis of underidentification of the Kleibergen-Paap LM test can always be rejected at the 5- or 10-% level. The Kleibergen-Paap Wald F statistic points to the instrument's strength: The comparison of the test values with the critical values proposed by Stock and Yogo (2005) suggests a bias relative to OLS on the demeaned data of less than 15% in the estimations reported in columns (5), (7) and (9), and of less than 20% for column (11). If the square of the constructed emigration rate is included as a second instrument in the FE IV estimations (not reported), the null hypothesis of valid instruments of the Hanson J statistic cannot be rejected at reasonable levels of statistical significance in specifications (5), (7) and (9). In specification (11) by contrast, the null hypothesis can be rejected at the 10-% level, casting some doubt on the exogeneity of the excluded instruments.

The estimation results from the robustness tests reported in table A.5 in the appendix reveal that the negative impact of the brain drain variable on the *ex post* level of human capital is non-linear (it is increasing in the brain drain rate, yet at a decreasing rate due to the natural logarithm). Furthermore, the net brain drain seems to have been less strong

for countries in Eastern Europe and Central Asia, in East Asia and the Pacific, and in South Asia compared to countries in Latin America and the Caribbean, in Sub-Saharan Africa, and in the Middle East and North Africa.

Results from Analysis with Occupation-Specific Data (ISCO-88 Sub-Major)

Table 3 reports the estimation results for equation (20). Columns (1) and (2) report results from POLS, columns (3)-(5), (8) and (9) from FE and RE, and columns (6), (7), (10) and (11) from FE IV and RE IV estimation. In the latter estimations, the brain drain variable is instrumented with the geographic share of emigration, which is constructed relying on the 16 developing countries for which employment data at ISCO-88 are available and the high-skilled sub-major occupational categories 21-24 and 31-34. All standard errors – except in column (11) – are adjusted for clustering on countries. Whereas the disaggregated employment data at the ISCO-88 sub-major level are in principle available for 16 developing countries, the sample considered in table 3 consists of only 11 countries due to unavailable wage data for 5 countries.

The reported results from the analysis relying on the more disaggregated data confirm the robust negative relation between the occupation-specific *ex post* level of human capital and the corresponding emigration rates in 2000: The estimated coefficient on the brain drain variable is statistically significant at the 1-% level in the specifications reported in columns (4)-(7), varying between -0.703 and -1.036 across FE, RE, FE IV, and RE IV estimation. The effect is not significant at any reasonable significance level in the POLS estimation. Since the hypothesis of the F test that all country fixed effects are equal to zero (not reported) can be rejected at the 1-% level for the specification reported in column (4), POLS seems inappropriate for the estimation of the model. The results from column (5) are the preferred ones out of (4)-(7), because the Hausman test for the FE and RE estimation (columns 4 and 5, not reported) favors RE estimation, and because the endogeneity test reported in column (6) suggests that the brain drain variable should be treated as exogenous.

Columns (8)-(11) report the estimation results from equation (20) including interaction terms of the brain drain variable with the occupation fixed effects. The reference category is sub-major 23 (*teaching professionals*). The Hausman test (not reported) comparing the estimates from columns (8) and (9) suggests that RE is the appropriate estimation technique. As explained below, the specifications that do not instrument the brain drain variable seem more appropriate. The preferred results from column (9) suggest a negative impact of the brain drain on *ex post* human capital only for *life science and health professionals* (sub-major 22). Whereas the estimated coefficient on the relevant interaction term is statistically significant only at the 10-% level in the preferred RE estimation, it is so at the 1-% level in the FE estimation (8). The results from the latter (FE) and the FE IV estimation in column (10) furthermore suggest a general negative impact of the brain drain on *ex post* human capital as found in columns (4)-(7), appearing smallest for *physical and engineering science associate professionals* and for *life science and health associate*

professionals (sub-majors 31, 32). However, it has to be stressed that these findings are neither very robust over the different estimation techniques, nor to slight variations in the sample.

The coefficient on the wage rate is positive and statistically significant at the 1- or 5-% level (except in columns 9 and 11) once heterogeneity at the country level is extracted and the brain drain variable is included as a further regressor: A 1 % higher average monthly wage is associated with a 0.3% to 0.8% higher occupation-specific employment share, *ceteris paribus*.

The first stage F tests of the FE IV estimations (columns 6 and 10) are quite high, pointing to the relevance of the employed instrument. In addition, the null hypothesis of underidentification of the Kleibergen-Paap LM test can be rejected at the 5-% level in both estimations. The Kleibergen-Paap Wald F statistic suggests that the instrument is strong (the bias relative to OLS on the demeaned data is less than 10% for column 6, and less than 15% for column 10). Introducing the square of the constructed emigration rate as a second instrument, the null hypothesis of valid instruments of the Hanson J statistic can be rejected at the 10- and 1-% level in specifications (6) and (10), respectively. This again questions the exogeneity assumption for the excluded instruments. The endogeneity tests reported in columns (6) and (10) suggest that the brain drain variable should be treated as exogenous (the null hypothesis of exogeneity cannot be rejected at reasonable significance levels).[29] Therefore, the results from FE and RE in columns (4), (5) and (8), (9) are preferred over those from FE IV and RE IV (columns 6, 7 and 10, 11).

The estimation results from the robustness tests with the disaggregated data included in table A.5 confirm the non-linearity in the negative impact of the brain drain on the *ex post* level of human capital that is obtained also with the data at the ISCO-88 major level.

---

[29] We also estimated the specification considered in columns (8)-(11) instrumenting all terms comprising the brain drain variable. To this end, we generated interaction terms of the constructed instrument and the occupation fixed effects in order to dispose of sufficient instruments. However, these instruments performed poorly – the first stage F values were very low. The results are not reported.

Table 3: Estimation results using data at the ISCO-88 sub-major level. Dependent variable: $\ln(h_{2000})$.

	(1) POLS	(2) POLS	(3) FE	(4) FE	(5) RE	(6) FE IV	(7) RE IV	(8) FE	(9) RE	(10) FE IV	(11) RE IV
$\ln(w_{2000})$	-0.016	0.302	-0.121	0.338**	0.502***	0.456**	0.794***	0.295**	0.304	0.353**	1.144*
	(0.120)	(0.229)	(0.280)	(0.112)	(0.153)	(0.190)	(0.224)	(0.129)	(0.231)	(0.157)	(0.622)
$\ln(m_{2000})$		-0.247		-0.783***	-0.703***	-0.985***	-1.036***	-0.904***	-0.263	-1.058***	-2.128*
		(0.172)		(0.111)	(0.115)	(0.171)	(0.177)	(0.137)	(0.218)	(0.185)	(1.253)
$\ln(m_{2000}) \times \delta_{21}$								0.005	-0.085	0.074	1.306
								(0.055)	(0.129)	(0.118)	(0.988)
$\ln(m_{2000}) \times \delta_{22}$								-0.149***	-0.056*	-0.086	1.392
								(0.038)	(0.029)	(0.094)	(1.030)
$\ln(m_{2000}) \times \delta_{24}$								-0.068	-0.046	0.017	1.703
								(0.054)	(0.132)	(0.098)	(1.248)
$\ln(m_{2000}) \times \delta_{31}$								0.263***	0.166	0.338***	1.570
								(0.046)	(0.113)	(0.113)	(0.987)
$\ln(m_{2000}) \times \delta_{32}$								0.266*	0.096	0.346***	1.591
								(0.135)	(0.109)	(0.126)	(1.042)
$\ln(m_{2000}) \times \delta_{33}$								0.102	-0.158	0.198	1.298
								(0.136)	(0.290)	(0.169)	(1.040)
$\ln(m_{2000}) \times \delta_{34}$								0.083	0.148	0.155	1.845
								(0.082)	(0.184)	(0.106)	(1.247)
Constant	-3.251***	-6.273**	-2.584	-8.617***	-9.384***			-8.818***	-6.356***	-12.584***	-19.193**
	(0.683)	(2.102)	(1.762)	(1.043)	(1.211)			(1.210)	(2.295)	(1.989)	(8.795)
Observations	83	83	83	83	83	83	83	83	83	83	83
No. of countries	11	11	11	11	11	11	11	11	11	11	11
R-squared (within)	0.480	0.520	0.666	0.869	0.864	0.856	0.850	0.915	0.794	0.911	0.540
First stage F						27.17				14.24	
Kleib.-Paap LM test						4.698				4.779	
Kleib.-Paap LM p-value						0.030				0.029	
Kleib.-Paap Wald F						27.17				14.24	
Endog. test						2.110				1.328	
Endog. test p-value						0.146				0.249	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors, adjusted for clustering on countries – except in column (11) –, in parentheses.

Occupation fixed effects are not reported in columns (3)-(11).

Sub-major category 23 (*teaching professionals*) serves as occupational reference category in the specifications considering the interacted brain drain variable.

In columns (6), (7), (10) and (11), $\ln(m_{2000})$ is instrumented with the constructed geographic component of emigration $(\widehat{m_{2000}})$, considering the 16 developing countries for which employment data at ISCO-88 is available and sub-major categories 21-24 and 31-34.

The considered sample includes developing countries (low- and middle-income countries according to the World Bank classification in 2000).

6 Summary and Conclusions

This paper has tested the hypothesis of a BBD relying on the novel datasets by Heuer (2010) which provide disaggregated data on the brain drain distinguishing between, respectively, two and eight high-skilled occupational categories according to the ISCO-88. As a benchmark analysis, the net effect of the brain drain on *ex post* human capital has been estimated relying on the panel dataset by Defoort (2008), which includes aggregate information on high-skilled emigration from developing countries to the six main OECD receiving countries.

The occupation-based cross-sectional data exhibit three important benefits when used to study the net effect of the brain on human capital in the sending economies: (i) They allow to test the hypothesis of a BBD for different types of human capital. (ii) In analogy to panel data, the cluster-sample structure of the data permits the extraction of observed and unobserved heterogeneity at the country level via the use of panel data estimation techniques. (iii) Excluding emigrated *professionals* and *technicians and associate professionals* who did not manage to find adequate jobs in the OECD, the occupation-specific emigration rates by construction account for the fact that skills are only imperfectly transferable internationally. Thus, compared to the conventionally used education-based emigration rates, the occupation-based rates are equally more realistic and hence also more conservative in capturing the potential incentive effect or brain gain.

Modeling anticipatory expectation-building and accounting for possible convergence forces in the accumulation process of human capital, the estimations with data classified at the ISCO-88 major level reveal a robust negative effect of the occupation-specific emigration rates on the sending countries' employment shares, which are used as a measure of occupation-specific *ex post* human capital. This finding suggests that – on average – the proposed brain gain effect was either inexistent or too small compared to the brain drain. Thus, this rejects the hypothesis of a BBD. The estimated average elasticity of the employment share of the two high-skilled occupational categories with respect to the relevant share of the occupation-specific native emigrant population in 2000 is larger (in absolute terms) than the elasticity estimated in the benchmark model relying on education-based panel data, and exceeds the estimates obtained in existing studies that use education-based cross-sectional data. One reason for this difference in magnitude might be the use of a more conservative measure of the incentive mechanism as explained above. In addition, the results from the existing cross-sectional analyses relying on OLS might be biased due to unobserved heterogeneity at the country level.

The obtained negative relation between the brain drain and the *ex post* level of human capital is robust to the instrumentation of the brain drain variable with the constructed geographic component of emigration, to the exclusion of the LDV, to the inclusion of the squared brain drain variable, to variations in the sample, and to the use of the more disaggregated data at the ISCO-88 sub-major level.

The negative effect of the brain drain on human capital turns out to be significantly

stronger for *professionals* – the occupational category with the highest educational requirements – compared to the occupational group of *technicians and associate professionals*. This finding implicitly suggests that the brain gain in terms of *professionals* was too small to compensate for the higher incidence of brain drain of the latter when compared to *technicians and associate professionals* (cf. section 3). In addition, the estimated negative effect appears stronger for countries in Latin America and the Caribbean, in Sub-Saharan Africa, and in the Middle East and North Africa compared to countries in Eastern Europe and Central Asia, in East Asia and the Pacific, as well as in South Asia.

The estimation results from the further disaggregated data suggest the existence of heterogeneity in the effect of the brain drain on human capital across high-skilled occupational categories differing manifestly with respect to their degree of international transferability of skills. The preferred specification suggests a negative impact of the brain drain on *ex post* human capital only for *life science and health professionals*. However, the estimation results from the more disaggregated data are not very robust and should hence be treated with caution.

As data availability improves, it would be desirable to re-assess the hypothesis of a BBD using richer data on occupation-specific employment, or on graduation by program of study. Furthermore, future work on this topic might include the analysis of spillover effects of occupation-specific brain drain on the human capital endowment in terms of other occupations: E.g., one might suppose that a higher probability of emigration of physicians incentivizes some individuals in developing countries, who otherwise would have studied law, to study medicine, *ceteris paribus*.

References

Adams, R. H. (2003). International Migration, Remittances and the Brain Drain: A Study of 24 Labor-Exporting Countries. World Bank Policy Research Working Paper No. 3069, June.

Alburo, F. A. and D. I. Abella (2002). Skilled Labour Migration from Developing Countries: Study on the Philippines. International Migration Papers 51, ILO, July.

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies 58(2)*, 277–297.

Batista, C., A. Lacuesta, and P. C. Vicente (2011). Testing the 'brain gain' hypothesis: Micro evidence from Cape Verde. *Journal of Development Economics, forthcoming.*

Beine, M., F. Docquier, and C. Oden-Defoort (2011). A Panel Data Analysis of the Brain Gain. *World Development 39(4)*, 523–532.

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

Beine, M., F. Docquier, and H. Rapoport (2003). Brain Drain and LDCs' Growth: Winners and Losers. IZA Discussion Paper No. 819, July.

Beine, M., F. Docquier, and H. Rapoport (2007). Measuring International Skilled Migration: A New Database Controlling for Age of Entry. *The World Bank Economic Review 21(2)*, 249–254.

Beine, M., F. Docquier, and H. Rapoport (2008). Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *The Economic Journal 118(528)*, 631–652.

Bhargava, A. and F. Docquier (2008). HIV Pandemic, Medical Brain Drain, and Economic Development in Sub-Saharan Africa. *The World Bank Economic Review 22(2)*, 345–366.

Bhargava, A., F. Docquier, and Y. Moullan (2011). Modeling the Effects of Physician Emigration on Human Development. *Economics and Human Biology 9(2)*, 172–183.

Bhorat, H., J.-B. Meyer, and C. Mlatsheni (2002). Skilled Labour Migration from Developing Countries: Study on South and Southern Africa. International Migration Papers 52, ILO, July.

Carrington, W. J. and E. Detragiache (1998). How Big Is the Brain Drain? IMF Working Paper 98/102, July.

Checchi, D., G. De Simone, and R. Faini (2007). Skilled Migration, FDI and Human Capital Investment. Centro Studi Luca d'Agliano Development Studies Working Papers No. 235, May.

Clemens, M. (2007). Do Visas Kill? Health Effects of African Health Professional Emigration. Center for Global Development Working Paper No. 114.

Clemens, M. A. and G. Pettersson (2008). New Data on African Health Professionals Abroad. *Human Resources for Health 6(1).*

Commander, S., R. Chanda, M. Kangasniemi, and L. A. Winters (2004). Must Skilled Migration Be a Brain Drain? Evidence from the Indian Software Industry. IZA Discussion Paper No. 1422, December.

Defoort, C. (2008). Long-term Trends in International Migration: An Analysis of the Six Main Receiving Countries. *Population 63(2)*, 285–317.

Di Maria, C. and E. Lazarova (2009). Migration, Human Capital Formation and Growth: An Empirical Investigation. Queen's University Management School Working Paper Series No. MS_WPS_ECO_10_3, December.

Docquier, F., O. Faye, and P. Pestieau (2008). Is Migration a Good Substitute for Education Subsidies? *Journal of Development Economics 86* (2), 263–276.

Docquier, F., B. Lowell, and A. Marfouk (2008). A Gendered Assessment of the Brain Drain. World Bank, Policy Research Working Paper 4613, May.

Docquier, F. and A. Marfouk (2006). International Migration by Education Attainment, 1990-2000. In Özden, C. and Schiff, M. (Ed.), *International Migration, Remittances & the Brain Drain*, Chapter 5, pp. 151–199. World Bank/ Palgrave Macmillan.

Docquier, F. and H. Rapoport (2008). Brain Drain. In Durlauf, S. N. and Blume, L. E. (Ed.), *The New Palgrave Dictionary of Economics* (The New Palgrave Dictionary of Economics Online ed.). Palgrave Macmillan. URL: $http : //www.dictionaryofeconomics.com/article?id = pde2008\_B000186$ [visited on 12–28–2008].

Durlauf, S. N., P. A. Johnson, and J. R. W. Temple (2005). Growth Econometrics. In Aghion, P. and Durlauf, S. N. (Ed.), *Handbook of Economic Growth*, Volume 1A, Chapter 8, pp. 555–677. Elsevier B.V.

Faini, R. (2004). Trade Liberalization in a Globalizing World. IZA Discussion Paper No. 1406, November.

Felbermayr, G. J., S. Hiller, and D. Sala (2010). Does Immigration Boost Per Capita Income? *Economics Letters 107*, 177–179.

Frankel, J. A. and D. Romer (1999). Does Trade Cause Growth? *The American Economic Review 89(3)*, 379–399.

Groizard, J. L. and J. Llull (2006). Skilled Migration and Growth. Testing Brain Drain and Brain Gain Theories. DEA Working Paper No. 20, December.

Groizard, J. L. and J. Llull (2007a). Brain Drain and Human Capital Formation in Developing Countries. Are there Really Winners? DEA Working Paper No. 28, June.

Groizard, J. L. and J. Llull (2007b). Skilled Migration and Sending Economies. Testing Brain Drain and Brain Gain Theories. mimeo, October.

Ha, W., J. Yi, and J. Zhang (2009, August). Brain Drain, Brain Gain and Economic Growth in China. Human Development Research Paper 2009/37.

Hagopian, A., M. Thompson, F. M., K. Johnson, and L. Hart (2004). The Migration of Physicians from sub-Saharan Africa to the United States of America: Measures of the African Brain Drain. *Human Resources for Health 2(17)*.

Heuer, N. (2010). Occupation-Specific South-North Migration. Tübinger Diskussionsbeitrag Nr. 328.

ILO (1990). *International Standard Classification of Occupations: ISCO-88*. International Labour Office, Geneva.

Kangasniemi, M., L. A. Winters, and S. Commander (2007). Is the Medical Brain Drain Beneficial? Evidence from Overseas Doctors in the UK. *Social Science & Medicine 65*, 915–923.

Meyer, J.-B., M. Brown, and D. Kaplan (2000). Assessing the South African Brain Drain. A statistical comparison. DPRU Working Paper No. 00/40, July.

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

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica 49(6)*, 1417–1426.

OECD (2008). *A Profile of Immigration Populations in the 21st Century: Data from OECD Countries*. Organisation for Economic Co-operation and Development, Paris.

Pellegrino, A. (2002). Skilled Labour Migration from Developing Countries: Study on Argentina and Uruguay. International Migration Papers 58, ILO, July.

Santos Silva, J. and S. Tenreyro (2006). The Log of Gravity. *Review of Economics and Statistics 88(4)*, 641–658.

Staiger, D. and J. H. Stock (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica 65(3)*, 557–586.

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

Stark, O., C. Helmenstein, and A. Prskawetz (1998). Human Capital Depletion, Human Capital Formation, and Migration: A Blessing or a 'Curse'? *Economics Letters 60*(3), 363–367.

Stock, J. H. and M. Yogo (2005). Testing for Weak Instruments in Linear IV Regression. In Andrews, D. W. K and Stock, J. H. (Ed.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pp. 80–108. Cambridge University Press. Working paper version: NBER Technical Working Paper 284.

Thomas-Hope, E. (2002). Skilled Labour Migration from Developing Countries: Study on the Caribbean Region. International Migration Papers 50, ILO, July.

Watanabe, S. (1969). The Brain Drain from Developing to Developed Countries. *International Labour Review 4*, 400–435.

Wooldridge, J. M. (2009). *Introductory Econometrics, A Modern Approach* (Fourth ed.). South-Western, Mason, Ohio.

Data Sources

Country groups by region
  The World Bank, Data & Statistics
  URL: *http://go.worldbank.org/D7SN0B8YU0*   [visited on 11-17-2008].

Database on Immigrants in OECD Countries (DIOC)
  OECD (2008)
  URL: *http://www.oecd.org/document/51/0,3343,en_2649_33931_40644339_1_1_1_1,00.html*
  [visited on 11-05-2008].

Employment for detailed occupational groups by sex (SEGREGAT)
Total employment, by occupation (Main statistics, annual, 2C)
  ILO, LABORSTA Internet
  URL: *http://laborsta.ilo.org/*   [visited on 10-15-2009].

ILO October Inquiry: Industry groups and occupations
  ILO LABORSTA
  URL: *http://laborsta.ilo.org/applv8/data/to1ae.html*   [visited on 01-11-2009].

Panel Data on International Migration, 1975-2000
  M. Schiff and M. C. Sjoblom
  The World Bank, International Migration and Development Program
  URL: *http://go.worldbank.org/4IJSEHH7P0*   [visited on 07-15-2010].

The Occupational Wages Around the World (OWW) Database
  R. B. Freeman and R. H. Oostendorp (2003)
  URL: *http://www.nber.org/oww/*   [visited on 11-16-2008].

Translation from US OCC 2000 to ISCO-88
  J. Elliott and V. Gerova (2006)
  Centre for Longitudinal Studies, Institute of Education, Research Archive
  URL: *http://www.cls.ioe.ac.uk/text.asp?section=00010001000500160002*   [visited on
11-20-2008].

World Bank GNI per capita Operational Guidelines & Analytical Classifications
(low, lower middle, upper middle, and high income countries in 2000)
  The World Bank, Data & Statistics
  URL: *http://go.worldbank.org/U9BK7IA1J0*   [visited on 01-22-2009].

World Development Indicators (WDI)
  The World Bank (2008).

Appendix

Table A.1: Summary statistics of data used in benchmark analysis

Variable		Mean	Std. Dev.	Min	Max	Observations
$\ln(h)$	overall	-3.544105	1.198595	-6.907755	-1.354796	N = 812
	between		1.135356	-6.792231	-1.688204	n = 150
	within		.5154545	-5.583155	-2.200137	$\bar{T}$=5.4
$\ln(m)$	overall	-2.343035	1.400567	-8.100799	0	N = 812
	between		1.467711	-7.393097	-.0866283	n = 150
	within		.4617914	-3.919986	-.1584565	$\bar{T}$=5.4
$\ln(h)$	overall	-2.68141	.769152	-5.298317	-1.496109	N = 121
	between		.6345884	-4.237291	-1.688204	n = 26
	within		.4849395	-3.991508	-1.653832	$\bar{T}$=4.7
$\ln(m)$	overall	-2.402608	1.426429	-6.527544	-.1575796	N = 121
	between		1.458211	-5.459018	-.2662939	n = 26
	within		.3850024	-3.471134	-1.334081	$\bar{T}$=4.6

Compared to the sample of 27 countries considered in the analysis with data at the ISCO-88 major level (cf. table A.4,a), data for Puerto Rico is unavailable from the Defoort (2008) dataset.

Table A.2: Summary statistics of data at the ISCO-88 major level

Variable		Mean	Std. Dev.	Min	Max	Observations
$\ln(h_{2000})$	overall	-2.432882	.4918905	-3.675529	-1.678355	N = 54
	between		.370496	-3.041591	-1.872035	i = 27
	within		.3275345	-3.163083	-1.702681	$\bar{j}=2$
$\ln(h_{1995})$	overall	-2.528645	.5820469	-4.415968	-1.433931	N = 54
	between		.3721267	-3.258809	-1.877407	i = 27
	within		.4504588	-4.019664	-1.037627	$\bar{j}=2$
$\ln(m_{2000})$	overall	-2.981677	1.407876	-6.161968	-.2997081	N = 54
	between		1.397445	-5.963824	-.6472211	i = 27
	within		.257117	-3.676567	-2.286786	$\bar{j}=2$
$\ln(h_{2000})$	overall	-2.711421	.742055	-6.210789	-1.678355	N = 108
	between		.6594744	-5.419217	-1.872035	i = 54
	within		.3461268	-3.667837	-1.755004	$\bar{j}=2$
$\ln(m_{2000})$	overall	-3.146559	1.428156	-7.660114	-.2997081	N = 108
	between		1.390986	-7.200251	-.6472211	i = 54
	within		.3505252	-4.257023	-2.036095	$\bar{j}=2$

Table A.3: Summary statistics of data at the ISCO-88 sub-major level

Variable		Mean	Std. Dev.	Min	Max	Observations
$\ln(h_{2000})$	overall	-3.962989	.9358952	-7.449187	-2.496365	N = 83
	between		.5021395	-5.198358	-3.55281	i = 11
	within		.7972708	-6.372831	-2.51841	$\bar{j}=7.5$
$\ln(w_{2000})$	overall	6.547545	.7982681	4.401116	7.627324	N = 83
	between		.8091966	4.500611	7.158494	i = 11
	within		.359204	5.81578	7.540114	$\bar{j}=7.5$
$\ln(m_{2000})$	overall	-3.602496	1.270599	-7.597597	-.8357415	N = 83
	between		1.209691	-6.554455	-1.695767	i = 11
	within		.6360487	-5.757858	-1.96514	$\bar{j}=7.5$

Table A.4: List of countries included in the different specifications

EASTERN EUROPE AND CENTRAL ASIA		LATIN AMERICA AND CARIBBEAN	
Azerbaijan	b	Argentina	ab
Bulgaria	b	Belize	ab
Croatia	ab	Bolivia	ab
Czech Republic	abc	Brazil	b
Estonia	abc	Costa Rica	ab
Georgia	ab	Dominica	ab
Hungary	abc	Dominican Republic	ab
Kazakhstan	b	Ecuador	b
Kyrgyzstan	b	El Salvador	ab
Latvia	abc	Mexico	ab
Lituania	abc	Nicaragua	b
Macedonia	b	Panama	b
Moldova	b	Peru	ab
Poland	abc	Puerto Rico	ab
Romania	ab	Saint Lucia	ab
Russian Federation	ab	Trinidad and Tobago	ab
Slovakia	abc	Uruguay	b
Turkey	b		
Ukraine	abc		
SUB-SAHARAN AFRICA		EAST ASIA AND PACIFIC	
Botswana	ab	Cambodia	b
Ethiopia	b	Malaysia	b
Mauritius	c	Mongolia	c
Namibia	b	Philippines	b
South Africa	b	Thailand	c
Tanzania	b	Viet Nam	b
MIDDLE EAST AND NORTH AFRICA		SOUTH ASIA	
Algeria	b	Maldives	b
Egypt	ab	Pakistan	b
Oman	ab	Sri Lanka	b

a: Countries included in the estimations using data at the ISCO-88 major level (27).
b: Countries of the enlarged sample included in the estimations with data at the ISCO-88 major level (54).
c: Countries included in the estimations using data at the ISCO-88 sub-major level (11).

Table A.5: Estimation results from robustness tests. Dependent variable: $\ln(h_{2000})$.

	Data at ISCO-88 major level			Data at ISCO-88 sub-major level		
$\ln(w_{2000})$				0.381**	0.357***	0.410***
				-0.121	-0.095	-0.115
$\ln(m_{2000})$	-1.457***	-0.843***	-1.157***	-1.401***	-0.831***	-1.487***
	(0.320)	(0.061)	(0.304)	-0.102	-0.172	-0.137
$[ln(m_{2000})]^2$	-0.107**		-0.052	-0.084***		-0.087***
	(0.048)		(0.048)	-0.013		-0.02
$\ln(m_{2000}) \times \delta_j$	-0.102*	-0.159***	-0.128**			
	(0.056)	(0.032)	(0.052)			
$\ln(m_{2000}) \times lac$		-0.400*	-0.352*			
		(0.234)	(0.198)			
$\ln(m_{2000}) \times mena$		-0.230	-0.236			
		(0.298)	(0.301)			
$\ln(m_{2000}) \times eeca$		0.436***	0.361**		0.091	0.13
		(0.151)	(0.153)		-0.239	-0.108
$\ln(m_{2000}) \times eap$		0.244***	0.209**			
		(0.069)	(0.079)			
$\ln(m_{2000}) \times sa$		0.361***	0.241*			
		(0.100)	(0.142)			
Constant	-6.075***	-5.035***	-5.507***	-9.886***	-8.679***	-10.011***
	(0.488)	(0.341)	(0.498)	-0.972	-0.945	-0.814
Observations	108	108	108	83	83	83
No. of countries	54	54	54	11	11	11
R-squared (within)	0.774	0.813	0.821	0.892	0.87	0.895

Country and occupation fixed effects are not reported. *lac* stands for Latin America and Caribbean, *mena* for Middle East and North Africa, *eeca* for Eastern Europe and Central Asia, *eap* for East Asia and Pacific, and *sa* for South Asia. Sub-Saharan Africa is the reference region.