Institut für Wirtschaftsforschung Halle
Halle Institute for Economic Research

Are there Gender-specific Preferences for Location Factors? A Grouped Conditional Logit-Model of Interregional Migration Flows in Germany

Lutz Schneider, Alexander Kubis

February 2009 No. 5

IWH-Diskussionspapiere
IWH-Discussion Papers

Are there Gender-specific Preferences for Location Factors? A Grouped Conditional Logit-Model of Interregional Migration Flows in Germany

Lutz Schneider, Alexander Kubis

February 2009 No. 5

Author: *Lutz Schneider*
Department: Department of Structural Economics, Halle Institute for Economic Research
Lutz.Schneider@iwh-halle.de
Tel.: (0345) 77 53-857

Author: *Alexander Kubis*
Department: Department of Structural Economics, Halle Institute for Economic Research
Alexander.Kubis@iwh-halle.de
Tel.: (0345) 77 53-851

Herausgeber:
INSTITUT FÜR WIRTSCHAFTSFORSCHUNG HALLE – IWH
Prof. Dr. Ulrich Blum (Präsident), Dr. Hubert Gabrisch (Forschungsdirektor)
Das IWH ist Mitglied der Leibniz-Gemeinschaft

Hausanschrift: Kleine Märkerstraße 8, 06108 Halle (Saale)
Postanschrift: Postfach 11 03 61, 06017 Halle (Saale)
Telefon: (0345) 77 53-60
Telefax: (0345) 77 53-8 20
Internetadresse: http://www.iwh-halle.de

Are there Gender-specific Preferences for Location Factors? A Grouped Conditional Logit-Model of Interregional Migration Flows in Germany

Abstract

The article analyses the question whether women and men differ in their tastes for location factors. The question is answered by quantifying the impact of location characteristics on interregional migration flows across Germany. The analysis is based on a grouped conditional logit approach. We augment the framework by controlling for violation of the independence of irrelevant alternatives assumption and for overdispersion. As a result, we find no differences in terms of direction of impact. However, the regressions confirm gender differences in terms of intensity, particularly regarding regional wage levels and the availability of educational institutions.

Key words: Labour Mobility; Gender Economics; Regional Migration; Discrete Choice Model

JEL classification: C25; J61; J16; R23

Are there Gender-specific Preferences for Location Factors? A Grouped Conditional Logit-Model of Interregional Migration Flows in Germany

Zusammenfassung

Die vorliegende Analyse untersucht die regionalen Determinanten geschlechtsspezifischer Binnenwanderungsflüsse und fragt, ob bestimmte Ausstattungsmerkmale eher von Frauen als von Männern geschätzt werden. Die Untersuchung basiert auf einem Grouped-Conditional-Logit-Ansatz, wobei die Probleme der Unabhängigkeit von irrelevanten Alternativen und der Überdispersion berücksichtigt werden. Im Ergebnis zeigt sich zunächst, dass bei Frauen und Männern dieselben Größen als Pull- oder aber als Push-Faktoren wirken. Geschlechtsspezifische Unterschiede werden indes hinsichtlich Stärke des Einflusses sichtbar, vor allem im Hinblick auf das regionale Lohnniveau und das Angebot an Bildungseinrichtungen.

Schlüsselwörter: Arbeitsmarktmobilität; Geschlechtsdifferenz; Binnenwanderung; Discrete-Choice-Modell

JEL-Klassifikation: C25; J61; J16; R23

Are there Gender-specific Preferences for Location Factors? A Grouped Conditional Logit-Model of Interregional Migration Flows in Germany

1 Introduction

Females are more migratory than males. The "seventh law of migration" was noted by Ernest George Ravenstein at the beginning of statistical migration analysis (Ravenstein 1885). Since then, migration research has widely neglected the gender dimension of spatial mobility – especially with respect to interregional migration. While the empirical literature dealing with international migration recently turned to gender issues (Dumont et al. 2007), only a few studies can be found that explicitly focus on gender-specific internal migration patterns and determinants.

This research gap is surprising since differing mobility rates usually translate into imbalances between regional in- and out-migration and, thus, into unequal regional net migration rates in terms of gender. In the long run regions mostly left by (or attracting) one sex will show substantially unbalanced sex ratios. Regarding younger age groups such trends can be observed in many countries in particular within transition regions. A lot of rural low growth areas face strong female out-migration followed by remarkable disequilibria in terms of sex ratios (Rees and Kupiszewski 1999). Due to the regional dualism between Eastern and Western regions this trend is particularly prevalent in Germany. Against this background our study tries to answer two questions:

i. Are women more mobile than men?
ii. Which regional location factors are women and men attracted by, i.e. what are the gender specific valuations of regional characteristics?

By answering these questions we focus on young adults aged 18 to 30 years. The main reason to restrict the analysis to this group is the distinction between household and individual migration. Between 18 to 30 years the people usually move as individuals therefore we do not expect a high share of "tied movers". Tied movers within households are problematic since their migration decision does not reveal *their* location preferences but the preferences of the person – usually the partner – benefiting from migration. Unfortunately, our data set does not allow distinguishing between household and individual migration, so we have to choose the option of analyzing only the young adults. However, this restriction is far from being critical since this age group is responsible for the largest part of geographic mobility.

The remainder of the paper is organized as follows: the next section outlines the recent empirical literature dealing with internal migration. Section three describes the econometric model, introduces the explanatory variables and characterizes the data set. Estimation results are presented in section four. A conclusion completes the paper.

2 Empirical Literature

The empirical literature analyzing the determinants of interregional migration can be divided into micro level and aggregate approaches. The micro concepts focus on the migration decision or intention of individuals or households whereas the interregional migration flows, i.e. the outcome of individuals' behaviour, are explained in aggregate level studies (Cushing and Poot 2004). Due to the improved availability of survey data during the last decades the literature more and more turned to the micro concepts as on this level the actual migration motives of individuals or households can be captured most closely (Cushing and Poot 2004). However, these survey based studies generally rely on a limited sample size so the impact of location factors at a small scale regional level cannot be examined. Thus, with respect to internal migration in Germany these micro studies usually focus only on the migration from East to West Germany (Burda 1993; Burda et al. 1998; Hunt 2000; Bruecker and Truebswetter 2007). Solely Hunt (2004) analyzes the determinants of regional migration on a smaller regional scale.

Due to our focus on the small scale regional dimension of internal migration, the paper is closer related to the empirical literature dealing with migration on the aggregate than on the micro level. The empirical literature analyzing the regional determinants of aggregate migration flows was traditionally focussed on the role of regional labour market conditions (Greenwood 1997). From a human capital perspective (Sjaastad 1962; Harris and Todaro 1970) the regional wage level as well as the unemployment rate were supposed to affect the regional migration balance. The third variable usually implemented in this type of analysis is distance. It is referred to as proxy for migration costs, thus, it should discourage migration. Many country studies confirm the positive effect of income levels and the negative impact of distance. In terms of regional unemployment rates the results are rather mixed.[1] For Germany the results are very similar (Burda and Hunt 2001; Parikh and Van Leuvensteijn 2003; Arntz 2006).

Regarding the topic of *gender-specific* migration patterns the literature is rare. A lot of micro studies implement a gender dummy; a lot of macro analysis calculates gender-specific mobility rates. But explanations of gender mobility gaps and analysis of gender-specific migration determinants are almost never provided so far. A remarkable exception is the work of Detang-Dessendre and Molho (2000) as well as the analysis of Faggian et al. (2007). The French authors investigate the migration patterns of young women in rural France after completing their education. They conclude that women might be more migratory than man since the woman usually moves to the man when a couple is formed. Women seem to be tied movers, in the sense of Mincer (1978). They move to regions where the male partner maximizes his income. Faggian et al. (2007) explore migration behaviour of university graduates in the UK. They also found higher

1 For the US see *Davies et al.* (2001) as well as *Cebula and Alexander* (2006); for Europe see *Huber* (2004); *Fidrmuc* (2004) as well as *Andrienko and Guriev* (2004).

mobility rates of young women. However, the higher mobility rates are not explained by partnership motives but by labour market factors. Faggian et al. derive the higher mobility of women in UK from the fact that migration is used as compensation mechanism for discrimination in the labour market. So, not only partnership consideration but also regional labour markets might be relevant for female migration decisions.

Aside from these two micro studies no econometric analysis can be found dealing with the gender-specific determinants of internal migration on the aggregate level. Our approach to identify gender-specific valuations of location factors and, thus, to explain the gender patterns of internal migration flows seems to be quite new and nonetheless of vital interest. We expand the framework of Davies et al. (2001) by implementing the gender issue, controlling for violations of basic assumption and adding explanatory variables. The last aspect is especially crucial since we do not solely focus on labour mobility but on other types of migration too.

3 Empirical Approach

3.1 The Model

Our empirical analysis is based on the micro econometric approach of McFadden (1974) known as Conditional Logit Model. Within the framework of a random utility model a probability function is derived which represents the likelihood of rational agents to choose a certain discrete alternative. As we will see, the concept is transferable to aggregate data of migration flows between regions which substantially lowers the computational effort. Otherwise estimations for millions of individuals and a choice set consisting of 439 alternatives (regions) could hardly be done.[2] Our analysis focuses solely on regional characteristics as determinants of gender-specific migration. Since our data set does not contain information on individuals beyond age and gender we abstract from additional individual attributes. Thus our estimation assumes homogenous agents at the regional level, i.e. the groups – men as well as women – only differ with respect to their origin region.

To describe the applied concept in more detail, we have to consider a one stage decision of individual *i* between *J* alternatives. The decision is categorized as one stage process since choice set *J* also contains the source region. In other words, the individual simultaneously decides if she moves and where she moves. Staying in the source region is no qualitatively different phenomenon than moving to a different destination. Strictly speaking, the model assumes that the decision to stay can be seen as decision to move from the source to the source region.

Utility *U* of a (representative) individual *i* moving to region *j* is given by the equation:

$$U_{ij} = \beta' X_{ij} + \varepsilon_{ij} \qquad i \in N, j \in J \tag{1}$$

Vector *X* contains attributes of destination *j* as well as attributes of individual *i*. All individuals face the same choice set *J*. According to the rationality condition, an individual chooses the region that maximizes her utility. Thus, the probability to move to region *j* is given by:

$$P(c_i = j) = P\left(U_{ij} > U_{ik}\right) \qquad \forall k \neq j \tag{2}$$

Given the statistical properties described in McFadden (1974) the probability of individual *i* to move to region *j* can be expressed as:

$$P(c_i = j) = p_{ij} = \frac{e^{\beta' x_j}}{\sum_{j=1}^{J} e^{\beta' x_j}} \tag{3}$$

2 Alternatively, we could estimate a count data model to overcome computational problems. However, these models lack a sound micro-economic foundation and, thus, a straightforward interpretation of coefficients.

According to Guimaraes and Lindrooth (2007), an indicator variable d_{ij} is defined which is set to one if individual i chooses option j or zero otherwise. Then, the likelihood function of the migration decision is given by:

$$L = \prod_{i=1}^{N} \prod_{j=1}^{J} p_{ij}^{d_{ij}} \tag{4}$$

In the case of grouped data, the likelihood can be concentrated if a group of individuals i can be treated equally and the choice set is the same for all individuals. The so called Grouped conditional logit model is formulated as (Guimaraes and Lindrooth 2007):

$$L = \prod_{g=1}^{G} \prod_{j=1}^{J} p_{gj}^{n_{gj}} \tag{5}$$

The exponent n_{gj} represents the number of individuals belonging to group g choosing region j. The probability p to move to region j depends solely on destination attributes of region j and group characteristics g. Individual heterogeneity of the members of a group is neglected in the model. The utility of individual i in group g deciding on destination j is given by:

$$U_{igj} = \beta' X_{gj} + v_{igj} \tag{6}$$

In our analysis we assign individuals to groups on the basis of their origin region. Therefore we obtain 439 groups and 439 potential choices. Thus, the data set consists of a 439x439 matrix. The log-likelihood function is:

$$ln\, L = \sum_{g=1}^{439} \sum_{j=1}^{439} n_{gj}\, ln \frac{e^{\beta' x_{gj}}}{\sum_{j=1}^{439} e^{\beta' x_{gj}}} \tag{7}$$

The factor n_{gj} refers to the number of individuals moving from region g to region j. Since the groups are generated according to the origin region, choice probabilities are solely determined by regional attributes captured by the vector X. Coefficient β can be interpreted as implicit price of the corresponding attribute X (Maddala 1983).

How is the gender dimension treated in the model? The simplest option is to realize two separate estimations for women and men. The drawback of this procedure is its failure to perform statistically proven tests of gender-specific differences. Therefore, we applied a dummy approach. The male and female dataset was combined, the dummy variable indicates whether the observation stems from the male or female part of the data set (zero = male, one = female). In the male part of the dataset the number of male migrants between to regions represents the left-hand side variable; in the female part the number of female migrants is used. The explanatory variables are duplicated – the duplicated part is multiplied with the gender dummy and contains the female values of a variable. Then, the estimated coefficients firstly show the pure male effect of a certain variable and, secondly, the female difference. The female effect itself can be calculated by the sum of the male and the difference coefficient of a variable. So, a special test of gender-specific differences is not necessary since the significance level of the difference coefficient itself provides this information.

3.2 Non-migrants, Overdispersion and IIA

Despite its sound microeconomic foundation the GCL model has some shortcomings: i) In the CGL approach the non-moving option is regarded as equal to the alternative to move to any of the 438 remaining destinations. ii) Due to unobserved group specific heterogeneity a correlation between decisions of group members might deflate the variance-covariance matrix and inflate z-statistics – a problem referred to as overdispersion (Giumaraes and Lindrooth 2007). iii) The model implies the independence of irrelevant alternatives (IIA), i.e. relative choice probabilities between two options are independent from existence and characteristics of other options (Maddala 1983; Dahlberg and Eklöf 2003).

i) Non-migrants. As Davies et al. (2001) argue, there might exist unobserved (fixed) costs of moving leading to a qualitative difference between migration and non-migration. As we estimate a one stage model this difference could bias the results. We follow Davies et al. who implement a dummy variable that indicates, if source and destination region are identical. This dummy variable captures the effect of non-moving. A large and statistically high significant parameter is expected.

ii) Overdispersion. To avoid overdispersion caused by unobserved group-specific effects Guimaraes and Lindrooth (2007) propose the implementation of a random variable capturing the ignored group heterogeneity. The modified utility equation (6) is:

$$U_{igj} = \beta' X_{gj} + \mu_{gj} + v_{igj} \tag{8}$$

The random effect μ is supposed to be gamma distributed with parameters ($\delta_g^{-1}\lambda_{gj}$,$\delta_g^{-1}\lambda_{gj}$) where δ_g represents a group-specific parameter. The authors show that choice probabilities p derived from (8) follow a Dirichlet distribution. The model can be estimated by ML-technique; the Likelihood function follows a Dirichlet-Multinomial multivariate distribution. Guimaraes and Lindrooth propose different methods to parameterize the random variable. We chose the option to treat δ_g as constant.[3]

iii) Independence of irrelevant alternatives. The IIA assumption of the GCL model is rather idealistic in an interregional migration context where a lot of destinations are undistinguishable from the individual perspective (Cushing and Cushing 2007). The weakness can be remedied at least partly if the lack of IIA is seen as omitted variable problem (Guimaraes et al. 2004). Then, the inclusion of an additional variable γ measuring the unobserved heterogeneity of every destination regions avoids estimation biases. Unfortunately, the implementation of a fixed effects variable measuring the unobserved heterogeneity of 439 destination regions is technically unfeasible. Instead of using dummies for every district, we construct dummy variables which aggregate regions showing huge similarities in terms of migration flows. Firstly, we distinguish between

3 The alternative procedure of determining the coefficient of correlation within the groups was also performed. The estimation results do not change.

east and west districts and, secondly, between seven types representing the regional planning category of a district.[4] Then, the IIA property is implied only between regions within these categories but not between them – a rather realistic implication.

3.3 Explanatory Variables

The log likelihood function has to be maximized with respect to the parameter vector β which measures the implicit prices of the choice-specific attributes $x \in X$. We model regional attributes as origin-destination relation. Therefore, the estimation does not include separate variables for origin and destination but only ratios or differences between them. With respect to these origin-destination-specific characteristics we distinguish between four groups of factors which we believe drive the migration behaviour of young adults: i) labour market, ii) education, iii) family and iv) amenities.

i) Labour market. As shown in the literature review, the labour market conditions are part of nearly all studies analyzing the determinants of interregional migration flows. We follow these approaches and implement the regional *average wage levels* – computed as gross wages per employee – and the *unemployment rates* in our estimation. Whereas unemployment rates can be disaggregated in terms of gender, we are not able to calculate gender-specific regional wage levels. However, it can be presumed that these gender wage differences only vary slightly between regions. To cover not only nominal income differences the implementation of the regional price level is necessary. Since appropriate regional price level data do not exist we include the building *land prices*. This variable represents at least partly regional rents which seem to be the main source for purchasing power disparities.

ii) Education. Different age groups seem to have specific needs as well as specific ties driving residential mobility. Since our analysis focuses on the age group of 18-30 years educational migration motives should be of particular relevance. Young adults of about 20 years usually start their tertiary education and might choose a location depending on its educational institutions. In our analysis, migration effects of third level education as well as of vocational training are considered. The first aspect is reflected by gender-specific *college availability*, defined as the share of students per high school graduates. It is a measure for the regional capacity to absorb school graduates by the regional academic opportunities. The effects of vocational training are implemented by *vocational training* availability which represents the number of vacancies and mediated positions per person seeking for vocational training. Beyond these considerations, a further life cycle migration motive is related to the transition from education to employment. This transition typically happens until the age of 30 years. Spatial mobility seems to be a crucial requirement to find adequate jobs for university graduates whereas person with vocational training usually stay within their firm after completing their education. Thus,

4 These types are taken from a classification of the Federal Office for Building and Regional Planning distinguishing basically between agglomerations, urbanized areas and peripheral regions.

the regional labour market capacity to absorb university graduates is implemented as explanatory variable. The capacity is measured as *employment gap*, i.e. as gender-specific ratio of students per high qualified employees.

iii) Family. Despite our focus on young adults, family issues, particularly child care, might already be of some relevance. We presume two exemplary concepts: Either, a reconciliation of work and family is aimed at, i.e. both partners have a job and externalize significant parts of child care. Or tasks are split in employment and household production which means that the sole earner has to generate necessary monetary resources.[5] Egalitarian families should be attracted by regions with appropriate child care facilities. Traditional families have to maximize the income of the sole earner but are not affected by external child care. In addition a mixed family model is supposed to appreciates an adequate offer of part time jobs. To take these aspects into account we implement i) the number of *child care* places per children under six years and ii) the share of gender-specific *part time jobs*.

iv) Amenities. Due to their impact on life quality, the utility of residential choice is also affected by natural and cultural amenities. The cultural endowment is considered via the number of *concert halls*. Even if concerts are only a small part of cultural life it seems to be an appropriate proxy for the entire culture of a region. The extent of *park areas* as well as *near-nature areas* – defined in square kilometre per inhabitant – measures the natural component of amenities.

In addition to the variables belonging to the four basic categories we employ two structural variables usually implemented in migration estimations: *distance* and *gender-specific population ratio* between source and destination region. Distance acts as proxy for migration costs. The variable is defined as the time in minutes required when going by car from the administrative centre of the source to the administrative centre of the destination district. Since the impact of distance on migration is usually found to be non-linear a second order term is implemented. Because there might exist a structural break between short-distance and long-distance moves – short-distance moves might only affect residence and not jobs –, we estimate a separate regression where only long-distance moves are considered as migration. The frontier between short- and long-distance moves is set to 75 minutes travelling time since a duration below 75 minutes is officially regarded as reasonable daily commuting distance (Section 121 (4) German Social Security Code Book III). The population variable measures the potential stock of in- and out-migrants. A more populous region is supposed to attract more young adults and vice versa more young adults should leave such a region. The effect on net migration is ambiguous.

5 *Juerges* (2006) shows the relevance of the distinction between sole earner ("traditional") and double income ("egalitarian") couples for their migrations decisions. Furthermore, *Zaiceva* (2007) provides evidence that women reduce their work supply after migration but do not give it up entirely in case of migration in East-West-direction.

3.4 Data

In our analysis, we use the migration data set for 2005 stemming from the migration statistics of the Federal Statistical Office. The data are based on the official register of residence and comprise all residential movements across district borders within Germany. It enables us to analyze migration flows at the small scale level of NUTS-3 regions. Since the migration data are laid down as a 439 origin-destination matrix, we know where the migrants come from and where they go. An individual is classified as migrant if she transfers her first residence from one NUTS-3 region to another during the year 2005. With respect to age and gender the data set differentiates between migration flows of men and women as well as certain age groups. Since our analysis focuses on young adults, we explore migration flow data for individuals aged from 18 to 30 years. A drawback of the official register of residence is its lack of information about crucial individual attributes, e.g. the educational status of movers cannot be observed. So, our analysis mainly has to focus on regional characteristics. We have to abstract from the impact of individual attributes as well as interactions between individual and regional level.

To avoid endogeneity bias, the explanatory variables in general refer to the year 2004. They are taken from different sources. Regional wage levels come from the German National Accounting of the Federal States (VGR der Länder); unemployment rates are provided by the German Federal Employment Office. The distance variable measuring the travelling time between two regions is computed in ArcGIS on the basis of a detailed German road map. The educational variables, child care availability, and amenity variables are taken from the INKAR statistics of the Federal Office for Building and Regional Planning (BBR). Information on gender-specific part-time- and high-qualification jobs within a region stem from a comprehensive data set of the German Federal Employment Office. These data contain records for every employee registered in National Security System, i.e. for approximately 2/3 of total employment.[6] Every record includes information on employee's job location, gender, employment status (part-time/full-time) and qualification. So, this data gives a very detailed description of regional labour markets. Table one presents a short illustration of all explanatory variables. Note that summary statistics refer to values of NUTS-3 regions while the estimation uses the computed ratios or differences between source and destination region.

6 Self-employed workers, civil servants and people working in liberal professions (e.g. lawyer, doctors, and artists) are not covered by the data. However, there should be a high correlation between the characteristics of the included and not-included workforce. Furthermore, this issue concerns only few and less relevant industries so our analysis should be widely unaffected by this problem.

Table 1
Description of explanatory variables [a]

Variable	Description	Men			Women		
		Mean	Min	Max	Mean	Min	Max
Gender specific variables							
Unemployment rate ♀♂	Gender-specific unemployment rate (in per cent)	13.53	4.22	30.70	13.14	4.64	29.19
College availability ♀♂	Gender-specific number of students per high school graduates	3.44	0.00	31.85	1.32	0.00	14.71
Employment gap ♀♂	Gender-specific number of students per high qualified job [b]	0.23	0.00	4.05	0.33	0.00	4.38
Part-time jobs ♀♂	Gender-specific number of high qualified part-time jobs per high qualified jobs [b]	0.01	0.00	0.08	0.05	0.02	0.16
Population ♀♂	Number of gender-specific population aged 18 to 30 years	13.46	2.60	273.99	13.05	2.42	277.41
Gender unspecific variables		Mean		Min		Max	
Average wage level	Average regional gross wage per employee and year (in Euro)	24.61		18.25		37.93	
Land price	Price of building land (Euro/m²)	94.08		5.41		886.21	
Vocational training	Number of per persons looking for training per offered training positions (in percent)	94.98		80.00		104.30	
Child care	Number of kindergarten places per children under six years (in percent)	73.1		40.4		139.9	
Near-nature area	Near-nature area (m² per inhabitant)	51.73		2		864	
Park area	Recreation area (m² per inhabitant)	43.81		7		222	
Concert halls	Number of concert halls and opera	0.24		0		4	
Distance	Distance between centres of two regions required when travelling by car (in minutes)	261.3		0		703.3	

[a] In the estimations, the variables are implemented as origin-destination relation. Instead, table values refer to values of region itself not to the relation of regions. Otherwise an adequate interpretation would be difficult. – [b] High qualification jobs are defined as jobs filled by high qualified employees = employees with academic degree.

Source: Own calculation.

4 Results

Tables two and three present the estimation results of the GCL method controlling for overdispersion.[7] Table two contains the results for the entire migration flows of the 18-30 years old between German districts whereas in table three the short-distance migration (below 75 minutes driving time) is filtered out.

If we firstly turn to table two, the estimations are not in favour of gender-specific impacts of location factors in terms of sign. Focusing on the labour market, high wages and low unemployment attract young men *and* young women. Also the price level variable shows the same sign for males and females – but in a surprising direction: young adults move to regions with high land prices. Since individuals care about real wages, this unexpected sign can be consistent with utility maximization when price levels are high where wages are high. Then, the wage effect might overweigh higher price levels and, consequently, individuals move to high price level regions. Regarding the educational motives the expected signs predominate. Young adults are attracted by adequate facilities for college education and vocational training. Likewise, the absorption capacity of regional labour markets for university graduates stimulates in-migration. An unpredicted effect has to be noticed on the family field: Child care facilities have negative impacts on net migration. One reason might be our focus on the age group at this age child care considerations are of no crucial interest for this group. Then, the negative sign results if there is a substantial negative correlation between child care facilities on the one hand and labour market conditions as well as educational institutions on the other hand.[8] With respect to amenities the estimated coefficients confirm the importance of cultural infrastructure – both for men and women. In addition, recreation areas seem to operate as pull factor.

The impact of distance on migration behaviour is u-shaped for both sexes. This finding is firstly driven by the large propensity of moving to adjacent regions. Secondly, the farer people move the lesser the binding impact of proximity becomes.[9] Moreover, estimations are in favour of an obvious agglomeration effect. Young adults – men as well

7 If the GCL Model is estimated without consideration of overdispersion, the standard errors are much smaller and all coefficients are highly significant. Since confidence intervals become narrower, nearly all factors exhibit gender-specific differences – as we will see for the most part a statistical result. See appendix tables one and two for the corresponding estimation output.

8 The correlation coefficient between regional wage level and child care facilities is –0.54, so, the conjecture might be true. However, the correlation coefficients within the eastern part and the western part of the migration relations are positive. Therefore, an east-west effect might cause this negative relationship. But in the regressions, this effect is controlled for by the destination East-West dummy variable implemented to guarantee the IIA assumption.

9 Surprisingly, after a distance of 370 minutes the impact turns to a positive direction. However, only 5% of migrants and 0.4% of the total sample move over 370 minutes. So, the right tail of the distribution may be not well identified.

as women – are attracted by more populous regions. And, not surprisingly, the option not to migrate exhibits an exceptionally high probability.

Table 2
GCL-Regression with random effects: Entire migration[a]

	Male Effect		Female Difference		≠
	Coefficient	Stand. Err.	Coefficient	Stand. Err.	
Labor market					
Average wage level	0.3674*	0.0209	0.0953*	0.0175	●
Unemployment rate ♀♂	-0.0125*	0.0008	-0.0023	0.0011	
Land price	0.0171*	0.0007	-0.0019	0.0010	
Education					
College availability ♀♂	0.0184*	0.0008	0.0374*	0.0021	●
Vocational training	0.0048*	0.0006	-0.0012	0.0008	
Employment gap ♀♂	-0.1552*	0.0102	0.0170	0.0127	
Family					
Part-time jobs ♀♂	1.7992*	0.3747	-1.1316*	0.4094	●
Child care	-0.0024*	0.0002	-0.0003	0.0003	
Amenities					
Near-nature area	0.0000	0.0000	0.0000	0.0000	
Park area	0.0005*	0.0001	0.0003	0.0001	
Concert halls	0.1793*	0.0072	-0.0008	0.0098	
Structural characteristics					
Distance	-0.0199*	0.0001	-0.0004*	0.0001	●
Distance²	0.0000*	0.0000	0.0000	0.0000	
Population ♀♂	0.0466*	0.0006	-0.0054*	0.0007	●
Stay Dummy	5.9443*	0.0106	-0.2638*	0.0143	●
Regions	192,721 (439x439)				
Individuals	11,634,142				
Log Likelihood	-568,746.9*				

[a] * 1% significance level; # 5% significance level; ● represents gender-specific differences to 5% significance level. The gender symbols indicate that a variable contains gender-specific values.

Source: Own calculation.

Even if our analysis does not reveal substantial differences between men and women regarding their preferences for location factors, some dissimilarity can be observed at least in terms of intensity. Differences can be seen with respect to wages, college availability part-time jobs, distance, population and the stay dummy. Compared to men young women are supposed to be more migratory (stay dummy) and to choose rather nearby and less populous regions. To a larger extent, they are attracted by regions with adequate availability of universities. And, somewhat surprisingly, women stronger react

on the ratio of regional wages whereas part-time jobs primarily concerns men and not women.

Table 3:
GCL-Regression with Random effects: Long-distance migration[a]

	Male Effect		Female Difference		≠
	Coefficient	Stand. Err.	Coefficient	Stand. Err.	
Labor market					
Average wage level	0.2442*	0.0233	0.1513*	0.0216	●
Unemployment rate ♀♂	-0.0120*	0.0008	-0.0007	0.0011	
Land price	0.0193*	0.0006	-0.0018	0.0009	
Education					
College availability ♀♂	0.0204*	0.0009	0.0498*	0.0023	●
Vocational training	0.0056*	0.0006	-0.0008	0.0008	
Employment gap ♀♂	-0.2113*	0.0113	0.0035	0.0142	
Family					
Part-time jobs ♀♂	1.5964*	0.4001	-1.3331*	0.4372	●
Child care	-0.0006*	0.0002	-0.0002	0.0003	
Amenities					
Near-nature area	0.0001	0.0000	0.0000	0.0000	
Park area	0.0003	0.0001	0.0002	0.0001	
Concert halls	0.2682*	0.0077	0.0036	0.0103	
Structural characteristics					
Distance	-0.0132*	0.0001	-0.0003	0.0002	
Distance²	0.0000*	0.0000	0.0000	0.0000	
Population ♀♂	0.0455*	0.0005	-0.0054*	0.0007	●
Stay Dummy	7.3240*	0.0158	-0.2040*	0.0208	●
Regions	192,721 (439x439)				
Individuals	11,634,142				
Log Likelihood	-456,942.1				

[a] * 1% significance level; ● represents gender-specific differences to 5% significance level. The gender symbols indicate that a variable contains gender-specific values.

Source: Own calculation.

From a regional policy perspective it might be interesting if preferences change when only long-distance movers are considered. Table three displays the estimation results where only relocation with at least 75 minutes distance is counted as migration. Thus, effects caused by suburbanization trends or by arbitrarily fixed administrative borders are filtered out. And migrants seem to be less bounded by social networks and private loyalties when decisions have to be made solely between distant regions.

However, findings remain unchanged, with few exceptions. Park areas are no longer a significant pull factor when long distant migration is considered. This result seems quite obvious since suburbanisation movements are filtered out. Additionally the negative effect of distance becomes smaller and the non-migration option is more attractive when the choice set is restricted to distant regions. Regarding gender effects only one variation can be noticed. Distant migrations are no longer a privilege of young men when the nearby movements are sorted out. Young women seem to heave a strong preference for short distance moves – if only long distant migration is considered no gender-specific effect occurs.

5 Discussion

What are the essential findings of our analysis? To answer the first question of the introduction we have to admit that young women are more mobile than men even after the main location factors of their origin and potential destination regions are taken into account. The second question of gender-specific tastes for location factors requires a less simple answer. Generally, women and men seem to be attracted by the same location factors. Some characteristics are higher valued by men than by women and vice versa. But the direction of valuation is for the most part identical. Young adults prefer regions with high wage levels, low unemployment rates, appropriate jobs for university graduates and adequate educational institutions. The cultural infrastructure also acts as a pull factor.

Regarding gender-specific differences we somewhat surprisingly found a higher valuation of the regional wage level by women. Secondly, women are stronger attracted by educational institutions, in particular by universities. However, our estimations cannot fully verify whether this difference is finally attributable to gender. Educational motivated migration is not a gender phenomenon, but a consequence of educational potential. Women might be more migratory not because they are women, but for their better education.

As a final point, some limitations of our work have to be mentioned. Data stem from an aggregate migration statistics, hence, a lot of desirable information about migrating individuals as well as their life cycle position are not given and ecological fallacies cannot completely ruled out. An essential improvement of our analysis – which of course requires (longitudinal survey) data not available for Germany – would be the implementation of individual level variables and their interaction with regional characteristics. Of primary interest are education and life cycle attributes. Then, we could distinguish between education-, labour market- and family-related migrations and we could analyse if high qualified movers differ from low skilled migrants in their location preferences. Furthermore, discrimination between individual and household migration would be very useful. Otherwise, a huge share of tied movers within households could bias results since their migration behaviour is driven by household and not individual preference.

Appendix

Table A1
GCL-Regression without Random Effects: Entire migration[a]

	Male Effect		Female Difference		≠
	Coefficient	Stand. Err.	Coefficient	Stand. Err.	
Labor market					
Average wage level	0.5897*	0.0143	0.1022*	0.0191	●
Unemployment rate ♀♂	-0.0182*	0.0004	-0.0028*	0.0006	●
Land price	0.0171*	0.0004	-0.0020*	0.0005	●
Education					
College availability ♀♂	0.0198*	0.0003	0.0428*	0.0008	●
Vocational training	0.0099*	0.0003	-0.0014*	0.0004	●
Employment gap ♀♂	-0.2082*	0.0044	0.0095	0.0054	
Family					
Part-time jobs ♀♂	1.1275*	0.1728	0.2533	0.1890	
Child care	0.0006*	0.0001	-0.0005*	0.0002	●
Amenities					
Near-nature area	-0.0001*	0.0000	0.0000	0.0000	
Park area	-0.0010*	0.0001	0.0004*	0.0001	●
Concert halls	0.2822*	0.0026	-0.0113*	0.0034	●
Structural characteristics					
Distance	-0.0325*	0.0000	-0.0010*	0.0001	●
Distance²	0.0000*	0.0000	0.0000*	0.0000	●
Population	0.0622*	0.0002	-0.0079*	0.0003	●
Stay Dummy	4.8163*	0.0030	-0.2878*	0.0041	●
Regions	19,2721 (439x439)				
Individuals	11,634,142				
Log Likelihood	-8,229,520.9				

[a] * 1% significance level; # 5% significance level; ● represents gender-specific differences to 5% significance level. The gender symbols indicate that a variable contains gender-specific values.

Table A2
GCL-Regression without Random Effects: Long-distance migration[a]

	Male Effect		Female Difference		≠
	Coefficient	Stand. Err.	Coefficient	Stand. Err.	
Labor market					
Average wage level	0.6140*	0.0201	0.1320*	0.0270	●
Unemployment rate ♀♂	-0.0177*	0.0006	-0.0019	0.0008	
Land price	0.0166*	0.0004	-0.0027*	0.0006	●
Education					
College availability ♀♂	0.0229*	0.0005	0.0535*	0.0013	●
Vocational training	0.0116*	0.0004	-0.0007	0.0005	
Employment gap ♀♂	-0.2649*	0.0070	0.0019	0.0085	
Family					
Part-time jobs ♀♂	0.1905	0.2439	1.7645*	0.2692	●
Child care	-0.0019*	0.0001	-0.0003	0.0002	
Amenities					
Near-nature area	0.0000*	0.0000	0.0000	0.0000	
Park area	0.0000	0.0001	0.0002	0.0001	
Concert halls	0.3351*	0.0036	-0.0138*	0.0047	●
Structural characteristics					
Distance	-0.0171*	0.0001	-0.0006*	0.0001	●
Distance²	0.0000*	0.0000	0.0000*	0.0000	
Population	0.0550*	0.0003	-0.0071*	0.0004	●
Stay Dummy	6.7207*	0.0091	-0.2438*	0.0125	●
Regions	192,721 (439x439)				
Individuals	11,634,142				
Log Likelihood	-4,582,665.4				

[a] * 1% significance level; # 5% significance level; ● represents gender-specific differences to 5% significance level. The gender symbols indicate that a variable contains gender-specific values.

References

Andrienko, Y.; Sergei, M. G. (2004): Determinants of interregional mobility in Russia., The Economics of Transition 12, pp. 1-27.

Arntz, M. (2006): What Attracts Human Capital? Understanding the Skill Composition of Interregional Job Matches in Germany. ZEW Discussion Paper No. 06-062. Mannheim.

Bruecker, H.; Trübswetter, P. (2007): Do the best go West? An analysis of the self-selection of employed East-West migrants in Germany. Empirica 34, pp. 371-395.

Burda, M. C. (1993): The Determinants of East-West-Migration: Some First Results. European Economic Review 37, pp. 452-461.

Burda, M. C.; Haerdle, W.; Mueller,M.; Werwatz, A. (1998): Semiparametric Analysis of German East-West-Migration Intentions: Facts and Theory. Journal of Applied Econometrics 13, pp. 525-541.

Burda, M. C.; Hunt, J. (2001): From reunification to economic integration: productivity and the labor market in Eastern Germany. Brookings Papers on Economic Activity 2, pp. 1-92.

Cebula, R. J.; Alexander, G.M. (2006): Determinants of Net Interstate Migration 2000-2004. The Journal of Regional Analysis & Policy 36, pp. 116-123.

Cushing, C.; Cushing, B. (2007): Conditional Logit, IIA, and Alternatives for Estimating Models of Interstate Migration. Paper presented at the 46 annual meeting of the Southern Regional Science Association. Charleston, SC.

Cushing, B.; Poot, J. (2004): Crossing Boundaries and Borders: Regional Science Advances in Migration Modelling. Papers in Regional Science 83, pp. 317-338.

Dahlberg, M.; Eklöf, M. (2003): Relaxing the IIA Assumption in Locational Choice Models: A Comparison between Conditional Logit Model, Mixed Logit, and Multinomial Probit Models. Working Paper No. 9-2003. Department of Economics: Uppsala University.

Davies, P. S.; Greenwood, M. J.; Li, H. (2001): A Conditional Logit Approach to U.S. State-to-State Migration. Journal of Regional Science 41, pp. 337-360.

Détang-Dessendre, C.; Molho, I. (2000): Residence spells and migration: a comparison for men and women. Urban Studies 37, pp. 247–260.

Dumont, J.-C.; Martin, J. P.; Spielvogel, G. (2007): Women on the Move: The Neglected Gender Dimension of the Brain Drain. IZA Discussion Paper No. 2920. Bonn.

Faggian, A.; McCann, P.; Sheppard, S. (2007): Some Evidence That Women Are More Mobile Than Men: Gender Differences in U.K. Graduate Migration Behavior. Journal of Regional Science 47, pp. 517-539.

Fidrmuc, J. (2003): Migration and regional adjustment to asymmetric shocks in transition economies. CEPR Discussion Paper No. 3798. London.

Greenwood, M. J. (1997): Internal Migration in Developed Countries, in: M. R. Rosenzweig; O. Stark (eds), Handbook of Population and Family Economics, Vol. 1B. North-Holland: Elsevier, pp. 647-720.

Guimarães, P.; Figeiredo, O.; Woodward, D. (2004): Industrial Location Modeling: Extending the Random Utility Framework. Journal of Regional Science 44, pp. 1-20.

Guimarães, P.; Lindrooth, R. C. (2007): Controlling for Overdispersion in Grouped Conditional Logit Models: A Computationally Simple Application of Dirichlet-Multinomial Regression. Econometrics Journal 10, pp. 439-452.

Harris, J. R.; Todaro, M. P. (1970): Migration, Unemployment & Development: A Two-Sector Analysis. American Economic Review 60, pp. 126-42.

Huber, P. (2004): Inter-Regional Mobility in Europe: A Note on the Cross-Country Evidence. Applied Economics Letters 11, pp. 619-624.

Hunt, J. (2000): Why do People Still Live in East Germany? NBER Working Paper No. 7564. Cambridge, MA.

Hunt, J. (2004): Are Migrants more Skilled than Non-Migrants? Repeat, Return, and Same-Employer Migrants, Canadian Journal of Economics 37, pp. 830-849.

Juerges, H. (2006): Gender ideology, division of housework, and the geographic mobility of families. Review of Economics of the Household 4, pp. 299-323.

Maddala, G. S. (1983): Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press: Cambridge.

McFadden, D. (1974): Conditional logit analysis of qualitative choice behaviour, in: P. Zarembka (ed.), Frontiers in Econometrics. Academic Press: New York, pp. 105-142.

Mincer, J. (1978): Family Migration Decisions. Journal of Political Economy 86, pp. 749-773.

Parikh, A.; Leuvensteijn, M. V. (2003): Interregional Labour Mobility, Inequality and Wage Convergence. Applied Economics 35, pp. 931-941.

Ravenstein, E. G. (1885): The Laws of Migration, Journal of the Statistical Society of London 48, pp. 167-235.

Rees, P.; Kupiszewski, M. (1999): Internal migration and regional population dynamics in Europe – a synthesis. Council of Europe Publishing: Strasbourg.

Sjaastad, L. A. (1962): The Costs and Returns of Human Migration. Journal of Political Economy 70, pp. 80-93.

Todaro, M. P. (1969): A Model of Labour Migration and Urban Unemployment in Less Developed Countries. American Economic Review 59, pp. 138-148.

Zaiceva, A. (2007): East-West Migration and Gender: Is there a Double Disadvantage" vis-à-vis Stayers? IZA-Discussion Paper No. 2810. Bonn.