

Does Discrimination Explain Occupational Segregation by Sex? Evidence from Local Gender Imbalances in Post-Unification Germany

Melanie Arntz
ZEW Mannheim

Holger Bonin
ZEW Mannheim and IZA

Felix Hörisch
ZEW Mannheim

(Preliminary, do not quote without permission)

Abstract

The paper contributes to the discussion whether demand or supply for labour drives occupational segregation by sex. It exploits a demographic phenomenon unique among industrialized nations: massive imbalance in local sex ratios developing through female-dominated regional migration flows in post-unification Germany. Estimates are on the basis of administrative panel data and control for unobserved heterogeneity both at the individual and regional level. The evidence suggests that when men become relatively scarce, the probability of women to access to male-dominated occupations becomes larger. Thus there appears to be a demographic window of opportunity for women, which is consistent with demand theories of occupational segregation. Our hypothesis is that if firms face a relative shortage of men, they find it increasingly difficult to discriminate against women.

Keywords: Occupational Segregation, Discrimination, Gender Ratio, Germany

JEL Classification: J71, J16, R23

1. Introduction

This paper contributes to the discussion whether occupational segregation by sex is a demand side or supply side phenomenon. The empirical analysis exploits large changes in local sex ratios in Germany post unification. As a result of female-dominated East-West migration flows, the dimension of local gender imbalances in the population at employable age has become unique among the industrialized countries. Our empirical estimates suggest that when men become scarce, the probability of women to access to male-dominated occupations becomes larger. Thus there appears to be a demographic window of opportunity for women, which is consistent with demand theories of occupational segregation by sex. Our hypothesis is that if there is a relative shortage of men, firms find it increasingly difficult to discriminate against women.

Occupational segregation by sex appears as an unwanted feature of labour markets for a number of reasons. First of all, the fact that a certain part of the population is excluded from, or at least has difficulties to obtain access to, a certain range of occupations may yield labour market rigidity, and therefore reduce the ability of an economy to adjust to change. Second, occupational segregation is wasteful of human resources. If it drives education choices, certain abilities in the population might remain undeveloped. Finally, it may be directly detrimental to women. It supports perpetuation of gender stereotypes, which may have an adverse effect on many economic and social variables, e.g. poverty and income inequality (Macpherson and Hirsch, 1995).

The literature explaining occupational segregation by sex distinguishes between labour supply and labour demand related factors, see Anker (1997) for an overview. The labour supply side explanation relies on the hypothesis that women generally have a special preference for certain job characteristics that makes them cluster in certain occupations. For example, women may prefer occupations that make it easier to work flexible hours, or to interrupt work around birth.

The labour demand side explanations rely on the hypothesis that employers prefer to hire male and female workers for specific occupations. In part, this may be a reflection of specific job requirements, like education and experience, meeting gender-specific worker characteristics.

However, another possible reason is – statistical – discrimination. If there are average differences in productivity, skills etc. of men and women, and if there are search and information costs associated with hiring and promotion decisions, employers have a rationale to discriminate against women. This requires that sustaining differences between male and female workers costs less than identifying suitable individual workers.

There is rather little empirical evidence on the role of discrimination in explaining occupational segregation. Gill (1989) compares the probability that an individual will choose an occupation and the probability that an individual will be hired for a desired job. His results suggest that occupational segregation by ethnic race can be attributed to discrimination by employers. Beller (1982) provides an indirect test that discrimination impacts occupational segregation, by showing evidence that enforcement of equal opportunity programs boosts women's

probability to be employed in a male occupation relative to men's probability. Reilly et al. (2006) explore establishment level data. Their results strongly suggest that both supply and demand aspects must be considered in explaining gender segregation, but that discrimination effects are not substantial.

In this paper, we provide an indirect test of the discrimination hypothesis by studying the impact of relative scarcity of men (or women) on the local labour market on the propensities that women will work in a "male" occupation, i.e. an occupation with a male employment share exceeding 70 per cent. The idea behind our approach is that a falling sex ratio, defined as the ratio of men to women at employable age, in a given labour market will raise costs for employers to sustain differences between male and female workers, whereas the costs of identifying suitable workers irrespective of gender should not be affected. Therefore, if employers behave discriminatorily, we would expect that the share of women in male-dominated occupations increases as the share of men in the labour market decreases.

The demographic development in Germany post unification allows an empirical test of this hypothesis. It has a unique feature – migration flows between eastern and western Germany have been highly selective with regard to gender. This has led to imbalances in regional sex ratios across Germany of a dimension unique in industrialized countries. We exploit this huge variation across regions and time in order to identify whether a relative shortage of men at the local level changes occupational segregation by gender. In order to exclude supply side factors, we rely on panel data on workers' occupation histories, which allows

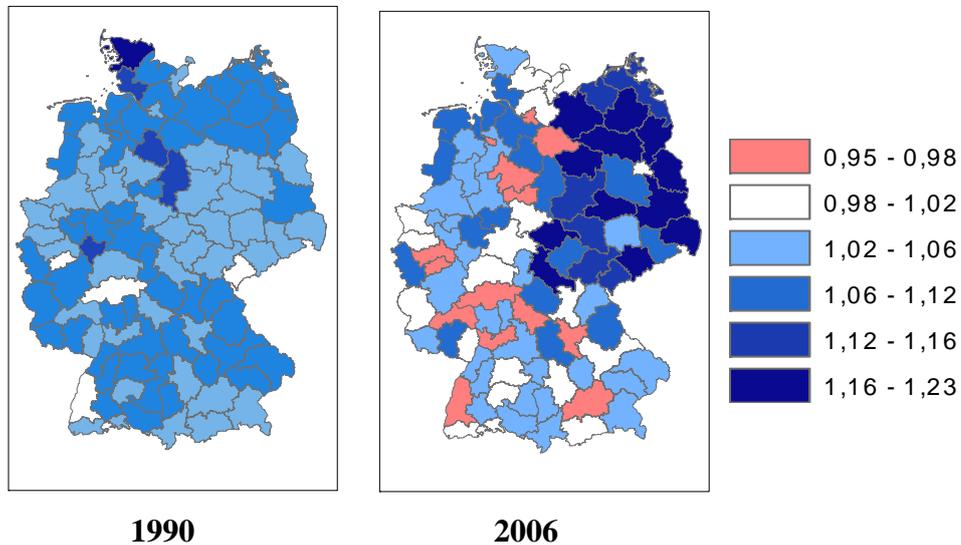
controlling for observed and unobserved heterogeneity at the individual level. At the regional level, in order to reduce the possibility of spurious correlation, we control for a range of observable regional characteristics and unobserved fixed effects.

The remainder of the paper is structured as follows. Section 2 provides the demographic background of gender-selective regional migration flows in post-unification Germany. Section 3 describes our data and empirical approach. Section 4 discusses the estimation results. Section 5 concludes.

2. Gender-selective migration in post-unification Germany

Since re-unification, 1.5 million people left eastern Germany and mainly moved to western Germany (Mai 2006). In particular, net migration from eastern to western Germany peaked around re-unification and has been dropping markedly until the mid 1990s, a trend that has been associated with a rapid wage convergence in the early 1990s (Hunt, 2000). Since the mid 1990s, however, the pace of economic recovery of eastern Germany has been slowing down and east-west migration has again started to rise (Heiland, 2004). At the same time, there have been migration flows within western and eastern Germany, with the rural regions experiencing net migration losses.

Figure 1: Regional Sex Ratio (Age 18 to 30 Years) in Germany 1990-2006



Apart from the reallocation of population, a main feature of these migration flows was their selectivity with regard to age, gender and education. In the main focus of recent research was the selectivity of migration flows with respect to the education and qualification of the migrants and a corresponding brain drain from eastern to western Germany (Hunt, 2004; Brücker and Trübswetter, 2004; Arntz, 2009). What has been mainly missed by the recent literature, however, is its high degree of selectivity with respect to gender. Many more young women leave eastern German for the western part of the country than young men.

Figure 1 shows the sex ratio, i.e. the number of men per women aged 18 to 30, at the level of regional planning units (*Raumordnungsregionen*) at the time of German re-unification and in 2006. In 1990, the sex ratio for individuals aged 18

to 30 years was quite balanced across regions, but showed a slight overhang of men in this age group on average. Only sixteen years later, the picture has changed dramatically. Due to gender-selective migration, the sex ratio in most eastern regions has dramatically increased, while some urban regions in western Germany were able to disproportionately attract women and show a surplus of young women.

The intensity of this process and its result are without comparison in Europe. Even regions along the arctic circle in North Sweden and Finland which have been suffering from a disproportionate migration to the cities among young women for a long time do not reach sex ratios as high as the ones that can be found in eastern Germany (Kröhnert and Klingholz 2007). In an attempt to find an explanation for the selectivity of east-west migration with respect to gender, Schneider and Kubis (2007) examine the relevance of labour market factors, educational reasons, family reasons and amenities in shaping the destination decisions of young men and women. For most of these factors, however, they cannot establish any gender-specific differences. Both young women and young men tend to migrate from eastern to western Germany and from rural to urban regions with good apprenticeship and work opportunities. However, Kubis and Schneider (2007), and Mai (2006) identify gender differences concerning the level of migration. Apparently, young women in Germany exhibit much higher migration rates than their male counterparts. Among those aged 18 to 25, nearly twice as many young females moved between regions. Hence, while the underlying reasons for these different migration levels among men and women

still remain a puzzle, it cannot be stressed enough that the resulting regional imbalance within Germany is a historically unique situation that is without precedent among current industrialized countries. In order to examine the impact of male and female labour supply on occupational segregation, these imbalances thus provide a valuable source of regional variation that we exploit in our empirical approach.

3. Data and Empirical Approach

3.1 Data

We use the IAB employment sample 1975-2004 - regional file (IABS-R04) which is described in detail by Drews (2008). This administrative data set contains information on a 2 % sample of the population working in jobs that are subject to social insurance payments. In particular, we have daily information on employment periods and periods for which the individual received unemployment compensation from the Federal Employment Agency. For the purpose of our analysis, we construct a panel data set based on this spell information by generating one observation per individual and year that contains the employment status, type of job and other individual characteristics as of May 15th of each year.

For each woman in our sample, we construct a binary dependent variable that indicates whether a women works in a male-dominated (=1) or not (=0). We classify occupations to be female-dominated if the share of female employees that we observe in this occupation in Germany in the year preceding our observation

period exceeds 70%. In contrast, occupations with a female workforce of less than 30% are considered as male-dominated occupations.

Occupations can be identified up to the 3-digit level and encompass 132 occupation groups. In line with the literature, we find that women are overcrowded in fewer occupations than men (Sorensen, 1990; Lewis, 1996). Thus, only 29 out of 132 occupations are defined as female-dominated according to our definition, while 66 out of the 132 occupations.

We use the binary indicator of women's employment in male-dominated occupations in order to examine whether the regional imbalance has an impact on women's employment chances in a male-dominated job. In order to give some examples of the jobs that are contained by these definitions, Table 1 shows the ten most female- and male-dominated jobs in Germany at the beginning of the observation period in 1994.

For each employment spell, the IABS-R04 contains information on the NUTS3 region of the workplace, i.e. the county level (*Kreis*). Since individuals need not work and live within the same county, we decided to aggregate the county level information up to functional urban areas, the so called *Raumordnungsregionen*, (ROR), that lump together three to four counties that are closely linked by commuting ties. In terms of the regional hierarchy, the ROR are between the NUTS2 and the NUTS3 level of aggregation and encompass 97 German regions.

Table 1: Typical female and male jobs in Germany

Female-Dominated Jobs			Male-Dominated Jobs		
Rank	Job	Share of female employees	Rank	Job	Share of female employees
1	Receptionist	0.996	1	Plumber	0.002
2	Kindergarten teacher	0.980	2	Bricklayer	0.003
3	Stenographer; shorthand writer	0.966	3	Roofer	0.004
4	Sewer	0.956	4	Concrete Worker	0.004
5	Medical-technical assistant	0.953	5	Motor Mechanic	0.008
6	Home help	0.953	6	Blacksmith; pipe fitter	0.010
7	Hairdresser	0.942	7	Locksmith; shipbuilder	0.011
8	Cleaner	0.938	8	Pavior	0.011
9	Data entry operator	0.935	9	Miners	0.011
10	Cashier; Collector	0.913	10	Factory mechanic	0.012

Based on this regional classification, we merge information on the regional sex ratio and the regional economic structure. The sex ratio is defined as the number of men per women in the working age population between 18 and 60 years of age. In addition, we merge regional information on the share of industrial jobs, the unemployment rate and the population density in order to control for the regional economic conditions that – apart from the sex ratio - might also be

related to occupational segregation. Since a consistent time-series on the regional information is not available before 1995, we restrict our sample to women during the years 1995 and 2004.

Table 2: Descriptive Statistics by type of occupation

	Male-dominated job		Other job	
	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics				
Age	38.9	10.9	38.0	11.5
Age squared	1631.2	871.7	1574.0	907.3
Vocational training ^D	0.538	0.499	0.723	0.448
University degree ^D	0.135	0.342	0.058	0.234
East Germany ^D	0.245	0.430	0.198	0.399
Regional characteristics				
Gender ratio	104.20	2.32	104.06	2.27
Industry share	18.9	4.3	18.9	4.3
Unemployment rate	10.9 %	4.6 %	10.6 %	4.3 %
Population density (residents per km ²)	433.1	445.0	449.8	455.9
Observations	108,785		1,042,930	

Note: ^D indicates that the variable is a dummy. Male dominated jobs are defined as all jobs with a female employment share below 25% in the year 1994. BBR is the Federal Office for Building and Regional Planning in Germany and INKAR its database for regional statistics.

Summary statistics for both the individual and regional level information are shown in Table 2. From the descriptive statistics, only few covariates appear to be related to a woman's type of occupation. What is striking is that a higher

share of women from eastern Germany works in a male-dominated job as compared to women from western Germany. This likely reflects the historically different cultural norms concerning female employment in both parts of the country.

In line with the literature, we also find that women working in a male-dominated job tend to have a higher educational attainment. Among the regional characteristics, only the population density seems to be related to women's occupations. Women in a male-dominated occupation live in regions with a lower population density, thus probably reflecting the higher share of typical female service jobs in denser regions. The gender ratio has no strong relationship to women's occupational choice. However, a multivariate analysis may nevertheless reveal some interesting findings.

3.2 Econometric Approach

We consider a framework in which a woman $i = 1, \dots, N$ in region $r = 1, \dots, R$ can either work in a female-dominated-job in period $t = 1, \dots, T$ or not. Our dependent variable thus is a binary outcome with $y_{itr} = 1$ if a woman works in a female-dominated job in period t and $y_{itr} = 0$ otherwise. Using this indicator, we explore the determinants of being employed in a female-dominated job by modelling the probability of working in a female-dominated job as a function of $k = 1, \dots, K$ explanatory variables x_{kitr} . In general, this model can be written as

$$\Pr[y_{itr} = 1 \mid x_{kitr}] = F(\alpha + x_{kitr}'\beta)$$

with F as a monotone function ranging from 0 to 1, α as an unknown constant, β as a $K \times 1$ vector of unknown coefficients, and x_{itr} as a $K \times 1$ vector. In most applications and textbooks, F is the cumulative logistic or normal distribution function and the models are referred to as logit and probit, respectively. We follow this literature by assuming that the true function F is logistic. The K explanatory variables are a combination of individual- and region-specific characteristics. We estimate the coefficients by means of different methods and model specifications using STATA. In particular, we apply pooled and fixed effects methods.

The pooled model does not exploit the panel data structure, but includes year dummies in order to capture aggregate trends that might affect a women's probability of working in a female-dominated job. This first model specification can thus be written as

$$\Pr[y_{itr} = 1 \mid x_{itr}] = F(\alpha + x_{itr}'\beta + \eta_t)$$

where η_t corresponds to the unknown coefficients for the year dummies. In this model, standard statistics are not valid and need to be corrected due to serial correlation of the individual specific errors over time. While this does not affect consistency, potential correlation between the regressors and unobserved individual specific effects (such as ability or motivation) or unobserved regional heterogeneity does. In particular, we may expect that the regressor of interest, the gender ratio, to be correlated to other regional economic conditions that also affect women's employment chances. While a number of region-specific

covariates are supposed to capture some important regional conditions, we nevertheless augment the previous specification by further including region dummies that capture time-constant unobserved regional effects. This extended model specification with regional dummies can be written as

$$\Pr[y_{itr} = 1 \mid x_{itr}] = F(\alpha + x_{itr}'\beta + \eta_t + \eta_r).$$

where η_r refers to the unknown coefficients for the regional dummies. However, this model still results in biased estimates if there is unobserved heterogeneity at the individual level that is correlated with the regressors. As an example, the gender ratio may be correlated to unobserved individual characteristics such as motivation if the migration flows that result in the unbalanced allocation of women and men across space is selective also with regard to such unobserved characteristics.

In order to remedy these problems, our preferred specification exploits the panel structure of the data and augments the previous model by including individual fixed effects α_i :

$$\Pr[y_{itr} = 1 \mid x_{itr}] = F(\alpha_i + x_{itr}'\beta + \eta_t + \eta_r).$$

In contrast to the linear fixed effects (FE) panel estimator, the logit FE panel estimator uses period data from individuals only for whom the value of the dependent variable switches between two periods. For the observations generated by these individuals, it is essentially a pooled logit estimator with period changes of regressors (Baltagi, 2005). Therefore, similar to the linear FE model, it does not

yield estimates for time constant variables such as gender and it does not reveal any information about the individual fixed effect. Since our covariates of interest are time-varying though, this is of no major concern. Moreover, as its main advantage, the logit FE panel estimation produces consistent estimates even if the individual time constant effect has a non-zero population covariance with the observed regressors.

In the following result section, we therefore present three model specifications. The first includes neither regional dummies nor individual fixed effects, whereas the second includes the former and the last specification takes account of both unobserved time-constant individual and regional heterogeneity. In addition to the described dependent variable, we run the same specifications also for a second dependent variable that indicates whether a woman works in a male-dominated job. This indicator is supposed to yield insights into whether the regional gender ratio affects not only the probability of working in a female-dominated job, but also the probability of entering jobs that are predominantly male. This second indicator thus is a binary outcome with $y_{itr} = 1$ if a woman works in a male-dominated job in period t and $y_{itr} = 0$ otherwise.

4. Results

Table 3 shows the estimates for the three model specifications that have been discussed in the previous section. While specification (1) and (2) are essentially pooled estimates, specification (3) displays estimates for the fixed effects logit estimator and thus uses much less observations than the pooled estimators.

Table 3: Logit Estimates of Probability of Working in Male Job

	(1)	(2)	(3)
<i>Individual characteristics</i>			
Age	1.091***	1.073***	-
Age squared	0.999***	0.999***	-
Vocational training	0.402***	0.395***	0.721***
University degree	2.022***	1.993***	1.576***
<i>Regional characteristics</i>			
Gender ratio	0.962***	0.956***	0.944***
Industry share	1.036***	1.087***	1.089***
Unemployment rate	0.990**	0.973**	0.975*
Population density	1.000	1.003*	1.003
East Germany	2.274***	1.842**	1.371
Year dummies	yes	yes	yes
Regional dummies	no	yes	yes
Individual fixed effects	no	no	yes
# of observations	1,151,715	1,151,715	68,310
# of women	250,292	250,292	12,603
Log-Likelihood	-149,359.3	-149,284.6	-25,988.4

Significance level: *** 1%, ** 5%, * 10%.

Interestingly, the estimates across specifications turn out to be surprisingly robust both concerning the individual as well as concerning the regional characteristics. Among the individual characteristics, women with a university degree are much more likely to work in a male-dominated job than women with

lower educational degrees, while women with a vocational training are less likely to work in a male-dominated profession than unskilled women.

This pattern is robust with respect to controlling for unobserved individual heterogeneity. Thus, even for the same woman, obtaining a university degree increases the likelihood of working in a male-dominated occupation. For the pooled estimates, there is additional evidence that the likelihood of working in a male-dominated job increases with age albeit at a diminishing rate.

The estimation results for the regional characteristics clearly demonstrate that demand-side factors have an important impact on women's employment outcomes. Women in regions with a relatively high share of industrial jobs are more likely to work in a male-dominated occupation even when controlling for both regional and individual heterogeneity. For a given ratio of women and men in the local workforce, a higher demand for industrial occupations thus seems to open up chances for women to enter typical male occupations.

In contrast, women's probability of working in a male-dominated job diminishes significantly with an increasing unemployment rate. This suggests that an excess supply of labour makes it easier for firms to discriminate against women and to slam the door on women.

The major finding in Table 3, however, concerns the effect of the local sex ratio. Irrespective of the specification, we find evidence in favour of a demand-driven discrimination against women. The higher the sex ratio, i.e. the more men outnumber women, the lower is the probability that women work in a male-

dominated job. Put differently, access to male-dominated professions depends on demand-side conditions in the sense that a firm's ability to discriminate against women diminishes if men – in a given regional economic context with regard to unemployment, industry structure, population density and other time-constant regional characteristics - become scarce.

While this finding for specification (2) might still be driven by unobserved individual heterogeneity due to a selection of certain individuals in particular markets, the fact that this result also holds when including individual fixed effects emphasizes the robustness of our findings.

Against these findings, however, one might put forward that these results are driven by differences between eastern and western Germany that are not adequately modelled by the inclusion of a few controls and regional dummies. Since both parts of Germany, albeit increasingly integrated, have experienced different economic shocks during the past years that may have affected both the degree of occupational segregation as well as the selectivity of migration flows with respect to gender, such unobserved shocks could bias our estimates. We therefore run separate estimations for model specification (3) in order to test whether similar conclusions can be drawn without exploiting the variation across both parts of the country.

Table 4: Logit Estimates of Probability of Working in Male Job by Region

	East Germany only (4)	West Germany only (5)
Individual characteristics		
Vocational training	0.866**	0.687***
University degree	2.227***	1.300***
Regional characteristics		
Gender ratio	0.855***	0.947
Industry share	1.086**	1.088***
Unemployment rate	0.975	0.966
Population density	0.992	1.004
Number of observations	15,618	50,466
Number of persons	3,090	9,162
Log-Likelihood	-5,859.9	-19,272.7

Note: Estimates include year and regional dummies.
Significance level: *** 1%, ** 5%, * 10%.

Table 4 shows the corresponding estimates for western (4) and eastern Germany (5). While the individual characteristics appear to be robust to the split of our estimation sample, some of the regional characteristics turn insignificant. In particular, the gender ratio significantly affects women's employment chances in eastern Germany only, while the corresponding estimate for western Germany misses significance (p-value: 0.147). Therefore, the strength of our previous findings to some extent depends on exploiting the variation across both parts of Germany. Nevertheless, the finding does not completely disappear when looking at western and eastern Germany separately.

5. Conclusions

The demographic development in Germany post unification has a unique feature. Migration flows between East and West Germany have been highly selective with regard to gender, thus resulting in a historical imbalance in regional sex ratios across the country that is unique among industrialized countries. This regional variation can be exploited in order to examine whether the relative shortage of men affects women's position in the labour market.

This paper focuses on a special gender aspect in the labour market – occupational segregation by sex. Segregation can be attributed to supply or demand factors. As we use individual panel data, we can control for individual fixed effects, i.e., individual workers preferring specific occupations. Thus our empirical analysis controls supply side factors and provides indirect evidence on the demand side dimension of occupational segregation.

Our results indicate that the probability of being employed in a male-dominated job significantly increases if men are relatively scarce. This result is robust, even if one controls for unobserved heterogeneity both at the individual and regional level. Thus there appears to be a demographic window of opportunity for women, which is consistent with the statistical discrimination theory of occupational segregation by sex. It appears increasingly costly for employers to exclude women from certain occupations, if there is a relative shortage of men.

References

- Anker, R. (1997), Theories of Occupational Segregation by Sex: An Overview, *International Labour Review* Vol. 136, 315-339.
- Arntz, M. (forthcoming), What attracts Human Capital? Understanding the Skill Composition of Interregional Job Matches in Germany, *Regional Studies*.
- Baltagi, B. H. (2005), *Econometric Analysis of Panel Data*, 3rd Edition, Wiley.
- Beller, A. H. (1982), Occupational Segregation by Sex: Determinants and Changes, *Journal of Human Resources* Vol. 17, 371-392.
- Brücker, H. and P. Trübswetter (2004), *Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany*, IZA Discussion Paper No. 986, Bonn.
- Drews, N. (2008), *Das Regionalfile der IAB-Beschäftigtenstichprobe 1975-2004*, FDZ Methodenreport No. 02/2008, IAB Nürnberg.
- Gill, A. (1989), The Role of Discrimination in Determining Occupational Structure, *Industrial and Labor Relations Review* Vol. 42, 610-623.
- Heiland, F. (2004), Trends in East-West German Migration from 1989 to 2002, *Demographic Research* Vol. 11, 173–194.
- Hunt, J. (2000), *Why do People Still Live in East Germany?*, DIW Discussion Paper No. 201, Berlin.
- Hunt, J. (2004), *Are Migrants More Skilled than Non-Migrants? Repeat, Return and Same-Employer Migrants*, DIW Discussion Paper No. 422, Berlin.
- Kröhnert, S. and R. Klingholz (2007), *Not am Mann. Von Helden der Arbeit zur neuen Unterschicht?*, Berlin-Institut für Bevölkerung und Entwicklung, Berlin.
- Lewis, D.E. (1996), Occupational crowding, *The Economic Record* Vol. 72, 107-117.
- Macpherson, D. A. and B. T. Hirsch (1995), Wages and Gender Composition: Why do Women's Jobs Pay Less? *Journal of Labor Economics* Vol. 13, 426-471.

- Mai, R. (2006), Age-selective out-migration from Eastern Germany, *Raumforschung und Raumordnung* 5/2006, 355-369.
- Reilly, Kevin T., Jaume Garcia, Pedro J. Hernandez, Angel Lopez-Nicolas and Luisa Zanchi (2006), *The Why of More or Less: Evidence from Sapin on Gender Segregation at the Establishment Level of the Firm*, Mimeo.
- Schneider, L. and A. Kubis (2007), *Are there gender-specific preferences for location factors? A Grouped Conditional Logit-Model of interregional migration flows in Germany*, Discussion Paper No. 12, Halle Institute for Economic Research.
- Sorensen, E. (1990), The Crowding Hypothesis and Comparable Worth, *Journal of Human Resources* Vol. 25, 55-89.