IAB Discussion Paper
Beiträge zum wissenschaftlichen Dialog aus dem Institut für Arbeitsmarkt- und Berufsforschung

No. 26/2007

# Inequality and the (Self-)Selection of International Migrants

## Theory and Novel Evidence

*Herbert Brücker, Cécily Defoort*

Bundesagentur für Arbeit

# Inequality and the (Self-)Selection of International Migrants

## Theory and Novel Evidence

*Herbert Brücker (IAB, Nuremberg and IZA, Bonn)*
*Cécily Defoort (EQUIPPE University of Lille and IRES – University of Louvain)*

Auch mit seiner neuen Reihe „IAB-Discussion Paper“ will das Forschungsinstitut der Bundesagentur für Arbeit den Dialog mit der externen Wissenschaft intensivieren. Durch die rasche Verbreitung von Forschungsergebnissen über das Internet soll noch vor Drucklegung Kritik angeregt und Qualität gesichert werden.

Also with its new series "IAB Discussion Paper" the research institute of the German Federal Employment Agency wants to intensify dialogue with external science. By the rapid spreading of research results via Internet still before printing criticism shall be stimulated and quality shall be ensured.

# Inequality and the (Self-)Selection of International Migrants: Theory and Novel Evidence

Herbert Brücker[*] Cécily Defoort[†]

This version: October 2007

**Abstract**

This paper analyses the (self-)selection of migrants between countries which have substantial differences in the inequality of earnings and income levels. In an extended version of the Roy-model we consider migration costs, which tend to grow less than proportional with the income level. As a consequence, migrants can be favourably self-selected although the inequality of earnings is larger in the destination relative to the sending country. Based on a novel panel data set, covering migration from 143 sending countries all over the world into the 6 main receiving countries in the OECD from 1975 to 2000, we examine the factors which drive the selection bias of the migrant population empirically. The descriptive statistics indicate that migrants tend to be positively (self-)selected although the inequality in earnings is larger in the destination relative to the sending countries. Our estimation results suggest that both, a higher inequality in the distribution of earnings in the receiving *and* the sending country increases the skill level of the migrant population relative to that of the population in the sending countries. Moreover, the positive selection bias decreases with the income level of the sending country at a given income differential. Finally, migration barriers and distance affect the selection bias positively.

*Keywords*: International migration, self-selection of migrants, inequality, human capital of migrants.
*JEL code*: F22.

[*]IAB, Nürnberg, and IZA, Bonn.
[†]EQUIPPE Universities of Lille and IRES – University of Louvain.
Corresponding author: Herbert Brücker, IAB, Weddigenstr. 20-22, Nürnberg, Tel.: +49-911-179-3807, Fax: +49-911-3298, Email: herbert.bruecker@iab.de.

# 1 Introduction

The skill bias of migration is highly relevant from both the perspective of the sending and receiving countries. In the traditional brain drain literature, which dates back to the 1960s and 1970s, economists and policy-makers were concerned that the loss of human capital associated with international migration is detrimental to economic development in the sending countries (Bhagwati and Hamada, 1974; Grubel and Scott, 1966; Kwok and Leland, 1982). Although the "new economics of the brain drain" literature suggests that international migration might foster human capital investment in the sending countries (Beine, Docquier, and Rapoport, 2001; Stark, Helmenstein, and Prskawetz, 1997), which in turn may increase economic growth and welfare (Mountford, 1997), there are still fears that labour mobility involves a net loss in human capital for the senders.

In contrast, from the perspective of the receiving countries, there are increasing concerns that the skill levels of migrants are declining over time, which in turn generates more and more problems in terms of social and economic integration. The academic background for these concerns forms the seminal paper by George Borjas (1987), which applies the classical Roy (1951) model to the migration context. The Roy model offers a rigorous and theoretically powerful framework to analyse the self-selection of individuals. According to the Roy model, self-selection is driven by comparative advantage of individuals. As a consequence, the distribution of income in receiving and sending countries determines whether individuals with higher or lower abilities tend to migrate: if the distribution of income in the host country is more equal than in the home country, and if the incomes of (potential) migrants in both locations are sufficiently positively correlated, migrants are chosen from the lower tail of the income distribution and vice versa (Borjas, 1987, pp. 551-52).

This has important policy consequences: since rich countries have on average a higher equality in the distribution of earnings than poor countries, the Roy model predicts that migrants from poor countries are unfavourably selected with regard to their skills and other abilities relevant for their labour market performance. The negative selection bias of the migrant population in OECD countries may increase over time, since more and more migrants come from poor countries.

In its original formulation the Roy (1951) model does not consider any switching costs, and the Borjas (1987) model assumes that moving costs are proportional to the income level. Pecuniary and non-pecuniary migration costs play however an important role in migration decisions, and it is reasonable to assume that skills and abilities relevant for the labour market performance of individuals affect moving costs. The same human capital characteristics which yield higher returns in the labour market are likely to reduce individual moving costs. At least it is likely that the share of migration costs in income tends to decline with increasing income. However, if we assume that the share of income which has to be spend for migration costs is decreasing in the income level, migrants may be chosen from the upper tail of the income distribution although the distribution of income in the receiving country is more equal than in the sending country.[1]

In the context of international migration, it was difficult to falsify the predictions of the Roy-model since micro data sets which contain individuals in sending and receiving countries were not available. Meanwhile novel data sets exist (Carrington and Detragiache, 1998; Defoort, 2006; Docquier and Marfouk, 2005), which provide macro information on the skill levels of migrants by country of origin. More specifically, these data sets distinguish migrants in OECD countries by skill levels. The overwhelming share of these migrants stems from developing countries. The data on the skill composition of the migrant population can be related to the skill composition of the native population in the sending countries. Although unobservable abilities, which might be relevant for the labour market performance, are not covered, these data sets allow at least to analyse the selection-bias of migrants with regard to their observable education levels.

At first glance, it looks as if migrants tend to be positively self-selected. Table 1 displays for the 6 main receiving countries in the OECD (Australia, Canada, France, Germany, USA, UK) the share of migrants in the labour force of 143 sending countries distinguished by skill levels from 1975 to 2000. The share of migrants is among the skilled workers around three times larger than among the unskilled workers, although the inequality in earnings measured by the Gini-coefficient is larger in the sending countries than in the receiving countries. Moreover, this tendency seems to be rather stable over time.

---

Table 1 about here

---

However, this does not say that the inequality of earnings does not affect the (self-)selection of migrants. An increase in the inequality of earnings

[1] A similar point has been made by Chiswick (2000): He demonstrates in a numerical example, that the strong implications of the Roy model are relaxed if we assume that moving costs are a fixed amount rather than a constant share of income. The point we make here is more general: We only assume that migration costs tend to grow less than proportional with the income level.

in the receiving relative to the sending country may increase the favourable selection bias and vice versa.

In this paper, we first analyse, drawing on a similar approach by Brücker and Trübswetter (2006), in an extended Roy-model how differences in the structure of earnings affect the skill distribution of migrants. The model determines the skill bias of migrants in dependence on the average income level, the inequality of earnings, and migration costs. We find, in contrast to the predictions of the standard Roy model, that the inequality of earnings has an ambiguous impact on the self-selection of migrants.

Second, we examine the correlation between the inequality of earnings, income levels and factors which may affect migration costs and the selection bias of earnings empirically. Our regression results indicate (i) that an increasing inequality of earnings in both, the receiving *and* the sending country increases the favourable selection bias of migrants, (ii) that a higher difference in per capita income levels has a negative impact on the selection bias, and (iii) that higher migration costs and higher migration barriers tend to be positively correlated with the skill bias of the migrant population.

The remainder of the paper is organised as follows: Section 2 outlines our version of a generalised Roy model, which considers the correlation between migration costs and skill levels. Section 3 presents the data set and our empirical results. Section 4 concludes.

# 2 The Model

Suppose that $w_1$ is the wage of residents in the home country (country 1), and $w_2$ the wage of residents in the host country (country 2). Assume that log wages in country 1 and country 2 have a joint normal distribution, such that

$$\ln w_1 = \mu_1 + \varepsilon_1, \tag{1}$$

where $\mu_1$ is the mean of the log wage in country 1 and $\varepsilon_1$ a normally distributed disturbance with zero mean and variance $\sigma_1^2$. Analogously,

$$\ln w_2 = \mu_2 + \varepsilon_2, \tag{2}$$

where $\varepsilon_2$ is normally distributed with zero mean and variance $\sigma_2^2$. The Roy model focuses on the impact of selection bias on the disturbances $\varepsilon_1$ and $\varepsilon_2$, which can be interpreted as the premium for skills and other abilities. Since we cannot distinguish in our data between the impact of individual abilities and education levels, we assume here that a mapping between individual wages and skill levels exists, such that the wage is monotonically increasing in the skill level of individuals in both countries.

The original Roy model ignores all switching costs, i.e. an individual from country 1 migrates into country 2 if $w_2 > w_1$, while Borjas (1987) treats moving costs as a constant share of the income level for all individuals. However, it is reasonable to assume that moving costs exist and that they are related to human capital characteristics and other abilities of individuals. Suppose that $c$ represents the pecuniary and non-pecuniary costs of migration as a proportion of home income. Migration occurs if $\frac{w_2 - w_1}{w_1} > c$, or, approximately, if $\ln w_2 - \ln w_1 > c$. Assume that $c$ is normally distributed with mean $\gamma$ and disturbance $\eta$, i.e.

$$c = \gamma + \eta, \tag{3}$$

and that $\eta \sim N(0, \sigma_\eta^2)$. The decision to migrate is then determined by the sign of the index function, $I^*$, which contains the wage gain from moving minus the costs of migration:

$$I^* = \mu_2 - \mu_1 - \gamma + \varepsilon_2 - \varepsilon_1 - \eta, \tag{4}$$

i.e. an individual migrates if $I^* > 0$, and stays at home if $I^* \leq 0$.

Define

$$\sigma^* = \sqrt{\mathrm{Var}(\varepsilon_2 - \varepsilon_1 - \eta)}, \quad z = -\frac{\mu_2 - \mu_1 - \gamma}{\sigma^*}, \quad \text{and} \quad \epsilon = \frac{\varepsilon_2 - \varepsilon_1 - \eta}{\sigma^*}.$$

Migration occurs if $\epsilon > z$. Under the normality assumptions, the share of migrants in the population, $mst$, is given by

$$mst = Pr(\epsilon > z) = 1 - \Phi(z), \tag{5}$$

where $\Phi()$ is the cumulative distribution function of the standard normal. Using the standard sample selection formula (Heckman, 1976, 1979), the wage of a migrant in the home country can be written as

$$E(\ln w_1 | I^* > 0) = \mu_1 + \sigma_{1\epsilon}\lambda(z), \tag{6}$$

and the wage in the host country as

$$E(\ln w_2 | I^* > 0) = \mu_2 + \sigma_{2\epsilon}\lambda(z), \tag{7}$$

where $\sigma_{1\epsilon}$ and $\sigma_{2\epsilon}$ are the covariance of $\varepsilon_1$ and $\epsilon$, and the covariance of $\varepsilon_2$ and $\epsilon$, respectively, and

$$\lambda(z) = \frac{\phi(z)}{1 - \Phi(z)}$$

is the inverse of Mills' ratio and $\phi()$ the density of the standard normal.

Whether migrants are favourably or unfavourably self-selected depends on the sign of the second term in the equations (6) and (7). Since $\lambda(z) \geq 0$ by definition, the average migrant has higher skills than the average person in the home country if $\sigma_{1\epsilon} > 0$, and, analogously, higher skills than the average person in the host country if $\sigma_{2\epsilon} > 0$ – if we ignore the limiting case that $\lambda(z) = 0$.

An interpretation of these conditions requires that we decompose $\sigma_{1\epsilon}$ and $\sigma_{2\epsilon}$. Using the definition for the covariance, we can rewrite $\sigma_{1\epsilon}$ as

$$\sigma_{1\epsilon} = \frac{\sigma_{12} - \sigma_1^2 - \sigma_{1\eta}}{\sigma^*},$$

and $\sigma_{2\epsilon}$ as

$$\sigma_{2\epsilon} = \frac{\sigma_2^2 - \sigma_{12} - \sigma_{2\eta}}{\sigma^*}.$$

Thus, we can derive two fundamental conditions for the favourable self-selection of migrants: firstly, migrants are better off than the average person in the home population if $\sigma_{12} > \sigma_1^2 + \sigma_{1\eta}$, or if

$$\frac{\sigma_2}{\sigma_1} > \frac{1}{\rho_{12}} + \frac{\rho_{1\eta}}{\rho_{12}}\frac{\sigma_\eta}{\sigma_1}, \tag{8}$$

where $\rho_{12}$ is the correlation coefficient between $\varepsilon_1$ and $\varepsilon_2$, and $\rho_{1\eta}$ the correlation coefficient between $\varepsilon_1$ and $\eta$. We assume for the further analysis that $\rho_{12} > 0$, since a negative correlation would imply that individuals which have a higher income in the source country have a lower in the receiving country and vice versa, which would not make much sense economically. Note that the second term on the right-hand side captures the correlation between income and moving costs. Since we assume that labour-market abilities and the share of moving costs $c$ are negatively correlated, i.e. that $\rho_{1\eta} < 0$, the second term is negative, and, hence, increases the probability of a favourable selection of migrants relative to the average person in the home population for a given variance of earnings in the host and the home country.

Secondly, the migrant is better off than the average person in the host country if $\sigma_2^2 > \sigma_{12} + \sigma_{2\eta}$ , or if

$$\frac{\sigma_2}{\sigma_1} > \rho_{12} + \rho_{2\eta}\frac{\sigma_\eta}{\sigma_1}, \tag{9}$$

where $\rho_{2\eta}$ is the correlation coefficient between $\varepsilon_2$ and $\eta$. Once again, since we assume that $\rho_{2\eta} < 0$, the second term on the right-hand side increases the probability of a favourable selection of migrants relative to the average person in the host population for a given variance of earnings in the host and the home country.

## Comparative Statics

Consider now the implications of the model for a change in the economic conditions underlying the (self-)selection of migrants. We can write the selection bias of migrants relative to the average person in the home population as

$$S_j = S_j\left(\omega, c, \sigma_1, \sigma_2, \eta, \rho_{12}, \rho_{1\eta}, \rho_{2\eta}\right), \quad j \in \{1, 2\}$$

where $\omega \equiv \mu_2 - \mu_1$ is the difference in mean incomes between the host and the home country. The second terms in equations (6) and (7) show that the selection bias in the home country is given by

$$S_1 = \sigma_{1\epsilon}\lambda(z),$$

and in the host country by

$$S_2 = \sigma_{2\epsilon}\lambda(z).$$

We can thus write the impact of a change in any variable $x$ on the change in $S_1$ and $S_2$ as

$$\frac{\partial S_1}{\partial x} = \frac{\partial \sigma_{1\epsilon}}{\partial x}\lambda + \frac{\partial \lambda}{\partial x}\sigma_{1\epsilon}, \tag{10}$$

and as

$$\frac{\partial S_2}{\partial x} = \frac{\partial \sigma_{2\epsilon}}{\partial x}\lambda + \frac{\partial \lambda}{\partial x}\sigma_{2\epsilon}. \tag{11}$$

The first term on the right hand side in equations (10) and (11) captures the composition effect for a constant scale of migration, and the second term the scale effect for a given composition of the migrant population (Borjas, 1987).

We focus here on the selection bias of migrants relative to the average person in the home country. Define $k = \sigma_2\rho_{12} - \sigma_1 - \sigma_\eta\rho_{1\eta}$. $k$ has a positive sign if $\frac{\sigma_2}{\sigma_1} > \frac{1}{\rho_{12}} + \frac{\rho_{1\eta}}{\rho_{12}}\frac{\sigma_\eta}{\sigma_1}$, i.e. if migrants are positively selected, and a negative one, if otherwise.

Consider first the impact of a change in the inequality of earnings on the selection bias. The derivation of $S_1$ with respect to $\sigma_1$ yields

$$\frac{\partial S_1}{\partial \sigma_1} = \frac{2\sigma_1\, k^2 \;-\; (\sigma_1 - k)\,\sigma^{*2}}{\sigma^{*3}}\,\lambda + \;\frac{\sigma_1 k^2}{\sigma^{*3}}\,\frac{\partial \lambda}{\partial z}\,z, \tag{12}$$

where the sign of the first term – the composition effect – is positive if $2\sigma_1 \left(\sigma_1 - \rho_{12}\sigma_2 + \rho_{1\eta}\sigma_\eta\right)^2 > \left(2\sigma_1 - \rho_{12}\sigma_2 + \rho_{1\eta}\sigma_\eta\right) \sigma^{*2}$, which depends on the value of the parameters.[2]

The impact of the second term – the scale effect – depends on the sign of $z$. If the net difference in mean earnings (incl. moving costs) is positive (i.e. $z < 0$), the scale effect is negative, and positive if otherwise. Intuitively, a positive (negative) selection bias tend to disappear since the share of migrants from the upper (lower) end of the skill distribution tend to decline with an increasing number of migrants.

The effect of an increasing inequality of earnings in the host country is again ambiguous. Analogously to equation (12), a derivation of $S_2$ with respect to $\sigma_2$ gives

$$\frac{\partial S_1}{\partial \sigma_2} = \frac{\sigma_1 \left(\rho_{12}\, \sigma^{*2} \;-\; k\, n\right)}{\sigma^{*3}}\, \lambda \;-\; k\, n\, \frac{\sigma_1}{\sigma^{*3}}\, \frac{\partial \lambda}{\partial z}\, z, \tag{13}$$

where $n$ is defined as $\sigma_2 + \rho_{12}\sigma_1 - \rho_{23}\sigma_\eta > 0$. The composition effect has a positive sign if $\rho_{12}\sigma^{*2} > kn$, which is always the case if a negative selection bias of the migrant population exists. In the converse case the sign of the composition effect depends on the sign of the individual parameters. The scale effect is positive if migrants are positively selected *and* the net difference in earnings between the host and the home country is positive (i.e. $z < 0$), and negative in the converse case. Thus, an increase in the inequality of earnings in the host country strengthens a given selection bias in both directions via the scale effect if the difference in net earnings is positive, while it reduces a negative selection bias via the composition effect, and is ambiguous if a positive selection bias exists.

Lets consider now a change in the difference of earnings between the host and the home country. Using equation (10) it can be shown that

$$\frac{\partial S_1}{\partial \omega} = -\, \frac{\sigma_1}{\sigma^{*2}}\, \frac{\partial \lambda}{\partial z}\; k, \tag{14}$$

i.e. a change in the income differential affects the composition of migrants only via the scale effect. An increase in the difference of earnings between the host and the home country reduces the positive (negative) selection bias of the migrant population if they are positively (negatively) selected. The

[2]Note that the sign of the composition effect remains ambiguous if we neglect the correlation between labour market abilities and moving costs: The derivation of $\sigma_{1\epsilon}$ with respect to $\sigma_1$ yields in this case, analogously to the first term in equation (12), $\frac{\partial \sigma_{1\epsilon}}{\partial \sigma_1} = \frac{2\sigma_1(\sigma_1 - \rho_{12}\sigma_2)^2 + (\sigma_1 - \rho_{12}\sigma_2)(\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2)^2}{(\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2)^3}$. It can be shown that this is positive if $\rho_{12}(\sigma_1^2 + \sigma_2)^2 > 2\sigma_1\sigma_2$.

intuition behind this result is that a higher difference in earnings increases the share of migrants in the population, which in turn reduces the selection bias in both directions, since migrants are increasingly drawn from the mean parts of the income distribution.

Increasing the mean costs of migration has the opposite effect, i.e.

$$\frac{\partial S_1}{\partial c} = \frac{\sigma_1}{\sigma^{*2}} \frac{\partial \lambda}{\partial z} k, \tag{15}$$

since increasing moving costs reduces the share of migrants in the population, which in turn increases the selection bias of the migrant population.

Finally, we can assess the implications of a change in the correlation coefficients. The derivation of the change in $S_1$ with respect to a change in the correlation coefficients are given by

$$\frac{\partial S_1}{\partial \rho_{12}} = k \frac{\sigma_1^2 \sigma_2}{\sigma^*} \lambda - k \frac{\sigma_1^2 \sigma_2}{\sigma^{*3}} \frac{\partial \lambda}{\partial z} z, \tag{16}$$

$$\frac{\partial S_1}{\partial \rho_{1\eta}} = k \frac{\sigma_1^2 \sigma_\eta}{\sigma^*} \lambda - k \frac{\sigma_1^2 \sigma_\eta}{\sigma^{*3}} \frac{\partial \lambda}{\partial z} z, \tag{17}$$

and

$$\frac{\partial S_1}{\partial \rho_{2\eta}} = k \frac{\sigma_1 \sigma_2 \sigma_\eta}{\sigma^*} \lambda - k \frac{\sigma_1 \sigma_2 \sigma_\eta}{\sigma^{*3}} \frac{\partial \lambda}{\partial z} z. \tag{18}$$

In all three equations the composition effect and the scale effect have the same sign if the net difference in mean earnings (incl. moving costs) between the host and the home is positive ($z < 0$), and the converse sign if the net difference in mean earnings is negative ($z > 0$). Thus, an increasing (positive) correlation between earnings in the home and the host country strengthens the selection bias both via the composition effect and the scale effect if the net difference in earnings is positive. In contrast, an increasing (negative) correlation between labour market abilities and moving costs weakens the selection bias if the net difference in earnings is positive.

# 3 Empirical evidence

The results from our analysis of the generalised Roy model which considers migration costs can be summarised as follows: (i) a higher variance of earnings in the home country does not necessarily affect the selection bias of the migrant population unfavourably, (ii) a higher variance of earnings in the host country does not necessarily affect the selection bias of the migrant population favourably, (iii)increasing the difference in average earnings between the host and the home country weakens a given selection bias of the migrant population, (iv) increasing the mean migration costs raises a given selection bias of the migrant population, and (v) a higher negative correlation between migration costs and earnings affects the selection bias of the migrant population favourably.

Thus, in contrast of the original Roy model which ignores switching costs, we find an ambiguous impact of the inequality of earnings on the selection bias of the migrant population. In this section we analyse how the distribution of earnings in the receiving countries relative to the sending countries affects and some institutional variables affect the skill composition of migrants. The empirical analysis is based on a novel set of macro migration data, which allows to distinguish migration stocks by their educational attainment. This data is used to calculate the share of skilled and unskilled migrants in the home population. We can thus examine whether the (i) inequality of earnings in the host country, (ii) the inequality of earnings in the home country, (iii) the difference in mean earnings, and (iv) various approximations for migration costs affect the self-selection of migrant with respect to their skill levels.

## Specification of the estimation equation

Specifically, we estimate

$$\ln(s^h_{jkt}/s^l_{jkt}) = a_0 + a_1\theta_{jt} + a_2\theta_{kt} + \gamma' X_{jkt} + \eta' Y_{jt} + \lambda' Z_{kt} + \epsilon_{jkt}, \tag{19}$$

where $s^h$ and $s^l$ are the shares of skilled and unskilled migrants, respectively, residing in receiving country $j$ as a share of the skilled labour force and the unskilled labour force, respectively, in sending country $k$, $\theta$ is a measure for the inequality in earnings, $X$, $Y$ and $Z$ are sets of variables which may affect the benefits and costs of migration in the host country, home country and both, $\gamma$, $\eta$ and $\lambda$ are the associated vectors of coefficients, and $\epsilon$ is the error term. The index $j = 1, 2...6$ denotes the receiving country, the index $k = 1, 2...143$ the sending country, and the index $t = 1, 2...6$ the time period.

Following Hsiao (1986), we specify the error term as a two-way error components model, i.e. as

$$\epsilon_{jkt} = \mu_{jk} + \nu_t + \varepsilon_{jkt}, \tag{20}$$

where $\mu_{jk}$ is a bilateral fixed effect, $\nu_t$ a time-specific fixed effect and $\varepsilon_{jkt}$ white noise. We also estimate a pooled version of the model without fixed effects.

As a measure for the inequality of earnings we use the Gini-coefficients in the respective countries. Other inequality measures such as the variance of earnings might be more appropriate for our purposes, but we rely on the Gini-coefficient since this is the only measure which is available for a broad set of countries.

In the most parsimonious specification of the model we consider only the Gini-coefficients as explanatory variables. Step by step we extend the model by other variables which may affect the skill distribution of migrants. First, we use the log of the per capita income differential between the receiving and the sending country and the log GDP in the sending country as additional explanatory variables. According to our theoretical model, the income differential should weaken a given selection bias via the scale effect. Since the descriptive statistics indicate that the migrants from most of our countries have a favourable skill-bias, we shall expect a negative or insignificant sign for this variable.

The per capita income in the sending country may affect liquidity constraints, and, hence, the composition of the migrant population (Faini and Venturini, 1995). Since the relaxation of liquidity constraints increases the opportunities of unskilled individuals to migrate more than proportional, we expect again a negative sign for this variable.

Second, we consider variables which may affect the costs of migration and migration barriers. The first among these variables is geographical distance, which is considered as an approximation for transport and communication costs. In general it can be expected that unskilled migrants benefit more from an overall reduction of migration costs than skilled migrants, such that we expect a negative coefficient for this variable. As institutional variables which approximate migration barriers we consider colonial links between the sending and the receiving country, bilateral migration agreements and the free movement of workers within the EU. All three variables tend to reduce migration barriers, and, hence, increase the number of migrants. We expect that these variables affect the skill composition of migrants negatively for three reasons: First, analogously to transport and communication costs, reduced migration barriers imply lower costs for migration, which increases the incentives and opportunities of unskilled individuals more than proportional. Second, migration barriers have in many cases a skill-bias, i.e. many countries which restrict migration tend to out-select unskilled individuals. Hence, a reduction of migration barriers increases the migration chances of unskilled individuals more than proportional. Third, reducing migration barriers affects the composition of the migrant population also via the scale affect. If migrants are favourable self-selected on average, than this favourable selection bias disappears with an increasing number of migrants. Altogether, we expect a negative coefficient for the distance variable and all three variables, which approximate a reduction of institutional migration barriers.

Finally, we consider a democracy index as a variable which should capture the political 'push' factors in the sending countries. This variable may affect migration incentives differently for skilled and unskilled individuals and, hence, affect the skill composition of the migrant population. However, it is hard to predict ex ante in which direction the selection bias is affected by this variable. It depends on whether high skilled individuals are more or less affected by push factors such as political instability or a lack of political freedom. The expected signs of the parameters of the model are summarised in Table 2.

---

Table 2 about here

---

The model is estimated in most regressions in static form. This seems to us a reasonably approach, since we have in our data set only one observation every five years and the main explanatory variables such as the Gini-coefficients and per capita income levels have a high stability over time. However, we estimated the model also in dynamic form. In order to address the so-called Nickell-bias we used the GMM-system estimator developed by Arellano and Bover (1995).

## Data

The data set which is employed here has been collected by Defoort (2006), and builds on previous contributions by Docquier and Marfouk (2005). The data set uses data from OECD countries on the skill levels of the migrants population for residents from (almost) all countries of the world. In addition to previous contributions, this data set expands the time-dimension and provides data for 6 receiving countries (Australia, Canada, France, Germany, UK, USA) for the years 1975 to 2000 (one observation each 5 years). This allows to carry out a panel analysis. For a detailed description of the data set see the Annex A.

For balanced panel estimation, we consider in this paper only bilateral pairs where we have data for all 6 time periods, which gives for the 143 sending and 6 receiving countries 705 bilateral relations and a total of 4,230 balanced panel observations.

We distinguish two types of individuals: skilled and unskilled. An individual is defined as skilled if it has an educational attainment of 9 years of schooling or more, and as unskilled if it has 8 years or less. We assume that this distinction corresponds to the distinction between individuals which have completed secondary education and those which have not in our sample. Note that around 70 per cent of the population in the sending countries belong to the non-skilled group. For the skill level of the migrant population data on the educational structure by country of birth for the working-age population has been used from OECD sources (see Annex A).

The most important explanatory variable is the measure for the inequality of earnings. We employ the Gini-coefficients provided by the World Development Reports from the World Bank. The missing observations have been estimated using a model which regresses the Gini-coefficient against the per capita GDP level and a country specific fixed effect. The explanatory power of this simple model has been relatively high, such that the estimated Gini-coefficient can be used as a relatively good proxy for the missing observations.

For the income variable, we use the GDP per capita in purchasing power parity at constant prices from the World Development Indicators of the World Bank. The distance variable captures the distance between the capitals of the sending countries and the destination country in km. Moreover, we consider the following institutional variables which affect migration restrictions: a free movement dummy, which has a value of 1 if there is free movement between the sending and the receiving country,[3] an agreement dummy, if there is a bilateral guestworker agreement or another bilateral agreement which enables migration between the sending the receiving country at least partially, and a colonial link dummy, if the sending country is a former colony of the receiving country. Finally, we consider a variable which captures the political freedom in the sending country: the democracy index from the World Development Indicators of the World Bank. This index is scaled between 0 and 1.

The descriptive statics of the variables are displayed in annex table A1.

## Estimation results

We estimate two versions of the model in equation (19): First, a pooled model without fixed effects, and, second, a model with country-specific fixed effects. Although our tests suggest a fixed-effects specification (see below), we estimate in the first step a pooled version of the panel model. Note that the coefficients are identified in the fixed-effects model by the within-dimension, which may generate an identification problem if the variance of our variables is small over time. Since the Gini-coefficients and income variables are relatively stable, this might be an issue in our sample.

[3]Note that free movement is only relevant for the receiving countries in the EU

The results of the pooled regressions are displayed in Table 3. The sign of the coefficients for the most important variables, the Gini-coefficients of the host and the home countries, are *both* positive and highly significant. However, the size of the estimated parameter for the Gini-cofficient in the receiving countries is much larger than that for the sending countries.

Thus, an increasing inequality of earnings in both, the host and the home country, has a positive impact on the selection bias of the migrant population. While the first result confirms the predictions of the standard version of the Roy model, the second results clearly contradicts it. However, this result is consistent with the expectations of our extended Roy model, which allows for a negative correlation between skill levels and individual moving costs.

---

Table 3 about here

---

The results for the income variables are also in line with our theoretical expectations. We obtain a negative sign for the income differential between the home and the host country as well as for the income level in the home country. Thus, increasing the income differential between countries reduces the favourable selection bias. Moreover, the negative sign of the coefficient for the income level in the sending countries indicates that liquidity constraints matter, i.e. that more and more unskilled migrants can afford to migrate if the per capita income level tends to increase in the source country.

Finally, we receive a positive sign for the distance variable and negative ones for the bilateral guestworker agreement and colonial tie dummy variables, indicating that higher transport and communication costs increase the positive selection bias, while relaxing migration barriers reduce it. We obtain only an unexpected sign for the free movement dummy, which suggests that removing migration barriers within the EU is associated with a higher share of skilled migrants. Moreover, the democracy-index has a negative sign, indicating that a higher level of political freedom involves a lower proportion of skilled migration.

Our specification tests suggest that the fixed effects are both, individually and jointly, significant. Moreover, the Hausman-test clearly rejects a random-effects specification. We estimated the fixed effects model with a feasible GLS estimator, which allows for group-specific heteroscedasticity. Our LR-Test clearly suggests that heteroscedasticity is present in our data set (see the regression diagnostics in Table 4).[4]

[4]However, the LR-Test does not reject the hypothesis of no contemporary correlation.

Table 4 about here

In the static specification of the fixed-effects model, the sign of the Gini-coefficient in the receiving country remains positive and highly significant. However, the sign of the Gini-coefficient in the sending country has also a positive sign, but appears insignificant. This might be caused by the relatively low within-variance in our data. The coefficients for the income differential have again a negative sign but are insignificant as well. Only home income has a significant negative sign. The free movement dummy and the guestworker agreement variable have a positive sign.

The dynamic version of the model is estimated with the GMM-System estimator in order to control for the simultaneous equation bias (Nickell, 1981). Note that GMM-estimation is appropriate in our data set with a small time dimension ($T = 5$) and large group dimension ($N = 705$) (Judson and Owen, 1999).

The estimation results of the dynamic model are largely in line with those from the pooled OLS regressions: Both the Gini-coefficients in the host and the home country have positive signs and appear highly significant, while the income variables have negative signs and are significant as well. Moreover, both the guestworker agreement and the free movement dummies have negative signs and appear as significant.

Given the rather short-time dimension of our panel and possible adjustment processes, the GMM-estimates of the dynamic model might be more reliable than the fixed effects estimates of the static model. Thus, we conclude that a higher inequality in earnings in the host countries can increase the favourable selection bias, while the same holds true for the sending countries. Moreover, increasing the costs of migration affects the selection bias positively, while relaxing the barriers to migration has the opposite effect.

# Conclusion

In this paper, we have first analysed on basis of a generalised Roy model the forces which may affect the selection bias of the migration population. In contrast to the traditional Roy model, which ignores moving costs or treats them as a uniform share of the income, we find that the inequality of earnings has an ambiguous impact on the self-selection of the migrant population if we consider that the share of income which has to be spend for migration is declining in the income level. This has important consequences for the sending countries: Increasing the inequality of earnings in source countries may yield an even higher brain drain for countries which suffer already from emigration of the skilled population.

Given the ambiguity of our theoretical findings, it is important to examine the empirical evidence. Based on our novel panel data set we find first descriptive evidence that migrants tend to be favourably self-selected with regard to their skill level on average, although the earnings inequality is much higher in the sending countries relative to the recipients. This descriptive evidence seems to contradict the standard predictions of the Roy-model.

However, our regression analysis which allow to control for certain factors provides a number of additional insights: We find evidence that (i) a higher inequality of earnings in the receiving countries, and (ii) a higher inequality of earnings in the sending countries is both positively correlated with a favourable selection bias. The latter result seems, against the background of our theoretical considerations, to indicate that moving costs have a non-trivial impact on the self-selection of migrants. This is also supported by the fact that higher migration costs tend to be positively correlated with a favourable selection bias, while lower migration barriers affect the self-selection unfavourably.

Finally, the favourable selection bias tends to decline with an increasing income difference between the home and the host country and a higher income level in the home country. The rationale behind the first result is that a higher income difference yields a higher share of migrants in the population, and, hence, dilutes the positive selection bias. The latter result supports the hypothesis that liquidity constraints play an important role in migration: While in poor countries only the rich and better educated parts of the population can afford to migrate, the share of the unskilled is increasing in the income level of the source countries, i.e. when liquidity constraints are relaxed.

Thus, altogether, our theoretical and empirical findings seem to indicate that a high inequality of earnings in the source countries does not prevent a positive selection bias of the migrant population. While this is reassuring news for the receiving countries, this need not to be the case for the sending countries: They may suffer from a 'brain drain.' However, as the 'new economics of the brain drain' literature suggests, a positive selection of the migrant population might be associated with additional human capital investment, and need therefore not necessarily be a burden for the sending countries. This, however, is a question for further research.

# References


Adams, R. (2003): "International migration, remittances and the brain drain: a study of 24 labor-exporting countries," World Bank Policy Research Working Paper No. 2972.

Arellano, M., and O. Bover (1995): "Another Look at the Instrumental Variable Estimation of Error Component Models," *Journal of Econometrics*, 68, 29–51.

Barro, R. J., and J. Lee (2000): "International Data on Educational Attainment: Updates and Implications," CID Working Papers 42, Center for International Development (Harvard University).

Beine, M., F. Docquier, and H. Rapoport (2001): "Brain drain and economic growth: theory and evidence," *Journal of Development Economics*, 64(1), 275–289.

Bhagwati, J. N., and K. Hamada (1974): "The Brain Drain, International Integration of Markets for Professional and Unemployment: A Theoretical Analysis," *Journal of Development Economics*, 1(1), 19–24.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 77(4), 531–553.

Brücker, H., and P. Trübswetter (2006): "Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany," *Empirica*, 33, (forthcoming).

Carrington, W., and E. Detragiache (1998): "How big is the brain drain?," IMF Working paper WP/98/102.

Chiswick, B. R. (2000): "Are Immigrants Favorably Self-Selected? An Economic Analysis," IZA Discussion Paper No. 131.

Defoort, C. (2006): "Tendances de long terme des migrations internationales : Analyse  partir des 6 principaux pays receveurs," Mimeo, University of Lille.

Docquier, F., and A. Marfouk (2004): "Measuring the international mobility of skilled workers - Release 1.0," Policy Research Working Paper n. 3382, World Bank, August.

——— (2005): "International migration by educational attainment, 1990-2000 (Release 1.1)," in *International Migration, Remittances and Development*, ed. by C. Ozden, and M. Schiff. McMillan and Palgrave: N.Y.

Faini, R., and A. Venturini (1995): "Migration and Growth: The Experience of Southern Europe," CEPR Discussion Paper No. 964.

Grubel, H., and A. Scott (1966): "The international flow of human capital," *American Economic Review*, 56, 268–74.

Heckman, J. J. (1976): "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Model," *Annals of Economic and Social Measurement*, 15, 475–492.

——— (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153 – 161.

Hsiao, C. (1986): *Analysis of Panel Data*. Cambridge University Press, Cambridge MA.

Judson, R., and A. Owen (1999): "Estimating dynamic panel data models: a guide for macroeconomists," *Economic Letters*, 65, 9–15.

Kwok, W., and H. Leland (1982): "An economic model of the brain drain," *American Economic Review*, 72(1), 91–100.

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.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135 – 146.

Stark, O., C. Helmenstein, and A. Prskawetz (1997): "A brain gain with a brain drain," *Economics Letters*, 55, 227–34.

# A Description of the data set

For the skill levels of the residents in the sending countries, we use population data from the United Nations and education data from Barro and Lee (2000). We consider only individuals aged 25 or more. For countries where the Barro and Lee measures are missing (about 70 countries in 2000), we have transpose the skill share of the neighbouring country with the closest human development index regarding education. This method gives good approximations of the brain drain rate, broadly consistent with anecdotal evidence.

Regarding migrants, there has been no systematic empirical assessment of the educational structure of international migration until recently. Despite numerous case studies and anecdotal evidence, many institutions consider the lack of harmonized international data on migration by country of origin and education level as the major problem for monitoring the scope and impact of brain drain in developing areas. An exception can be found in Carrington and Detragiache (1998) who provided estimates of the emigration stocks and rates of tertiary educated workers for 61 developing countries in 1990. These estimates are based on three main statistical sources (US Census data on the skill structure of immigration, OECD data on immigration per country of origin, Barro and Lee (2000) data describing the skill structure in sending countries). Unfortunately, these estimates rely on two very strong assumptions: First, for non-US countries, they use OECD migration statistics which report limited information on the origin of immigrants. Second, they transpose the skill structure of US immigrants on the total immigration stock in the OECD. Adams (2003) used the same methodology to update the emigration rates of 24 labor-exporting countries in 2000. Docquier and Marfouk (2004) and Docquier and Marfouk (2005) revisit the methodology by collecting data on the immigration structure by educational attainment and country of birth from all OECD receiving countries. They use harmonised definitions of educational attainment and distinguish the working-age migration stock by the country of birth in 1990 and 2000. Thus, the time dimension of this data set is two small for a panel analysis.

The data set which is employed in this paper has been collected by Defoort (2006). This data set extends the time dimension of Docquier and Marfouk's data set but focuses on a limited set of receiving countries. It considers the six major immigration countries in the OECD (USA, Canada, Australia, UK, Germany, France), which represents about 75 percent of the OECD stock of working-aged migrants. For these countries, they rely on Census data available in 1980, 1990 and 2000 (as well as in 1985 and 1995 in the case of Australia). These Census data give the structure of immigration by country of birth (country of citizenship in the case of Germany) and by educational attainment. Individuals which have less than 9 years of schooling are defined as unskilled, and individuals which have 9 years of schooling or more are considered as skilled.

# Tables

**Table 1**: Mean share of skilled and unskilled migrants in % of home labour force and Gini-coefficients, 1975-2000

| variable | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 |
|---|---|---|---|---|---|---|
| mean share of skilled mig. | 2.43 | 2.18 | 2.23 | 2.22 | 2.22 | 2.50 |
| mean share of unskilled mig. | 0.71 | 0.70 | 0.74 | 0.75 | 0.86 | 0.69 |
| ratio skilled/unskilled mig. sh. | 3.42 | 3.11 | 3.01 | 2.96 | 2.58 | 3.62 |
| Gini host country | 31.14 | 29.36 | 30.68 | 31.53 | 31.65 | 31.78 |
| Gini home country | 41.94 | 41.53 | 41.40 | 42.24 | 43.09 | 43.07 |
| ratio Gini host/ Gini home c. | 0.74 | 0.71 | 0.74 | 0.75 | 0.73 | 0.74 |

*Sources*: Defoort (2006) and own calculations.

**Table 2**: Expected signs of coefficients

| description | variable | expected sign |
|---|---|---|
| Gini coefficient host country | $\theta_{jt}$ | + - |
| Gini coefficient home country | $\theta_{kt}$ | + - |
| GDP per capita differential | $\ln(y_{jt}/y_{kt})$ | - |
| GDP per capita home | $\ln(y_{kt})$ | - |
| distance | $\ln(dist_{jk})$ | + |
| colonial tie dummy | $COLON_{jkt}$ | - |
| bilateral guestworker agreement dummy | $AGREE_{jkt}$ | - |
| free movement dummy | $FREE_{jkt}$ | - |
| democracy index | $DEM_{kt}$ | ? |

**Table 3**: Pooled estimation results

| variable | (1) | (2) | (3) |
|---|---|---|---|
| $\theta_{jt}$ | 0.13*** | 0.14*** | 0.12*** |
| | (24.56) | (31.13) | (26.89) |
| $\theta_{kt}$ | 0.04*** | 0.02*** | 0.02*** |
| | (23.15) | (11.34) | (11.35) |
| $\ln(y_{jt}/y_{kt})$ | - | -0.47*** | -0.44*** |
| | - | (-6.13) | (-5.70) |
| $\ln(y_{kt})$ | - | -0.93*** | -0.87*** |
| | - | (-12.12) | (-11.13) |
| $\ln(dist_{jk})$ | - | - | 0.00*** |
| | - | - | (11.76) |
| $COLON_{kt}$ | - | - | -0.77*** |
| | - | - | (-8.95) |
| $AGREE_{jkt}$ | - | - | -0.07 |
| | - | - | (-1.14) |
| $FREE_{jkt}$ | - | - | 0.20** |
| | - | - | (1.98) |
| $DEMOCRACY_{kt}$ | - | - | -0.14** |
| | - | - | (-2.37) |
| $CONSTANT$ | -4.53*** | 4.38*** | 4.21*** |
| | (-23.74) | (5.76) | (5.49) |
| Dependent variable | $\ln(s^h_{jkt}/s^l_{jkt})$ | $\ln(s^h_{jkt}/s^l_{jkt})$ | $\ln(s^h_{jkt}/s^l_{jkt})$ |
| Observations | 4230 | 4230 | 4230 |
| Adjusted $R^2$ | 0.21 | 0.42 | 0.46 |
| RMSE | 1.38 | 1.18 | 1.13 |

*Notes*: ***,**,* indicate significance levels of 1%, 5%, and 10%, respectively.

**Table 4**: Fixed effects and GMM estimation results

| variable | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $\ln(s^h_{jk}/s^l_{jk})_{t-1}$ | - | - | - | 0.67*** |
| | - | - | - | (26.34) |
| $\theta_{jt}$ | 0.01*** | 0.01*** | 0.01*** | 0.04*** |
| | (9.28) | (8.55) | (8.88) | (9.97) |
| $\theta_{kt}$ | 0.00 | 0.00 | 0.00 | 0.00** |
| | (0.34) | (1.28) | (1.50) | (2.19) |
| $\ln(y_{jt}/y_{kt})$ | - | 0.00 | -0.01 | -0.50*** |
| | - | (0.35) | (-0.52) | (-5.50) |
| $\ln(y_{kt})$ | - | -0.10*** | -0.13** | -0.51*** |
| | - | (-4.48) | (-5.46) | (-5.28) |
| $AGREE_{jkt}$ | - | - | 0.13*** | -0.45*** |
| | - | - | (4.03) | (-3.99) |
| $FREE_{jkt}$ | - | - | 0.18*** | -0.40*** |
| | - | - | (5.19) | (-3.01) |
| $DEMOCRACY_{kt}$ | - | - | -0.02 | 0.14 |
| | - | - | (-0.87) | (1.41) |
| $CONSTANT$ | 0.00 | 0.00 | 0.00 | 5.78*** |
| | (0.00) | (0.00) | (0.00) | (5.93) |
| Dependent variable | $\ln(s^h_{jkt}/s^l_{jkt})$ | $\ln(s^h_{jkt}/s^l_{jkt})$ | $\ln(s^h_{jkt}/s^l_{jkt})$ | $\ln(s^h_{jkt}/s^l_{jkt})$ |
| Observations | 4230 | 4230 | 4230 | 3525 |
| Adjusted $R^2$ | 0.15 | 0.37 | 0.26 | - |
| F(704,3524)-test statistic | 31.95*** | 22.06*** | 20.90*** | - |
| Hausman $\chi^2$-test statistic | -2105.48?? | 1903.06*** | 2469.61*** | - |
| LR-test statistic for the hetero- vs. homoscedastic model | 2861.36*** | 2871.58*** | 2864.09*** | - |
| LR-test statistic for the auto- vs. the uncorrelated model | 141.02 | 134.27 | 129.20 | - |
| Hansen's $J$-test statistic for overidentifying restrictions | - | - | - | 492.02*** |
| Arellano-Bond test statistic for first-order autocorrelat. | - | - | - | -7.86*** |
| Arellano-Bond test statistic for second-order autocorrel. | - | - | - | 3.63*** |

*Notes*: ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively.

**Table A1**: Descriptive statistics

| variable | obs. | mean | std. dev. | min. | max. |
|---|---|---|---|---|---|
| $\ln(s^h_{jt}/s^l_{kt})$ | 4230 | 1.59 | 1.55 | -3.28 | 7.75 |
| $\theta_{jt}$ | 4230 | 31.10 | 3.92 | 23.70 | 39.40 |
| $\theta_{kt}$ | 4230 | 42.93 | 11.07 | 17.80 | 77.60 |
| $\ln(y_{jt}/y_{kt})$ | 4230 | 2.35 | 1.63 | -1.06 | 5.87 |
| $\ln(y_{kt})$ | 4230 | 7.59 | 1.63 | 4.31 | 10.93 |
| $\ln(dist_{jk})$ | 4230 | 8.74 | 0.79 | 5.56 | 9.86 |
| $COLONIAL_{kt}$ | 4230 | 0.09 | 0.28 | 0 | 1 |
| $AGREEMENT_{jkt}$ | 4230 | 0.22 | 0.41 | 0 | 1 |
| $FREE_{jkt}$ | 4230 | 0.05 | 0.21 | 0 | 1 |
| $DEMOCRACY_{kt}$ | 4230 | 0.47 | 0.36 | 0 | 1 |

## Recently published

| No. | Author(s) | Title | Date |
|---|---|---|---|
| 1/2004 | Bauer, T. K.<br>Bender, S.<br>Bonin, H. | Dismissal protection and worker flows in small establishments<br>published in: Economica, (2007) | 7/04 |
| 2/2004 | Achatz, J.<br>Gartner, H.<br>Glück, T. | Bonus oder Bias? : Mechanismen geschlechtsspezifischer Entlohnung<br>published in: Kölner Zeitschrift für Soziologie und Sozialpsychologie 57 (2005), S. 466-493 (revised) | 7/04 |
| 3/2004 | Andrews, M.<br>Schank, T.<br>Upward, R. | Practical estimation methods for linked employer-employee data | 8/04 |
| 4/2004 | Brixy, U.<br>Kohaut, S.<br>Schnabel, C. | Do newly founded firms pay lower wages? : first evidence from Germany<br>published in: Small Business Economics, (2007) | 9/04 |
| 5/2004 | Kölling, A.<br>Rässler, S. | Editing and multiply imputing German establishment panel data to estimate stochastic production frontier models<br>published in: Zeitschrift für ArbeitsmarktForschung 37 (2004), S. 306-318 | 10/04 |
| 6/2004 | Stephan, G.<br>Gerlach, K. | Collective contracts, wages and wage dispersion in a multi-level model<br>published as: Wage settlements and wage setting : results from a multi-level model. In: Applied Economics, Vol. 37, No. 20 (2005), S. 2297-2306 | 10/04 |
| 7/2004 | Gartner, H.<br>Stephan, G. | How collective contracts and works councils reduce the gender wage gap | 12/04 |
| 1/2005 | Blien, U.<br>Suedekum, J. | Local economic structure and industry development in Germany, 1993-2001<br>published in: Economics Bulletin, Vol. 15, No. 17 (2005), S. 1-8 | 1/05 |
| 2/2005 | Brixy, U.<br>Kohaut, S.<br>Schnabel, C. | How fast do newly founded firms mature? : empirical analyses on job quality in start-ups<br>published in: Michael Fritsch, Jürgen Schmude (Ed.): Entrepreneurship in the region, New York et al., 2006, S. 95-112 | 1/05 |
| 3/2005 | Lechner, M.<br>Miquel, R.<br>Wunsch, C. | Long-run effects of public sector sponsored training in West Germany | 1/05 |
| 4/2005 | Hinz, T.<br>Gartner, H. | Lohnunterschiede zwischen Frauen und Männern in Branchen, Berufen und Betrieben<br>published in: Zeitschrift für Soziologie 34 (2005), S. 22-39, as: Geschlechtsspezifische Lohnunterschiede in Branchen, Berufen und Betrieben | 2/05 |
| 5/2005 | Gartner, H.<br>Rässler, S. | Analyzing the changing gender wage gap based on multiply imputed right censored wages | 2/05 |
| 6/2005 | Alda, H.<br>Bender, S.<br>Gartner, H. | The linked employer-employee dataset of the IAB (LIAB)<br>published as: The linked employer-employee dataset created from the IAB establishment panel and the process-produced data of the IAB (LIAB). In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften 125 (2005), S. 327-336 (shortened) | 3/05 |
| 7/2005 | Haas, A.<br>Rothe, T. | Labour market dynamics from a regional perspective : the multi-account system | 4/05 |

| | | | |
|---|---|---|---|
| 8/2005 | Caliendo, M.<br>Hujer, R.<br>Thomsen, S. L. | Identifying effect heterogeneity to improve the efficiency of job creation schemes in Germany | 4/05 |
| 9/2005 | Gerlach, K.<br>Stephan, G. | Wage distributions by wage-setting regime<br>published as: Bargaining regimes and wage dispersion. In: Jahrbücher für Nationalökonomie und Statistik, Bd. 226, H. 6 (2006) | 4/05 |
| 10/2005 | Gerlach, K.<br>Stephan, G. | Individual tenure and collective contracts | 4/05 |
| 11/2005 | Blien, U.<br>Hirschenauer, F. | Formula allocation : the regional allocation of budgetary funds for measures of active labour market policy in Germany<br>published in: Economics Bulletin, Vol. 18, no. 7 (2006) | 4/05 |
| 12/2005 | Alda, H.<br>Allaart, P.<br>Bellmann, L. | Churning and institutions : Dutch and German establishments compared with micro-level data | 5/05 |
| 13/2005 | Caliendo, M.<br>Hujer, R.<br>Thomsen, S. L. | Individual employment effects of job creation schemes in Germany with respect to sectoral heterogeneity | 5/05 |
| 14/2005 | Lechner, M.<br>Miquel, R.<br>Wunsch, C. | The curse and blessing of training the unemployed in a changing economy : the case of East Germany after unification | 6/05 |
| 15/2005 | Jensen, U.<br>Rässler, S. | Where have all the data gone? : stochastic production frontiers with multiply imputed German establishment data<br>published in: Zeitschrift für ArbeitsmarktForschung, Jg. 39, H. 2, 2006, S. 277-295 | 7/05 |
| 16/2005 | Schnabel, C.<br>Zagelmeyer, S.<br>Kohaut, S. | Collective bargaining structure and its determinants : an empirical analysis with British and German establishment data<br>published in: European Journal of Industrial Relations, Vol. 12, No. 2. S. 165-188 | 8/05 |
| 17/2005 | Koch, S.<br>Stephan, G.<br>Walwei, U. | Workfare: Möglichkeiten und Grenzen<br>published in: Zeitschrift für ArbeitsmarktForschung 38 (2005), S. 419-440 | 8/05 |
| 18/2005 | Alda, H.<br>Bellmann, L.<br>Gartner, H. | Wage structure and labour mobility in the West German private sector 1993-2000 | 8/05 |
| 19/2005 | Eichhorst, W.<br>Konle-Seidl, R. | The interaction of labor market regulation and labor market policies in welfare state reform | 9/05 |
| 20/2005 | Gerlach, K.<br>Stephan, G. | Tarifverträge und betriebliche Entlohnungsstrukturen<br>published in: C. Clemens, M. Heinemann & S. Soretz (Hg.): Auf allen Märkten zu Hause, Marburg 2006, S. 123-143 | 11/05 |
| 21/2005 | Fitzenberger, B.<br>Speckesser, S. | Employment effects of the provision of specific professional skills and techniques in Germany<br>published in: Empirical Economics, Vol. 32, No. 2/3 (2007), S. 529-573 | 11/05 |
| 22/2005 | Ludsteck, J.<br>Jacobebbinghaus, P. | Strike activity and centralisation in wage setting | 12/05 |
| 1/2006 | Gerlach, K.<br>Levine, D.<br>Stephan, G.<br>Struck, O. | The acceptability of layoffs and pay cuts : comparing North America with Germany | 1/06 |
| 2/2006 | Ludsteck, J. | Employment effects of centralization in wage setting in a median voter model | 2/06 |

| | | | |
|---|---|---|---|
| 3/2006 | Gaggermeier, C. | Pension and children : Pareto improvement with heterogeneous preferences | 2/06 |
| 4/2006 | Binder, J.<br>Schwengler, B. | Korrekturverfahren zur Berechnung der Einkommen über der Beitragsbemessungsgrenze | 3/06 |
| 5/2006 | Brixy, U.<br>Grotz, R. | Regional patterns and determinants of new firm formation and survival in western Germany | 4/06 |
| 6/2006 | Blien, U.<br>Sanner, H. | Structural change and regional employment dynamics | 4/06 |
| 7/2006 | Stephan, G.<br>Rässler, S.<br>Schewe, T. | Wirkungsanalyse in der Bundesagentur für Arbeit : Konzeption, Datenbasis und ausgewählte Befunde<br>published as: Das TrEffeR-Projekt der Bundesagentur für Arbeit : die Wirkung von Maßnahmen aktiver Arbeitsmarktpolitik. In: Zeitschrift für ArbeitsmarktForschung, Jg. 39, H. 3/4 (2006) | 4/06 |
| 8/2006 | Gash, V.<br>Mertens, A.<br>Romeu Gordo, L. | Are fixed-term jobs bad for your health? : a comparison of West-Germany and Spain<br>published in: European Societies, Vol. 9, No. 3 (2007), S. 429-458 | 5/06 |
| 9/2006 | Romeu Gordo, L. | Compression of morbidity and the labor supply of older people | 5/06 |
| 10/2006 | Jahn, E. J.<br>Wagner, T. | Base period, qualifying period and the equilibrium rate of unemployment | 6/06 |
| 11/2006 | Jensen, U.<br>Gartner, H.<br>Rässler, S. | Measuring overeducation with earnings frontiers and multiply imputed censored income data | 6/06 |
| 12/2006 | Meyer, B.<br>Lutz, C.<br>Schnur, P.<br>Zika, G. | National economic policy simulations with global interdependencies : a sensitivity analysis for Germany<br>published in: Economic systems research, Vol. 19, No. 1 (2007), S. 37-55 | 7/06 |
| 13/2006 | Beblo, M.<br>Bender, S.<br>Wolf, E. | The wage effects of entering motherhood : a within-firm matching approach | 8/06 |
| 14/2006 | Niebuhr, A. | Migration and innovation : does cultural diversity matter for regional R&D activity? | 8/06 |
| 15/2006 | Kiesl, H.<br>Rässler, S. | How valid can data fusion be?<br>published in: Journal of Official Statistics, (2006) | 8/06 |
| 16/2006 | Hujer, R.<br>Zeiss, C. | The effects of job creation schemes on the unemployment duration in East Germany | 8/06 |
| 17/2006 | Fitzenberger, B.<br>Osikominu, A.<br>Völter, R. | Get training or wait? : long-run employment effects of training programs for the unemployed in West Germany | 9/06 |
| 18/2006 | Antoni, M.<br>Jahn, E. J. | Do changes in regulation affect employment duration in temporary work agencies? | 9/06 |
| 19/2006 | Fuchs, J.<br>Söhnlein, D. | Effekte alternativer Annahmen auf die prognostizierte Erwerbsbevölkerung | 10/06 |
| 20/2006 | Lechner, M.<br>Wunsch, C. | Active labour market policy in East Germany : waiting for the economy to take off | 11/06 |
| 21/2006 | Kruppe, T. | Die Förderung beruflicher Weiterbildung : eine mikroökonometrische Evaluation der Ergänzung durch das ESF-BA-Programm | 11/06 |
| 22/2006 | Feil, M.<br>Klinger, S.<br>Zika, G. | Sozialabgaben und Beschäftigung : Simulationen mit drei makroökonomischen Modellen | 11/06 |

| | | | |
|---|---|---|---|
| 23/2006 | Blien, U.<br>Phan, t. H. V. | A pilot study on the Vietnamese labour market and its social and economic context | 11/06 |
| 24/2006 | Lutz, R. | Was spricht eigentlich gegen eine private Arbeitslosenversicherung?<br>published als: Zeitschrift für die gesamte Versicherungswissenschaft, Jg. 96, H. 2 (2007), S. 169-208 | 11/06 |
| 25/2006 | Jirjahn, U.<br>Pfeifer, C.<br>Tsertsvadze, G. | Mikroökonomische Beschäftigungseffekte des Hamburger Modells zur Beschäftigungsförderung | 11/06 |
| 26/2006 | Rudolph, H. | Indikator gesteuerte Verteilung von Eingliederungsmitteln im SGB II : Erfolgs- und Effizienzkriterien als Leistungsanreiz? | 12/06 |
| 27/2006 | Wolff, J. | How does experience and job mobility determine wage gain in a transition and a non-transition economy? : the case of east and west Germany | 12/06 |
| 28/2006 | Blien, U.<br>Kirchhof, K.<br>Ludewig, O. | Agglomeration effects on labour demand | 12/06 |
| 29/2006 | Blien, U.<br>Hirschenauer, F.<br>Phan, t. H. V. | Model-based classification of regional labour markets : for purposes of labour market policy | 12/06 |
| 30/2006 | Krug, G. | Kombilohn und Reziprozität in Beschäftigungsverhältnissen : eine Analyse im Rahmen des Matching-Ansatzes | 12/06 |
| 1/2007 | Moritz, M.<br>Gröger, M. | The German-Czech border region after the fall of the Iron Curtain: Effects on the labour market : an empirical study using the IAB Employment Sample (IABS) | 1/07 |
| 2/2007 | Hampel, K.<br>Kunz, M.<br>Schanne, N.<br>Wapler, R.<br>Weyh, A. | Regional employment forecasts with spatial interdependencies | 1/07 |
| 3/2007 | Eckey, H.- F.<br>Schwengler, B.<br>Türck, M. | Vergleich von deutschen Arbeitsmarktregionen | 1/07 |
| 4/2007 | Kristen, C.<br>Granato, N. | The educational attainment of the second generation in Germany : social origins and ethnic inequality | 1/07 |
| 5/2007 | Jacob, M.<br>Kleinert, C. | Does unemployment help or hinder becoming independent? : the role of employment status for leaving the parental home | 1/07 |
| 6/2007 | Konle-Seidl, R.<br>Eichhorst, W.<br>Grienberger-Zingerle, M. | Activation policies in Germany : from status protection to basic income support | 1/07 |
| 7/2007 | Lechner, M.<br>Wunsch, C. | Are training programs more effective when unemployment is high? | 2/07 |
| 8/2007 | Hohendanner, C. | Verdrängen Ein-Euro-Jobs sozialversicherungspflichtige Beschäftigung in den Betrieben? | 2/07 |
| 9/2007 | Seibert, H. | Frühe Flexibilisierung? : regionale Mobilität nach der Lehrausbildung in Deutschland zwischen 1977 und 2004 | 2/07 |
| 10/2007 | Bernhard, S.<br>Kurz, K. | Familie und Arbeitsmarkt : eine Längsschnittstudie zum Einfluss beruflicher Unsicherheiten auf die Familienerweiterung | 2/07 |
| 11/2007 | Drechsler, J.<br>Dundler, A.<br>Bender, S.<br>Rässler, S.<br>Zwick, T. | A new approach for disclosure control in the IAB Establishment Panel : multiple imputation for a better data access | 2/07 |

| | | | |
|---|---|---|---|
| 12/2007 | Fuchs, J.<br>Söhnlein, D. | Einflussfaktoren auf das Erwerbspersonenpotenzial : Demografie und Erwerbsverhalten in Ost- und Westdeutschland | 3/07 |
| 13/2007 | Hartmann, J.<br>Krug, G. | Verknüpfung von Befragungs- und Prozessdaten : Selektivität durch fehlende Zustimmung der Befragten? | 3/07 |
| 14/2007 | Baltagi, B. H.<br>Blien, U.<br>Wolf, K. | Phillips Curve or wage curve? : evidence from West Germany: 1980-2004 | 4/07 |
| 15/2007 | Blien, U.<br>Gartner, H.<br>Stüber, H.<br>Wolf, K. | Expensive and low-price places to live : regional price levels and the agglomeration wage differential in Western Germany | 4/07 |
| 16/2007 | Jaenichen, U.<br>Stephan, G. | The effectiveness of targeted wage subsidies for hard-to-place workers | 6/07 |
| 17/2007 | Fuchs, J.<br>Weber, B. | Vollbeschäftigungsannahme und Stille Reserve : eine Sensitivitätsanalyse für Westdeutschland | 6/07 |
| 18/2007 | Haas, A.<br>Damelang, A. | Labour market entry of migrants in Germany : does cultural diversity matter? | 6/07 |
| 19/2007 | Wachter, T. von<br>Bender, S. | Do initial conditions persist between firms? : an analysis of firm-entry cohort effects and job losers using matched employer-employee data | 6/07 |
| 20/2007 | Reiter, J. P.<br>Drechsler, J. | Releasing multiply-imputed synthetic data generated in two stages to protect confidentiality | 6/07 |
| 21/2007 | Damelang, A. | Räumliche Mobilität von türkischen Arbeitnehmern : eine Analyse mit der IAB-Beschäftigtenstichprobe 2001 | 7/07 |
| 22/2007 | Pfeifer, C. | Homogene und heterogene Teilnahmeeffekte des Hamburger Kombilohnmodells : ein Verfahrensvergleich von Propensity Score Matching und OLS-Regression | 7/07 |
| 23/2007 | Bender, S.<br>Koch, S.<br>Meßmann, S.<br>Walwei, U. | Was muten sich Arbeitslose zu? : Lohnkonzessionen von ALG-II-Empfängern | 7/07 |
| 24/2007 | Bruckmeier, K.<br>Schnitzlein, D. | Was wurde aus den Arbeitslosenhilfeempfängern? : eine empirische Analyse des Übergangs und Verbleibs von Arbeitslosenhilfeempfängern nach der Hartz-IV-Reform | 9/07 |
| 25/2007 | Büttner, T. | Ankündigungseffekt oder Maßnahmewirkung? : eine Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit | 10/07 |

Stand: 12.10.2007

## Imprint

**IABDiscussionPaper**
**No. 26 / 2007**

**Editorial address**
Institut für Arbeitsmarkt- und Berufsforschung
der Bundesagentur für Arbeit
Regensburger Straße 104
D-90478 Nürnberg

**Editorial staff**
Regina Stoll, Jutta Palm-Nowak

**Technical completion**
Jutta Sebald

**All rights reserved**
Reproduction and distribution in any form, also in parts, requires the permission of IAB Nürnberg

Download of this DiscussionPaper:
http://doku.iab.de/discussionpapers/2007/dp2607.pdf

**Website**
http://www.iab.de

**For further inquiries contact the author:**
Herbert Brücker, Tel. 0911/179-3807,
or e-mail: herbert.bruecker@iab.de