Unequal Pay or Unequal Employment? What Drives the Self-Selection of Internal Migrants in Germany?*

Melanie Arntz[†]
ZEW Mannheim & University of Heidelberg
Terry Gregory [‡]
ZEW Mannheim
Florian Lehmer[§]
IAB Nürnberg

Preliminary Version: March 3, 2011

Abstract

This paper examines the determinants of internal migration in a context where wages tend to be rather inflexible at a regional scale so that regional labor demand shocks have a prolonged impact on employment rates. Regional income differentials, then, reflect both regional pay and employment differentials. In such a context, migrants tend to move to regions that best reward their skills in terms of both of these dimensions. As an extension to the Borjas framework, the paper thus hypothesizes that regions with a low employment inequality attract more unskilled workers compared to regions with unequal employment chances. By estimating a migration model for the average skill level of gross labor flows between 27 German regions, we find evidence in favor of this hypothesis. While rising employment inequality in a region raises the average skill level of an in-migrant, higher pay inequality in a region does not have a significant impact on the average skill level of its in-migrants. A higher employment inequality in Eastern as compared to Western Germany may, thus, be the missing link to explain the fact that East-West migrants tend to be rather unskilled.

Keywords: internal migration, self-selection, employment chances, Borjas model

JEL: R23, J31, J61

*Melanie Arntz and Terry Gregory would like to thank the German Research Foundation for financial support.

[†]Melanie Arntz is Assistant Professor at the University of Heidelberg and Senior Research Fellow at the Centre for European Economic Research (ZEW) Mannheim, L7, 1 D-68161 Mannheim, e-mail: arntz@zew.de, phone: +49 621 1235159, fax: +49 621 1235225.

[‡]Terry Gregory is Research Fellow at the Centre for European Economic Research (ZEW) Mannheim, L7, 1 D-68161 Mannheim, e-mail: gregory@zew.de, phone: +49 621 1235306, fax: +49 621 1235225.

[§]Florian Lehmer is Research Fellow at the Institute for Employment Research (IAB) of the Federal Employment Agency, Regensburger Str. 104, D-90478 Nuernberg, e-mail: Florian.Lehmer@iab.de.

1 Introduction

Since German reunification, there have been concerns that East Germany experiences a brain drain, thereby losing an important regional asset for job creation and economic growth.[1]. A major reason to believe that migrants from Eastern to Western Germany may be positively selected with respect to skills can be derived from the Borjas model on the self-selection of immigrants (1987). According to this framework, skilled workers should be attracted to regions that best reward their abilities by paying high returns to their skills as reflected in a high variance of the wage distribution. With the wage variance in Western Germany exceeding Eastern levels, we would thus expect east-west migrants to be positively selected, although the convergence in the relative wage inequality that has been found by Burda and Hunt (2001) and Gernandt and Pfeiffer (2008) points towards a reduced relevance of this selection mechanism. Empirical evidence for the 1990s suggests that East-West migrants are disproportionately skilled in terms of observable and unobservable skills (Burda and Hunt 2001, Hunt 2000, Bruecker and Truebswetter (2007)). In contrast, Granato and Niebuhr (2009) present evidence that the net migration rate of unskilled migrants even exceeds that of skilled migrants after 2000, a finding that contradicts the idea that the selection mechanism is solely based on regional differences in wage inequality.

One underlying reason for a wage-based selection mechanism to fail could be that central wage bargaining in Germany prevents a flexible wage adjustment at the regional level as has been found by Topel (1986) for the US. Consistent with this finding Möller (1995) and Mertens (2002) show that wages in Germany do not exhibit any reaction to regional labor demand shocks which therefore tend to have a prolonged impact on regional employment rates. Similar results have been found for Europe by Decressin and Fatas (1995) and Abraham (1996). In such a context, modeling the migration decision as a wage-maximizing process may be inadequate since income differentials that drive migration decisions may result from employment rather than wage differentials. Not surprisingly, therefore, empirical tests of the Borjas framework perform rather poor in the German context (Arntz 2010, Brücker and Trübswetter 2007) compared with studies by Borjas, Bronars, and Trejo (1992) and Hunt and Mueller (2004) that demonstrate its relevance for internal migration in the US. Hence,

[1]For studies on the effect of human capital on growth see for instance Fratesi and Riggi (2007) and Kanbur and Rapoport (2005).

in a context with regional wage rigidities, we need to base the analysis of migrant selectivity on an extended framework in the manner of Harris and Todaro (1969) that allows for regional differences in employment chances.

This paper therefore suggests that skill-specific regional employment chances may be a missing link to fully explain skill-selective migration in a context where wages are rather inflexible at a regional scale. In particular, we argue that workers choose regions that best reward their skills in terms of both employment chances and wages. For unskilled workers, for example, we hypothesize a self-selection into regions with a low employment inequality. This prediction results from the assumption that employment inequality is a measure of the returns to skills in terms of employment. We consider this a plausible assumption since unskilled individuals are the ones who are most prone to unemployment (Möller and Schmillen 2008, Reinberg and Hummel 2007, Boockmann and Steffes, 2010). This unemployment risk may reflect that unskilled individuals are more likely to be atypically employed than their skilled counterparts (Giesecke, 2009). Moreover, firms tend to hoard skilled rather than unskilled labor during economic downturns (Nickell and Bell, 1995). In line with such arguments, Mauro (1999) finds that adverse regional shocks increase unemployment mainly among unskilled workers. Thus regions with a high employment inequality penalize unskilled workers and should thus attract predominantly skilled workers.

Therefore, this paper contributes to the literature by extending the Borjas framework to allow for a selection mechanism based on both wage and employment differentials. This model predicts the average skill level of a migration flow to be a positive function of the wage and employment inequality in the destination as compared to the origin region. Moreover, unlike the Borjas framework, the model predicts that mean wage and employment differentials also induce a positive skill sorting. The second contribution of the paper thus is to test the predictions of this extended migration model for the average skill level of gross labor flows between 27 German regions. For this purpose, we make use of the full sample of the employment register data in order to be able to determine the skill content for each of these flows with regard to both observable and unobservable skills. We then regress these skill measures on the mean and the dispersion of the regional wage and employment distribution. Thus, instead of only conditioning on the regional unemployment rate as is done is most migration models, we capture not only the average risk of being unemployed, but also allow regions to differ in how this risk is spread among the local workforce. As a third contribution to the literature, we are able to exploit the panel dimension of our data. Rather than estimating the self-selection of migrants based on a cross-section only, we are able to include a fixed effect for each labor flow, thus conditioning on average time constant utility differentials between two regions (e.g. amenity differentials) that may otherwise bias the estimation results.

The findings suggest that, as expected, the average skill level of a migration inflow increases with the dispersion of employment chances in the destination compared to the origin region. Moreover, mean differentials in wages and employment also tend to increase the average skill level of a labor flow, while the wage dispersion has no significant impact. Apparently, the high skill level of an average West-East migrant and the relatively low skill level of an average East-West migrant may partially be attributed to the poorer employment chances for unskilled relative to skilled workers in Eastern compared to Western Germany. The structure of the paper is as follows. Section 2 proposes an extended theoretical framework for the self-selection of migrants. Section 3 describes the data base and the definition of the covariates used for the subsequent analysis. Section 4 presents descriptive evidence on gross migration flows in Germany as well as on the distributions of both wages and individual employment chances. The section also includes a discussion of German East-West Migration. Section 5 describes the estimation strategy for the empirical analysis and presents the findings together with several robustness checks for the estimations of the extended migration model. Finally, section 6 concludes.

2 Theoretical Framework

Our theoretical framework builds upon Borjas, Bronars, and Trejo (1992) who formalize the self-selection of interstate migrants and test their model in the US context. Their framework is linked to the self-selection of workers as described by Roy (1951) and the extension of this approach to the self-selection of immigrants as developed by Borjas (1987). However, while the latter approach focusses on the selection based on unobservable abilities, Borjas, Bronars, and Trejo (1992) focus on the selectivity of internal migrants with respect to both observable skills and unobservable abilities. Our framework extends their theoretical model by allowing for unemployment. As a consequence, migration decisions do not depend on differentials in regional wage distributions alone, but also hinge on the probability of receiving this wage, i.e. the probability of being employed as has already been discussed by Todaro (1969). As a consequence, differentials in regional income distributions that reflect both the employment and the wage distribution affect migration decisions and the selectivity of internal migrants.

Consider $j = 1, \ldots, J$ regions that only differ with respect to the income distribution. For ease of exposition, our theoretical framework thus abstracts from other utility differentials between regions such as regional amenities or disamenities including regional price differentials.[2] An income-maximizing individual chooses to live in region j if

$$\pi_j > \underset{i \neq j}{max}[\pi_i] \tag{1}$$

with π_j as the income in region j which is the product of the probability of being employed e_j on any particular workday in region j and the wage w_j paid in this region if employed. Now assume, analog to Borjas, Bronars, and Trejo (1992), that the population wage distribution can be decomposed into a part reflecting the mean wage μ_w that is independent of an individual's skills and abilities and a part that measures person-specific deviations from this mean wage that depend on individual i's skills υ_i and the returns to skills paid in region j. We assume that skills and abilities are perfectly transferable between all regions, an assumption that we consider justified for internal migration.[3] The population wage distribution in region j can then be written as

$$w_j \quad = \quad \mu_{w_j} + \eta_{w_j} \upsilon, \tag{2}$$

where υ is considered as a continuous random variable with mean zero and an unlimited range over the real numbers that reflects the region-invariant skill distribution of all skills and abilities. An individual's potential wage is thus determined by his position in the skill distribution and the region-specific returns to these skills η_{w_j}. We further assume an analog decomposition of the population employment distribution which can be written as

$$e_j \quad = \quad \mu_{e_j} + \eta_{e_j} \upsilon, \tag{3}$$

[2]Our empirical approach controls for time-constant regional differentials and thus takes account of much of these utility differentials.

[3]While this assumption may be unproblematic with Western and Eastern Germany, it is less clear whether the assumption can be applied to migration across the former German border since East German skills have partially been depreciated after re-unification. We will thus run some sensitivity analysis in section **??**.

with μ_{e_j} as the average probability of employment on an average workday and υ as defined above. Hence, an individual's employment probability is determined by the average employment probability and the region-specific returns to his skill level in terms of employment η_{e_j}. Typically, migration models only consider interregional differences in average employment chances by using the regional unemployment rate as a corresponding indicator. The disadvantage of this approach is that it only tells us something about average employment chances and nothing about how the risk of being out of work is spread across the local workforce and especially across different skill groups. The literature suggests, however, that unqualified individuals are more prone to unemployment, especially in the last two decades. Bynner and Parsons (2001) show for the U.K., for example, that contrary to older cohorts, the lack of certain qualifications significantly increases the incidence of unemployment by about 30% among younger cohorts born in the 1970s. Consistent with these findings, Möller and Schmillen (2008) find that the unemployment incidence is much higher among unqualified individuals in Germany and that the share of individuals who are at risk of being unemployed is increasing among younger cohorts. Similarly, Wilke (2005) finds skilled individuals to be much less prone to unemployment than their unskilled counterparts. All these findings are in line with a theoretical model developed by Helpman at al (2010) that suggests that the unemployment rate is decreasing in worker ability, whereas the average wage is increasing in worker ability. In light of the literature, we therefore assume that unskilled workers are situated at the bottom of the distribution of both wages and employment.

If we thus apply the two decompositions in equation (2) and (3), the income in region j can then be written as

$$\begin{aligned}\pi_j &= (\mu_{w_j} + \eta_{w_j}\upsilon) \cdot (\mu_{e_j} + \eta_{e_j}\upsilon) \\ &= \mu_{w_j}\mu_{e_j} + (\mu_{e_j}\eta_{w_j} + \mu_{w_j}\eta_{e_j} + \eta_{w_j}\eta_{e_j})\upsilon \qquad (4) \\ &= M_j + R_j\upsilon \qquad (5)\end{aligned}$$

where the first term M_j corresponds to the average income in region j, and the second term R_j reflects all region-specific returns to skills.[4] The returns to migration thus differ across skill groups,

[4]Note that, analog to Borjas, Bronars, and Trejo (1992), we assume relative prices of all skills to be region-invariant so that we do not have to operate with a multifactor model of ability.

thereby inducing a sorting of individuals into regions that best reward their particular skills. In particular, the utility differential between region j and k can be written as

$$U_{jk} = \pi_j - \pi_k = M_j - M_k + (R_j - R_k)\upsilon \tag{6}$$

and thus depends on an individual's skills and abilities. The utility differential now depends on the parameters of the wage and employment distribution in both regions. Note that changes in average employment and average wage level affect the mean differential as well as the differential in the returns to skills. The partial effect of increasing average employment and average wages in j can thus be written as

$$\frac{\partial U_{jk}}{\partial \mu_{w_j}} = \mu_{e_j} + \eta_{e_j}\upsilon \quad \text{and} \quad \frac{\partial U_{jk}}{\partial \mu_{e_j}} = \mu_{w_j} + \eta_{w_j}\upsilon. \tag{7}$$

Therefore, individuals will be rewarded by an increasing average wage to the extent that the individual is employed in region j. While region j becomes more attractive for all potential migrants, skilled individuals with a higher υ gain the most from an increasing mean wage so that the migration flow from k to j should become more skilled on average. The same argument can be applied to an increase in the average employment probability. While all potential migrants benefit from such an increase by reaching higher income levels, the gain is more pronounced the higher is the wage an individual receives for each additional day employed. For high-skilled individuals, this gain should thus exceed the gain for less-skilled individuals so that the migration flow from k to j should again become more skilled on average. Note that the opposite predictions hold for increases in the average wage and employment in region k.

Changes in the returns to skills in terms of both wages and employment chances only affect the second part of equation (6). In particular, the partial effects for changes in region j can be written as

$$\frac{\partial U_{jk}}{\partial \eta_{w_j}} = (\mu_{e_j} + \eta_{e_j})\upsilon \quad \text{and} \quad \frac{\partial U_{jk}}{\partial \eta_{e_j}} = (\mu_{w_j} + \eta_{w_j})\upsilon. \tag{8}$$

An increasing wage inequality will thus attract skilled individuals with a positive υ who can expect to benefit from the increasing returns to skills to the extent that they are employed in region j. In contrast, individuals with a below average skill level will experience an income and thus utility loss if wage inequality increases. The skill composition of the migration flow from region k to j should thus get more skilled on average if either wage inequality or employment inequality increases.

Note that we have $J(J-1)$ migration flows and thus any comparative static should take account of all the parameters in the model, i.e. the employment and wage distributions in each region. However, the partial effects derived in equation(7) and (8) can be considered to reflect the predictions of the model holding all other parameters in all other regions constant. In our later estimation approach, we will have to ensure that such a ceteris paribus condition holds.

In addition, three further issues warrant a short discussion. First of all, we neglect the case that any of the regions is unpopulated. Thus, we assume each region to make a competitive offer to at least some individuals. Thus, for any three regions $j-1, j, j+1$ that are adjacent in terms of their returns to skills with $R_{j-1} < R_j < R_{j+1}$, region j can only exist if the mean income in region j satisfies

$$M_j > \frac{(R_{j+1} - R_j)M_{j-1} + (R_j - R_{j-1})M_{j+1}}{R_{j+1} - R_{j-1}}. \tag{9}$$

If mean incomes where the same across al regions, all skilled individuals would prefer the region with the highest return to their skills whereas individuals lacking these skill would be attracted to regions with the lowest penalty from lacking these skills. Thus, a region that ranks in the middle in terms of the returns to skills can only exist if it offers a competitive mean income level that exceeds the mean income level in at least one of the neighbors. The existence condition thus rules out cases where the relationship between the mean income M and the returns to skills R is flat or even inversely U-Shaped, but allows for a monotonously decreasing or increasing as well as for a U-shaped relationship. While this does not differ from the insights in Borjas, Bronars, and Trejo (1992), the extended model suggests an important derivation. In particular, unlike the simple wage-based framework, the relationship between μ_w and η_w can be inversely U-shaped as long as the relationship between μ_e and η_e makes up for it by a U-shaped relationship. The extended model thus implies that a region can compensate a disadvantage in terms of its wage distribution and thus ensure its existence by a favorable employment distribution and vice versa.

Secondly, it may be helpful to discuss the reasons why we expect flows in opposing directions to exist. This may be the case because there are new cohorts entering the labor market each period among which a certain share are likely to be mismatched to their origin region in terms of their skills. Thus, while the most able individuals will leave their region for regions with a higher returns to their skills, less-skilled individuals may prefer the opposite direction in order to minimize the penalty from lacking skills. Moreover, individuals for whom a particular region once offered the optimal return to their skills need not be optimally matched forever if individuals shift their position in the skills distribution due to training effects or due to the depreciation of skills.

Finally, the model abstracts from a number of potential complications such as regional amenity differentials, price differentials as well as from migration costs. As long as these components are not correlated to the skill level, the key results remain unchanged. However, there are reasons to believe that migration costs decrease in abilities if abilities facilitate the gathering of information and reduce the psychological costs of migration. If this is the case, the key results of the theoretical model remain unchanged only conditional on such costs. Similarly, if individuals differ in how they value certain regional amenities and disamenities depending on their skill level as has been argued by Glaser et al. (2005), the key results also remain intact only conditional on regional differentials in amenities and disamenities. Our estimation approach thus needs to take account of these complicating factors.

3 Data

In the previous section we derived clear empirical predictions about the relationship between regional differences in the returns to skills and the resulting direction and skill composition of migration flows. In order to test these predictions, we analyze the employment register data (BeH) of the German Federal Employment Agency. This administrative data set contains information on the population working in jobs that are subject to social insurance payments, thus excluding civil servants and self-employed individuals. We are thus able to reconstruct individual employment histories including periods of employment and periods of unemployment benefit receipt on a daily basis. For each employment spell, the data contains individual and firm-level characteristics including the daily gross wage, the educational attainment as well as the micro-census region of the workplace. We are thus able to identify gross labor flows rather than migration flows between regions by comparing workplaces before and after an interregional job transition. Though the restriction to labor flows is a shortcoming, the data set is the only data set in Germany that allows for observing wages, skill levels and interregional moves on the basis of a large sample. Moreover, our theoretical predictions should be applicable to labor flows as well, although our results should not be interpreted as reflecting general migration flows in Germany. The sample is restricted to the time period between 1995-2004, since for the years before 1995 no reliable information on the workplace relocation between East and West Germany is possible. Furthermore, we focus on men between the age of 16 and 65 because women's lower labor force attachment would aggravate the selectivity of the sample used for the analysis.[5].

For all subsequent analyzes, we distinguish between 27 aggregated planning districts.[6] These regions lump together 97 German planning districts ('Raumordnungsregionen') that are defined according to commuting ranges and thus already comprise labor market regions that are relatively self-contained. In order to ensure a sufficient number of job moves between each of these regions for different skill levels, we further aggregate these planning districts to 27 larger regions. We do so based on an algorithm that reduces the remaining external commuting linkages, thereby ensuring that the regional division still reflects relatively self-contained labor markets.[7] For each of these 27 regions we estimate the key explanatory variables, namely the returns to skills reflected be the employment and wage distribution, for each year during the observed time period. Altogether, the 27 regions result in 702 gross labor flows whose size and skill composition we calculate for each year and region. The following subsection discusses the corresponding details.

3.1 Data on Level and Skill-Composition of Interregional Labor Flows

For the computation of interregional labor flows, we exploit information on the entire working population, i.e. we use the full employment register data (BeH) that is only available to researchers at the Institute for labor Market Research (IAB). For the computation of labor flows, we use yearly cross sections to the cut-off date June 30th and compare the workplace location between two consecutive years. We are thus able to calculate the gross labor flows by identifying the origin and the destination

[5]We exclude men attending military or civilian service since they are centrally registered so that the identification of their exact location is not possible and neglect apprentices and all employment spells with minor employment, since its definition changed in 1999.

[6]For a map and a complete list of regions see Figure **??** and Table **??** in the appendix.

[7]Details on the algorithm is available from the authors upon request.

region for all interregional job moves. Note that the identification of an interregional job move necessitates an individual to be employed on June 30th of two consecutive years. While the sample may thus include individuals who have been unemployed between these two cut-off days, long-term unemployed are clearly underrepresented in our data, a shortcoming that we have to be aware of when discussing our estimation results.

We then calculate the average skill level of each gross labor flow by constructing three alternative measures. The first measure corresponds to the average years of completed education among the movers of each labor flow. This measure reflects part of the differences in observable skills across labor flows. The calculation of the second and third measure is based on ranking individuals in the predicted and residual wage distribution. The underlying idea is that wages reflect the marginal product of labor and may thus proxy for abilities and skills.[8] More precisely, following the approach by Borjas, Bronars, and Trejo (1992), we estimate daily gross wages[9] with the following Fixed-Effects (FE) model for all individuals in the sample:

$$log\ w_{ijt} \quad = \quad \beta_1 REGION_{it} + \beta_2 YEAR + \beta_3 X_{it} + \varepsilon_{ijt} \tag{10}$$

where w_{ijt} is individual i's daily gross wage in region j and year t. The wage is a function of a vector of dummy variables indicating the workplace location (REGION), a vector of dummy variables indicating the year of the observation (YEAR), a vector of control variables (X). The control variables include all observable characteristics that affect an individual's level of productivity such as age, age squared, occupation (8 categories), industry (10 categories), establishment size, educational attainment (six levels) and tenure, defined as the number of working years in the current company. The control vector also includes a dummy for part-time employment, because we do not observe hourly wages, but only daily wages that may differ between fulltime and part time employees due to different working hours. We thus use the available information on part time employment to control for such differences. The error term ε_{ijt} depends on the match between workers and regions. By

[8]Ideally, we would rank individuals according to their income residuals. However, we are not able to estimate the income distribution for the full BeH data because the data is reduced to a cross-section that lacks information on the previous employment history. Extending the data to include the full employment history is impossible due to the resulting size of the data.

[9]Unfortunately, around 15% of all wages are top-coded at the contribution limit of the social security. Therefore, we impute the censored wages with an estimation procedure described by Gartner (2005). This procedure adds a randomly drawn error term to the predicted wage level and thereby avoids a strong correlation between the error term and the explanatory variables.

assuming that the individual's skill level is region-invariant, the error term can be written as

$$\varepsilon_{ijt} = \eta_j(\nu_i + u_{it}), \tag{11}$$

where ν_i comprises the remaining stock of *unobservable* person-specific characteristics, η_j refers to the region-specific returns to human capital and u_{it} denotes a random error term. Equation (11) thus states that the error term is proportional to η_j. When estimating the model in equation (10) after demeaning at the individual level, the average residual of an individual thus reflects ν_i as long as the regional dummies capture η_j. In other words, identification of the ν_i comes from individuals moving between regions and receiving a wage that is proportional to the region-specific returns to their region-invariant skills and abilities.

Based on this approach, we define two wage-based measures for the skill content of each labor flow. The first measure is based on the predicted wages of equation (10). For this we calculate the average predicted wage among the movers of each labor flow which corresponds to the number of standard deviations that their wage is above or below the mean wage due to deviations in observable abilities. We call this measure the "observed" skill measure. The second measure is based on estimating equation (10) without any controls in X. An individual's average position in the resulting residual wage distribution, ν_i, then reflects unobservable and observable skills that are time-constant. By computing the average of these individual fixed effect for all migrants following a particular migration path, we get a measure of the overall skill content of each flow. We call this measure the "overall " skill measure.

We are thus able to analyze three measures of the skill content of each labor flow which reflect either of the following three only: years of completed education, observable skills and overall skills including both observable and unobservable skills. Note that we have the corresponding skill measures for each labor flow on a yearly basis so that we can exploit the panel dimension in the subsequent analysis. In addition, we calculate the number of movers for each labor flow per quintile of the skill distribution according to the observed and unobserved skills measure as well as for each of the six education groups, in order to compare the skill structure of each labor flow in more detail.

3.2 Data on Regional Returns to Skills

The theoretical model presented in Section 2 implies that the sorting of skills across regions is largely determined by the region-specific returns to skills parameters R_j. According to the theoretical model, the region-specific returns to skills are reflected by the mean and dispersion of the regional wage and employment distribution. Since we distinguish between movers observable and unobservable skills, we need to construct appropriate regional returns to skills measures that takes this into account.

3.2.1 Regional Wage Distribution

For the regional mean wage we calculate the average predicted wage of the regional workforce that results from separate region- and year-specific OLS-regressions of equation (11). By estimating this model separately across years and regions, we allow for different returns to observable characteristics across space and time. For the regional wage dispersion we construct two different measures. As a first measure, we compute the standard deviation of the predicted wages for the regional workforce reflecting region- and year-specific returns to observable skills only. This measure of wage inequality should be able to explain the selectivity of migration with respect to observable skills. In order to explain the selectivity of labor flows with regard to the overall skill content, we simply use the wage variance at the regional level.

3.2.2 Regional Employment Distribution

For the regional employment distribution we construct similar measures. For this purpose, we first compute the number of days a worker is employed during a year based on the two percent random sample of the employment register data that contains full spell information on periods of employment and unemployment.[10] While the regional mean and the regional variance of the days employed can be used as a measure to explain the skill selectivity of labor flows with regard to overall skills, we again need a second measure that reflects returns to observable characteristics only. For this

[10] One problem is that for periods of unemployment without receipt of income, transfers from the German employment agency are not identifiable (see Fitzenberger and Wilke, 2010, and Lee and Wilke, 2009 for more details). Using the institutional setting in Germany, we define unemployment by all periods of nonemployment after an employment period which contain at least one period with income transfers by the German federal labor office (for more details see Fitzenberger and Wilke (2010)'s definition of 'nonemployment'.) Furthermore, we weigh part-time jobs according to their amount of hours worked. In particular, we weigh days, where part-time work with more than 18 hours, by 0.6 and those with less than 18 hours with 0.25.

purpose, we need to estimate a model of the number of days employed as a function of observable characteristics. Note, however, that the number of days employed during a year comes with mass points at 0 and 365 employed days. We thus need to take account of this inconvenient distribution by modeling the different cases separately.

Let $I_{ijt} = 0, 1, 2$ denote an individual-specific indicator function that depends on the number of days individual i is employed during a particular year t in region j, d_{ijt}:

$$I_{ijt} = \begin{cases} 2 & if \quad d_{ijt} = 365 \\ 1 & if \quad 0 < d_{ijt} < 365 \\ 0 & if \quad d_{ijt} = 0 \end{cases}$$

In turn, individual i's expected number of employed days depends on the probability of being employed all-year-long ($I_{ijt} = 2$), between 0 and 365 days ($I_{ijt} = 1$) and all year long ($I_{ijt} = 0$). According to the law of total probability, individual i's conditional expected employed days in region j at time t can be written as follows:

$$E[d_{ijt}|X_{it}] = P(I_{ijt} = 1|X_{it})E[d_{ijt}|I_{ijt} = 1, X_{it}] + P(I_{ijt} = 2|X_{it})E[d_{ijt}|I_{ijt} = 2, X_{it}] \quad (12)$$

where the conditional probabilities $P(I_{ijt} = 0|X_{it})$, $P(I_{ijt} = 1|X_{it})$ and $P(I_{ijt} = 2|X_{it})$ add up to unity. Note that $E[d_{ijt}|I_{ijt} = 0] = 0$ and $E[d_{ijt}|I_{ijt} = 2, X_{it}] = 365$. We calculate equation (12) for each region and year using predicted values of the conditional probabilities that are estimated within a multinomial logit framework.[11] The conditional expected values $E[d_{ijt}|I_{ijt} = 1, X_{it}]$ are predicted within region- and year-specific OLS-regressions. For the construction of the employment dispersion we again construct two different measures. As a first measure we calculate the standard deviation of the conditional expected employed days as follows:

$$sd\big[d_{ijt}|X_{it}\big] = \sqrt{E[d_{ijt}^2|X_{it}] - (E[d_{ijt}|X_{it}])^2} \quad (13)$$

Equation (13) then captures region- and year-specific returns to observable skills only. Arguing similarly as above, the higher are the rewards to observable skills characteristics in a region in terms of higher expected days employed, the higher is the dispersion of predicted employed days in such a region. As a second measure we calculate the dispersion of employment residuals. For this we first

[11] We thereby assume that the assumption of independence from irrelevant alternatives (IIA) is fulfilled.

calculate the dispersion of the residual employed days u_{ijt} as follows:

$$sd\big[u_{ijt}\big] \quad = \quad sd\big[d_{ijt} - E[d_{ijt}|X_{it}]\big] \tag{14}$$

Equation (14) then reflects region- and year-specific returns to unobservable skills only.

Thus, we receive wage- and employment-based region- and year-specific returns to skills that reflect either returns to observable or unobservable skills. These regional returns to skills measures will serve as explanatory variables in the following estimations of a gross migration model.

4 Descriptives

As we have derived from the theoretical framework in section 2, differences in the skill composition of gross migration flows may be explained by interregional differences in the regional returns to skills that are reflected by both the wage and the employment distribution. The present section provides corresponding descriptive statistics on migrants average skills characteristics, the size and skill composition of labor flows in Germany as well as on the regional returns to skills across labor market regions. The statistics aim to show that their exists large variation in the skill content of labor cross flows as well as in the regional returns to skills across regions, a fact that we want to exploit in our empirical analysis.

Table 4 shows means and standard deviations for movers, stayers and for the total amount of workers of the sending region. Standard deviations are in parenthesis. The values are mean values across the entire 10 year time period between the years 1995-2004. The table shows that, compared to the previous year, on average 2.61 percent of a region's labor force relocate to another labor market region. In total we are able to analyze 3.55 million job moves during the observed time period. According to Tabel 4 an average migrant has 8 months more completed education compared to an average stayer. Furthermore, looking on wage-based skill measures shows that migrants are a positive selection with regard to observable skills, i.e. the predicted wage of an average migrant is 0.26 log points higher compared to stayers. The result is in line with other studies on the self-selection of migrants based on observable skills characteristics as for instance the study by Cutillo and Ceccarelli (2010). However, migrants are a negative selection with respect to unobservable skills

Table 1: Means and Variables of Movers and Stayers

Variable	*Movers*	*Stayers*	*All*
Movers (in percent)	2.61	97.39	100
	(.01)	(.01)	(0)
Skill measures:			
Years of education	13.1	12.4	12.42
	(.4)	(.32)	(.32)
All skills	.005	-.029	-.028
	(.103)	(.092)	(.092)
Unobservable skills	-.277	-.068	-.074
	(.248)	(.234)	.234)
Observable skills	4.73	4.471	4.478
	(.284)	(.29)	(.29)
Selected observable human capital characteristics:			
Log daily wage	4.46	4.4	4.4
	(.14)	(.17)	(.17)
Age	37.05	39.95	39.88
	(.89)	(.81)	(.8)
Parttime workers (in percent)	2.01	3.15	3.12
	(.01)	(.01)	(.01)
Percent of workers with..			
no educational degree	7.58	11.67	11.57
vocational training degree	55.97	65.69	65.44
high school degree	.7	.67	.67
high school degree and vocational training degree	4.49	2.97	3.01
technical college degree	5.92	4.14	4.18
university degree	10.28	5.82	5.93
Percent of workers in age group..			
15<age<=21	1.46	2.27	2.25
21<age<=25	7.92	6.3	6.34
25<age<=35	40.14	29.21	29.5
35<age<=45	30.57	30.77	30.77
45<age<=65	19.91	31.44	31.14
Sample size (in 1000)	3,550	133,381	136,931

Note: Standard deviations in parenthesis.

Table 2: Cross-Correlations of Average Skill Measures of Gross Labor Flows

Variables	Obs. skills	Unob. skills	All skills	Years of education
Observable skills	1.000			
Unobservable skills	-0.798	1.000		
Overall skills	-0.181	0.714	1.000	
Years of education	0.212	0.191	0.641	1.000

characteristics, i.e. holding constant all observable characteristics, migrants earn 0.21 log points less than stayers. The latter may also be explained by the fact that a movers' decision to relocate is often caused by a company's shutdown. Moreover, in the empirical literature it has often been stated that sorting based on observable and unobservable skills need not go in the same direction (Dostie and Leger, 2009). Overall, the results suggest that migrants in Germany are a positive selection in observable terms of the population sending region . However, within groups with similar characteristics, migrants constitute a negative selection in unobservable terms.

Table 2 reports cross-correlations between all skill measures of migration flows that we have constructed. The table shows that the wage-based skill measure capturing overall skills, i.e. both observed and unobserved skills, is highly correlated with average years of education and the average unobservable skills measure. Both measures thus seem to capture the overall skill content of labor. Furthermore, Table 2 shows that the observed wage-based skills measure is negatively correlated with the unobserved measure of skills.

Tables 7 and 8 in the appendix show average out- and in-migration rates, absolute size of out- and in-migration and corresponding average skill levels for the entire time period 1995-2004. The tables show substantial interregional variation in the size and skill-composition between labor market regions. In particular, East-German regions have thus been loosing many of their low educated workers during the observed time period. The result is in line with Granato, Haas, Hamann, and Niebuhr (2009) who study East-West migration during the time period 2000-2006 and find that net migration rates of low qualified workers are higher than net migration rates of high qualified workers.

Table 3 shows average and dispersions of both wages and employment of all 27 labor market regions. For a regional map see Figure 6 in the appendix. For instance, in region 2, on average, 255 thousand

Table 3: Averages and Dispersions of Regional Wages and Employment across 27 Labor Market Regions in Germany

Region	*Average Population (in 1000)*	*Average daily wage (in Euro)*	*Average. employment (in days)*	*Wage dispersion (in Euro)*	*Employment dispersion (in days)*
			East German Regions		
2	255	59.9	284.9	1.46	127.36
7	176	60.5	283.2	1.44	130.73
8	876	71.3	290	1.59	129.57
16	371	59.6	294.6	1.44	121.33
17	384	61.5	286.9	1.47	129.61
18	530	60.3	294.4	1.47	122.79
Mean East	432.0	62.2	289.0	1.5	126.90
			West German Regions		
1	258	80.7	308.6	1.47	116.08
3	605	90.6	319.4	1.56	108.13
4	411	83.3	312.1	1.47	113.11
5	431	83.4	327.3	1.43	97.72
6	561	88.5	314.7	1.49	113.2
9	461	85.4	324.5	1.43	101.59
10	446	88.3	311.3	1.47	118.11
11	1082	92	314.3	1.52	115.08
12	714	92.5	321.1	1.54	107.85
13	336	83.4	324.5	1.42	100.96
14	445	82.6	320.2	1.43	105.79
15	933	97	329.1	1.57	97.72
19	615	90.4	322.9	1.5	105.15
20	318	86.4	332.6	1.43	91.47
21	751	99.3	333.5	1.51	92.13
22	342	88.4	335	1.45	87.21
23	537	89.3	333.9	1.47	90.86
24	422	86.9	326.8	1.51	97.03
25	456	80.7	327	1.42	91.4
26	294	90.5	331.4	1.45	92.61
27	699	98.7	333.7	1.59	89.12
Mean West	529	88.5	324.0	1.5	101.5

individuals participate in the labor market. The average daily gross wage of these workers is 60 Euros and their average number of days employed is 285 days. The dispersion of daily wages is 1.45 Euros and the dispersion of days employed is 127 days. Table 3 shows large variations in wages and employment within and between East and West Germany. In particular, average daily wages and average employed days are higher in West as compared to East Germany. However, the average wage dispersion is similar in East and West Germany and hardly shows any regional variation. The result is in line with Burda and Hunt (2001) and Gernandt and Pfeiffer (2008) who find evidence in favor of a convergence in the relative wage inequality between East and West Germany. In contrast, the dispersion of days employed varies markable between and within both parts of the country. In particular, East German regions show high levels of employment dispersion compared to West German regions. The result is in line with Grotheer and Struck (2005) who find a lower employment stability in East compared to West Germany. Thus, while markable regional differences in the dispersion of employment chances exist, regional differences in the dispersion of wages are less pronounced in Germany. The result suggests that a wage-based selection mechanism might not explain the large variation of migration selectivity that we observe. Rather, the descriptives suggest an employment-based selection mechanism might explain the differences in the skill composition across interregional labor flows in Germany.

Table 4 shows all 7020 gross labor flows ranked according to their average overall skill level as well as corresponding standardized interregional returns to skills characteristics. The idea is to see if flows with high average skill content are those flows where the regional returns to skills are higher in the destination region compared to the region of origin, as theory suggests. For this we create quintiles of a flow's average overall skill level and distinguish flows between and within East and West Germany. In this spirit, the first quintile in column one of Table captures the average skill level and corresponding interregional returns to skills of the 10 percent less qualified labor flows within West Germany. In contrast, the fifth column in Table describes the 10 percent best qualified labor flows within West Germany. According to Table , migrants among the first quintile of West-West flows are on average 0.05 standard units less skilled in terms of observable and unobservable skills compared to average workers in the region of origin. In terms of unobservable skills only, an average migrant among this flow is 0.32 standard units less skilled compared to natives in the

Table 4: Labor Flows Ranked According to Their Overall Skill Level and Corresponding Regional Returns

	Average log wage residuals with regard to:			*Standardized interregional differences in the regional returns to skills:*			
Quintile of Selectivity	*All skills*	*Unobservable skills*	*Years of education*	*Average wage*	*Average employment*	*Wage dispersion*	*Employment dispersion*
West-West Flows							
1. Quintile	-0.05	-0.32	0.51	-0.16	0.03	-0.38	-0.09
2. Quintile	0.03	-0.23	0.91	0.00	-0.02	0.04	0.02
3. Quintile	0.08	-0.17	1.18	-0.03	0.02	-0.06	-0.04
4. Quintile	0.13	-0.11	1.43	0.07	-0.01	0.20	0.05
5. Quintile	0.21	-0.02	1.68	0.12	-0.02	0.20	0.07
West-East Flows							
1. Quintile	-0.22	-0.48	0.40	-2.35	-2.13	-0.46	1.85
2. Quintile	-0.15	-0.40	0.68	-2.29	-2.06	-0.37	1.70
3. Quintile	-0.10	-0.34	0.89	-2.16	-1.83	-0.11	1.50
4. Quintile	-0.04	-0.25	1.16	-2.06	-1.95	0.03	1.62
5. Quintile	0.07	-0.15	1.74	-1.81	-1.98	0.62	1.73
East-West Flows							
1. Quintile	-0.06	-0.44	-0.44	2.02	2.35	-0.23	-2.19
2. Quintile	-0.00	-0.37	-0.20	2.09	2.15	-0.04	-1.90
3. Quintile	0.04	-0.30	-0.12	2.19	1.95	0.12	-1.60
4. Quintile	0.08	-0.24	0.14	2.24	1.83	0.28	-1.42
5. Quintile	0.17	-0.13	0.61	2.12	1.67	0.16	-1.27
West-West Flows							
1. Quintile	-0.04	-0.16	0.06	-0.65	-0.03	-1.33	-0.15
2. Quintile	0.02	-0.10	0.12	-0.02	0.02	-0.08	-0.04
3. Quintile	0.05	-0.09	0.24	0.17	0.05	0.32	0.02
4. Quintile	0.09	-0.03	0.36	0.22	0.01	0.47	0.04
5. Quintile	0.14	0.03	0.47	0.28	-0.05	0.59	0.13

Note: All average skill levels are relative to the population in the region of origin.

region of origin. Furthermore, migrants among this flow have on average 0.51 years more completed education compared to natives in the region of origin. For each quintile Table also shows standardized interregional returns to skills, i.e the standardized value in the destination minus the standardized value in the region of origin. Looking for instance again at the first quintile of all West-West flows, i.e. the lowest qualified flows within West Germany, shows that the average daily wage and the wage dispersion are 0.16 and 0.38 standard units lower in the destination region compared to the region of origin. Moreover, while average employment is 0.03 standard units higher, the employment dispersion is 0.09 standard units lower in the destination compared to the region of origin. Overall, Table shows that flows in the first quintile of selectivity, i.e. the 10 percent flows with the lowest average skill level, were flows where regional returns to skills in terms of wages and employment were higher in the destination compared to the origin region. In contrast, flows in the fifth quintile of selectivity, that is the 10 percent flows with the highest skill level were flows where regional returns to skills were lower in the destination region compared to the origin region. The descriptive results therefore suggest, the higher the interregional employment dispersion, average wage and average employment, the higher is the average skill level of a labor flow. In the following section we will test this hypothesis.

5 Empirical Analysis

Our main attempt is to identify the determinants of a migrants' average skill level. We do so by exploiting the panel dimension of the data that is given by the variation of skill-compositions across 7020 gross labor flows that we observe during the time period 1993-2004. Since an Ordinary Least Squares (OLS) regression would be biases due to unobserved effects such as utility differentials (e.g. amenity differentials) that are very likely correlated with regional wages and employment, we estimate the following Flow-Fixed-Effects model:

$$S_{ijt} \quad = \quad \beta_0 + \beta_1 RETURNS_{ojt} + \beta_2 NATIONAL_t + \nu_{oj} + \epsilon_{ojt}, \tag{15}$$

where $t = 1, ..., 10$ and $k \neq j$. The dependent variable S_{ijt} is the measure of migration flows' average skill level. The term $RETURNS_{kjt}$ is a set of regional covariates that contains all interregional differences in regional returns to skills. The latter contains the average values of both wages and employment as well as the wage and employment dispersion of a region. The regional characteristics are statistically standardized interregional values. A value of one for the interregional mean wage, for example, thus means that the mean wage is one standard deviation higher in the destination compared to the origin region. Furthermore, we control for national values of the regional returns $NATIONAL_t$. The above model thus controls for time-constant utility differentials.

Table 5: Estimating the Average Observable Skill Level of Labor Flows with Flow-Fixed Effects

	(1) Migration Rate (per 1000)	(2) Observable Skills	(3) Observable Skills (relative to region of origin)
Interregional values:			
Mean wage	0.147*	0.006	0.014
	(2.36)	(0.34)	(0.81)
Mean employment	0.174***	0.025***	0.048***
	(15.13)	(7.62)	(14.86)
Wage dispersion	0.062**	0.013	0.030***
	(2.64)	(1.89)	(4.57)
Employment dispersion	0.106***	0.021**	0.024***
	(4.54)	(3.10)	(3.68)
National values:			
National employment	-0.229***	-0.002	-0.009
	(-7.91)	(-0.25)	(-1.16)
National wage	0.034	0.163***	0.002
	(1.93)	(32.30)	(0.43)
National wage dispersion	0.040*	0.100***	-0.005
	(2.46)	(21.72)	(-1.16)
National employment dispersion	-0.317***	0.004	-0.020*
	(-10.89)	(0.52)	(-2.41)
Constant	1.004***	4.735***	0.257***
	(263.82)	(4378.97)	(242.39)
N	7020	7020	7020
R sq.	0.137	0.902	0.051

t-statistics in parentheses
* $p < 0.01$, ** $p < 0.05$, *** $p < 0.001$

Table 5 shows the results for estimations of (1) the migration rate (per 1000 workers in the origin region), (2) the average observable skill level of labor flows as well as (3) the average observable

Table 6: Estimating the Average Overall Skill Level of Labor Flows with Flow-Fixed Effects

	(4) Overall Skills	(5) Overall Skills (relative to region of origin)	(6) Years of education (relative to region of origin)
Interregional values:			
Mean wage	-0.016	0.024*	-0.233**
	(-1.50)	(2.32)	(-3.18)
Mean employment	0.022***	0.010	0.209***
	(4.18)	(1.82)	(5.64)
Wage dispersion	0.004	0.001	0.128***
	(1.02)	(0.15)	(4.44)
Employment dispersion	0.027***	0.025***	0.181***
	(5.31)	(5.05)	(5.12)
National values:			
National employment	-0.020***	-0.014**	-0.019
	(-4.07)	(-2.98)	(-0.55)
National wage	-0.065***	-0.045***	-0.092***
	(-21.43)	(-15.25)	(-4.36)
National wage dispersion	0.021***	0.028***	0.131***
	(7.46)	(10.20)	(6.85)
National employment dispersion	-0.007	-0.004	0.008
	(-1.36)	(-0.87)	(0.22)
Constant	0.013***	0.041***	0.869***
	(20.06)	(63.86)	(193.66)
N	7020	7020	7020
R sq.	0.390	0.114	0.033

t-statistics in parentheses
* $p < 0.01$, ** $p < 0.05$, *** $p < 0.001$

skill level relative to the origin region. The results of model (1) for absolute migration can be interpreted as follows: a one standard unit increase in the interregional mean wage increases absolute migration by 147 migrants. Column one shows that all interregional characteristics have a positive and significant effect on migration levels, with the effect of employment being stronger. The result is in line with Parikh and Leuvensteijn (2002) and Decressin (1994) who find that both unemployment differences and wage differences are important factors in determining migration levels.

Furthermore, model (2) shows that a one standard unit increase in the interregional mean wage increases the average observable skill level, i.e. the predicted log wage level of an average migrant, by 0.025 log points. According to model (2) a higher interregional value of both mean and dispersion of employment increases the average observable skill level of a labor flow, while the wage dispersion has no effect. However, model (3), where the dependent variable is the average observable skill level relative to the region of origin, indicates also a positive effect of in crease in the interregional wage dispersion as Borjas et. al. (1992) suggest.

Table 6 shows estimations of (4) the overall skill level, (2) the overall skill level relative to the region of origin as well as (3) average education years relative to the region of origin. Note that again in all estimations on the employment dispersion has a positive effect on the average skill level. Also, in model (2) and (6) average employment has a positive effect. Thus, the results suggest that a higher interregional difference in the average and dispersion of employment chances increases the average skill level of a migration flow. In contrast, we find only weak evidence for a positive effect on selectivity of a higher interregional wage dispersion. Tables 5 and 6 also shows the effects of national values of both the wage and employment distribution. In 4 out of 5 models a higher national wage dispersion has a positive effect on the average skill level of a migration flow. The reason may be that a higher national wage dispersion may increase the risk of a low-skilled worker that migration pays of. In particular, institutions such as welfare benefits and other protection laws may further decreases incentives for low-skilled workers to migrate in such an environment.

6 Conclusion

This paper examined the determinants of skill-selective internal migration in Germany in light of an extended framework for the self-selection of migrants. The extension hypothesized that self-selection in Germany may be driven by interregional differentials in the dispersion of individual employment chances rather than by the typically suggested interregional differentials in wage dispersion. The analysis is motivated by repeated hints in the existing literature that the scope for regional wage bargaining and thus interregional wage differentiation may be weak in labor markets that are dominated by central wage bargaining.

The findings suggest that, as expected, the average skill level of a migration inflow increases with the dispersion of employment chances in the destination compared to the origin region. Moreover, mean differentials in wages and employment also tend to increase the average skill level of a labor flow, while the wage dispersion has no significant impact. Apparently, the high skill level of an average West-East migrant and the relatively low skill level of an average East-West migrant may partially be attributed to the poorer employment chances for unskilled relative to skilled workers in Eastern compared to Western Germany. Thus, this paper suggests, that the Borjas framework needs to be complemented by a selection mechanism that works via the employment side of the labor market in order better understand the self-selection of internal migrants in context where where wages tend to be rather inflexible at a regional scale. These findings are relevant beyond Germany whenever regional wage rigidities prevent flexible wage adjustments, especially among unskilled workers, and thus generate skill-specific interregional disparities in employment chances.

Besides the theoretical implications, the present work is relevant for the understanding of labor mobility of unskilled workers in countries where increased global competition has forced firms and companies to increasingly make use of flexible employment instruments such as temporary and part-time employment as well as labor leasing. The reason is that unskilled workers are mostly affected by such nonstandard employment relations. The chances that migration pays off for unskilled workers therefore deteriorate and create a barrier to migrate. Therefore, an increasingly instable environment may even worsen the chances of unskilled workers as regional mobility constitutes a chance for workers to find a new or more stable job nationwide in the present of local labor market tightness. Policy makers should take this into account when designing policy instruments to encounter inequalities in light of a transition to a more flexible labor market.

As a further contribution, we are able to better explain the self-selection of migrants between East and West Germany. Looking at migration rates, we find that East-West migrants tend to be unskilled, while West-East migrants tend to be rather skilled. The results somewhat calm down the current debate on the feared brain drain from the Eastern to the Western part of the country. Apparently, East-German regions are not falling behind in attracting human capital into their regions. Moreover, the present analysis sheds light on the determinants of East-West migration selectivity. The results suggest that the pattern of East-West migration may partially be attributed to the higher employment instabilities in Eastern compared to Western Germany. In particular higher shares of nonstandard contracts in East compared to West Germany may explain why predominantly unskilled workers are distracted from East-German regions, while skilled workers seem to be unaffected by a more instable employment environment.

References

Arntz, M. (2010): "What attracts human capital? Understanding the skill composition of interregional job matches in Germany," *Regional Studies*, 44(4), 423 – 441.

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

Borjas, G. J., S. G. Bronars, and S. J. Trejo (1992): "Self-selection and internal migration in the United States," *Journal of Urban Economics*, 32(2), 159–185.

Bruecker, H., and P. Truebswetter (2007): "Do the best go west? An analysis of the self-selection of employed East-West migrants in Germany," *Empirica*, 34(4), 371–395.

Burda, M. C., and J. Hunt (2001): "From Reunification to Economic Integration: Productivity and the Labor Market in Eastern Germany," *Brookings Papers on Economic Activity*, 32(2), 1–92.

Cutillo, A., and C. Ceccarelli (2010): "The internal relocation premium: are migrants positively or negatively selected? Evidence from Italy," *Working Paper Dipartimento di Economia Pubblica, Sapienza Università di Roma, Roma, Italia.*

Decressin, J. W. (1994): "Internal Migration in West Germany and Implications for East-West Salary Convergence," *Weltwirtschaftliches Archiv*, 130(2), 231–256.

Dostie, B., and P. T. Leger (2009): "Self-selection in migration and returns to unobservables," *Journal of Population Economics*, 22, 1005Ũ1024.

Fratesi, U., and M. R. Riggi (2007): "Does Migration Reduce Regional Disparities? The Role of Skill-Selective Flows," *Review of Urban and Regional Development Studies*, 19(1), 78–102.

Gernandt, J., and F. Pfeiffer (2008): "Wage Convergence and Inequality after Unification: (East) Germany in Transition," SOEPpapers 107, DIW Berlin, The German Socio-Economic Panel (SOEP).

Granato, N., A. Haas, S. Hamann, and A. Niebuhr (2009): “Arbeitskräftemobilität in Deutschland - Qualifikationsspezifische Befunde regionaler Wanderungs- und Pendlerströme,” *Raumforschung und Raumordnung*, 67(1), 21–33.

Hunt, G. L., and R. E. Mueller (2004): “North American Migration: Returns to Skill, Border Effects, and Mobility Costs,” *The Review of Economics and Statistics*, 86(4), 988–1007.

Hunt, J. (2000): “Why Do People Still Live in East Germany?,” *IZA Discussion Paper*, (123).

Kanbur, R., and H. Rapoport (2005): “Migration selectivity and the evolution of spatial inequality,” *Journal of Economic Geography*, 5(1), 43–57.

Mauro, P., and A. Spilimbergo (1999): “How do Skilled and the Unskilled Respond to Regional Shocks?,” *IMF Staff Papers*, 46, 1–17.

Mertens, A. (2002): “Regional and Industrial Wage Dynamics in West Germany and the United States,” *Journal of Economics and Statistics*, 222(5), 584–608.

Moeller, J. (1995): “Empirische Entwicklung der Regionalentwicklung,” *Standort und Region*, 24, 197–230.

Moeller, J., and A. Schmillen (2008): “Verteilung von Arbeitslosigkeit im Er-werbsleben. Hohe Konzentration auf wenige Ű steigendes Risiko für all,” *IAB-Kurzbericht*, (24).

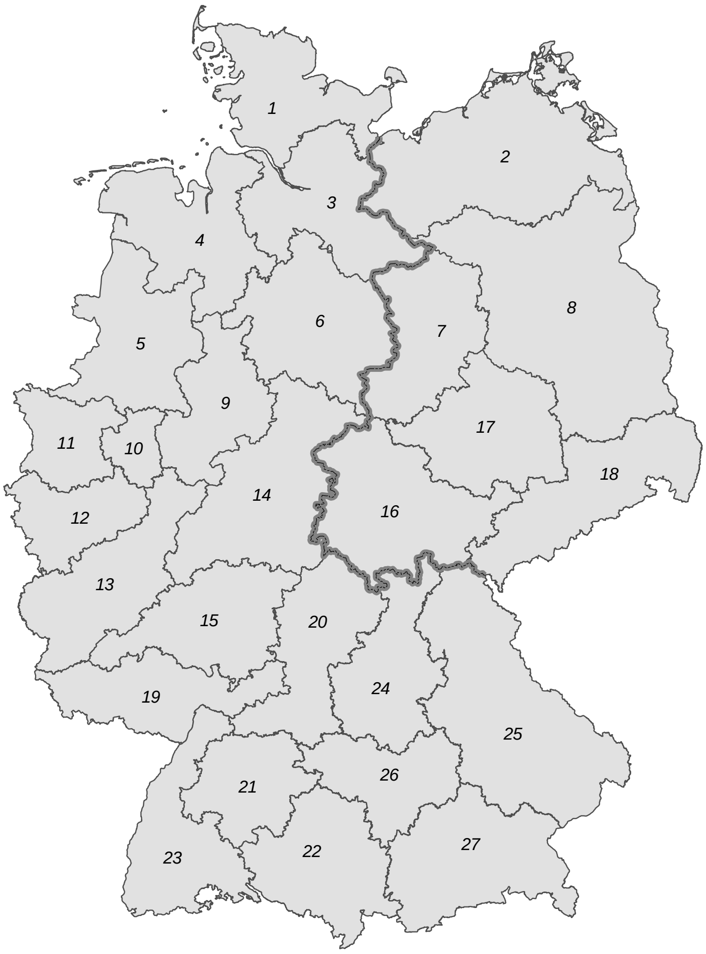
Nickell, S., and B. Bell (1995): “The Collapse in Demand for the Unskilled and Unemployment across the OECD,” *Oxford Review of Economic Policy*, 11(1), 40–62.

Parikh, A., and M. V. Leuvensteijn (2002): “Internal Migration In Regions Of Germany. A Panel Data Analysis,” Discussion paper, Economic Policy Research Institutes.

Roy, A. D. (1951): “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers,*, 3(2), 135–146.

Todaro, M. P. (1969): “A Model for Labor Migration and Urban Unemployment in Less Developed Countries,” *American Economic Review*, 59(1), 138–48.

Figure 1: Map of German Labor Market Regions



Note: The bold line indicates the border line between East and West German Regions.

Table 7: Out-Migration Rates and Average Skills of Out-Migrants

Region	*N* (*in 1000*)	*Rate* (*per 1000*)	*Av. educ.*	*Obs. skills*	*Unobs. skills*	*All skills*
			East German Regions			
2	74	1.11	12.8	4.77	-0.48	-0.15
7	69	1.49	12.65	4.74	-0.46	-0.16
8	213	0.93	13.31	4.67	-0.26	-0.02
16	116	1.2	12.68	4.73	-0.45	-0.16
17	149	1.49	12.63	4.69	-0.41	-0.16
18	138	1	12.8	4.72	-0.42	-0.14
Mean East	759	1.20	12.81	4.72	-0.41	-0.13
			West German Regions			
1	62	0.93	13.41	4.8	-0.36	0
3	164	1.04	13.52	4.68	-0.14	0.1
4	90	0.84	13.04	4.72	-0.3	-0.01
5	104	0.92	13.19	4.74	-0.29	0.01
6	127	0.87	13.56	4.72	-0.21	0.08
9	96	0.8	13.19	4.76	-0.29	0.04
10	157	1.35	13.41	4.71	-0.2	0.08
11	323	1.15	13.53	4.66	-0.11	0.12
12	201	1.08	13.84	4.71	-0.18	0.1
13	81	0.93	13.01	4.74	-0.3	0
14	111	0.95	13.34	4.76	-0.32	0.01
15	258	1.06	13.86	4.7	-0.13	0.13
19	138	0.86	13.61	4.72	-0.22	0.07
20	79	0.96	13.08	4.79	-0.36	0
21	169	0.87	13.64	4.74	-0.2	0.1
22	66	0.75	13.39	4.79	-0.33	0.03
23	110	0.79	13.7	4.76	-0.25	0.08
24	100	0.91	13.57	4.75	-0.23	0.08
25	90	0.76	13.01	4.78	-0.39	-0.04
26	78	1.02	13.12	4.77	-0.3	0.03
27	186	1.02	13.89	4.73	-0.19	0.12
Mean West	133	0.95	13.42	4.74	-0.25	0.05

Table 8: In-Migration Rates and Average Skills of In-Migrants

Region	*N* (in 1000)	*Rate* (per 1000)	*Years of education*	*Observable skills*	*Unobservable skills*	*All skills*
			East German Regions			
2	57	0.84	13.32	4.59	-0.38	-0.1
7	57	1.25	13.06	4.56	-0.37	-0.11
8	182	0.8	13.79	4.61	-0.23	0.02
16	94	0.97	13.01	4.55	-0.37	-0.12
17	119	1.18	13	4.55	-0.35	-0.11
18	117	0.84	13.27	4.56	-0.36	-0.09
Mean East	104	0.98	13.24	4.57	-0.34	-0.09
			West German Regions			
1	62	0.93	13.42	4.78	-0.35	0
3	181	1.15	13.55	4.77	-0.21	0.09
4	93	0.87	13.13	4.72	-0.3	-0.01
5	107	0.95	13.02	4.75	-0.3	0.01
6	131	0.9	13.54	4.74	-0.22	0.08
9	98	0.82	13.07	4.78	-0.29	0.03
10	168	1.45	13.15	4.65	-0.16	0.06
11	310	1.1	13.42	4.73	-0.15	0.1
12	207	1.11	13.61	4.77	-0.22	0.07
13	83	0.95	12.83	4.76	-0.31	0
14	104	0.9	13.25	4.73	-0.3	0
15	283	1.16	13.73	4.81	-0.19	0.1
19	134	0.84	13.49	4.78	-0.25	0.06
20	87	1.05	12.9	4.85	-0.38	-0.01
21	176	0.9	13.46	4.85	-0.27	0.07
22	73	0.82	13.16	4.84	-0.34	0.01
23	120	0.86	13.41	4.81	-0.29	0.05
24	103	0.93	13.42	4.78	-0.24	0.08
25	97	0.82	12.92	4.77	-0.35	-0.02
26	87	1.14	13.09	4.84	-0.33	0.02
27	219	1.2	13.77	4.91	-0.28	0.08
Mean West	139	0.99	13.30	4.78	-0.27	0.04