

Discussion Paper No. 09-014

**Occupational and
Locational Substitution:
Measuring the Effect of
Occupational and Regional Mobility**

Alisher Aldashev

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 09-014

**Occupational and
Locational Substitution:
Measuring the Effect of
Occupational and Regional Mobility**

Alisher Aldashev

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp09014.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.

Non-technical summary

Frictional and structural unemployment are inherent features of imperfect labor markets. Frictional unemployment is caused by imperfect information in the market such that the unemployed do not find jobs at instant but require a certain amount of time until a suitable job is found. Structural unemployment (also referred to as mismatch) appears when vacancy and job-searcher profiles do not coincide (e.g. in terms of qualification or location). Matching function is a convenient empirical tool which is used for the analysis of frictions or mismatch in the labor market in the aggregate.

With the availability of disaggregated regional and occupational data, the disaggregated matching function has been analyzed by the number of authors. All those studies mainly try to identify factors which affect the matching rate. For example, it has been found in the literature that availability of vacancies in neighboring locations increase the matching rate. However, quantification of the mismatch (with only a few exceptions) has not been done so far. In other words, what would the matching rate have been if workers could move across locations or change occupations at a lower (higher) cost?

Partially aggregated data, i.e. disaggregated only at regional or only at occupational level, do not enable splitting the mismatch rate into occupational and spatial components. Therefore, spatial mismatch, measured using regional data only, or qualification mismatch, measured using occupational data only, suffer from partial aggregation bias.

The matching function in this paper is modeled in such a way that vacancies are heterogeneous across two levels: regions and occupations. Using the monthly panel of the Federal Labor Office on unemployment, vacancies and flows from unemployment into employment disaggregated on regional and occupational levels the three-level constant-elasticity-of-substitution (CES) model is estimated, which has not been done in the literature so far. The three-level CES model allows different elasticities of substitution between spatial and occupational dimensions.

The estimates indicate higher regional mobility in West Germany than in the East, but higher occupational mobility in East Germany with respect to West Germany. The results show that in the absence of regional mobility, the number of matches would have been smaller by about 10-25% on average. If occupations were not substitutable the number of matches would have dropped by about 3-5% in the West and by 6-10% in the East. Perfect occupational mobility could increase matchings by 5-9%, which can be interpreted as the measure of occupational mismatch. Perfect regional mobility could increase it by 15% in the West and 5% in the East. This can be interpreted as spatial mismatch.

The paper also analyzes the bias of the estimates caused by partial aggregation. It is claimed that the aggregation bias arises due to biasedness of the elasticity of substitution estimates.

Das Wichtigste in Kürze:

Friktionale und strukturelle Arbeitslosigkeit sind inhärente Merkmale imperfekter Arbeitsmärkte. Friktionale Arbeitslosigkeit wird durch imperfekte Information auf dem Arbeitsmarkt verursacht, die dazu führt, dass Arbeitslose nicht sofort eine neue Stelle finden, sondern für die Suche nach einer passenden Anstellung eine gewisse Zeit benötigen. Strukturelle Arbeitslosigkeit (auch als "Mismatch" bezeichnet) tritt auf, wenn die Profile von Arbeitslosen und Arbeitsplätzen in räumlicher oder beruflicher Hinsicht nicht übereinstimmen. Die Matching-Funktion ist ein geeignetes empirisches Werkzeug, um Friktionen und Mismatches auf dem Arbeitsmarkt im Aggregat zu analysieren.

Auf Grundlage nichtaggregierter Regional- und Berufsdaten haben zahlreiche Autoren die nichtaggregierte Matching-Funktion untersucht. All diese Studien versuchen in erster Linie, Einflussfaktoren der Matching-Rate zu identifizieren. Eine Quantifizierung der Mismatches fehlt hingegen bisher. Es gilt also die Frage zu beantworten: Welche Matching-Rate hätte sich ergeben, wenn Arbeitnehmer in der Lage gewesen wären, ihre Standorte zu wechseln bzw. ihre Berufe zu geringeren (höheren) Kosten zu wechseln?

Räumlich aggregierte Daten, die lediglich auf Berufsebene nicht aggregiert sind, erlauben es nicht, die Rate der Mismatches in eine berufliche und eine regionale Komponente zu zerlegen. Aus diesem Grund weisen Schätzungen des räumlichen Mismatches, die ausschließlich auf Regionaldaten basieren, einen partiellen Aggregationsfehler auf. Gleiches gilt für Schätzungen des beruflichen Mismatches, die ausschließlich mit Berufsdaten durchgeführt werden.

Die Matching-Funktion wird in diesem Artikel so modelliert, dass offene Stellen Heterogenität in zwei Dimensionen aufweisen: Regionen und Berufe. Auf Grundlage eines monatlichen Panels über Arbeitslosigkeit, Vakanzen und Strömen von Arbeitslosigkeit in Beschäftigung der Bundesagentur für Arbeit wird ein dreistufiges Modell konstanter Substitutionselastizitäten (CES) geschätzt. Dies stellt eine Neuerung in der Literatur dar. Das dreistufige CES-Modell erlaubt unterschiedliche Substitutionselastizitäten zwischen räumlichen und beruflichen Dimensionen.

Die Schätzergebnisse deuten auf eine höhere regionale Mobilität in Westdeutschland und eine höhere berufliche Mobilität in Ostdeutschland hin. Die Ergebnisse zeigen, dass ohne regionale Mobilität die Anzahl der Matches durchschnittlich um 10 – 25% geringer gewesen wäre. Wenn Berufe nicht gegeneinander substituierbar wären, wäre die Anzahl der Matches um ca. 3 – 5% in West- und um 6 – 10% in Ostdeutschland gesunken. Vollkommene berufliche Mobilität könnte die Matchings um 5 – 9% steigern. Diese Werte kann man als ein Maß des beruflichen Mismatches interpretieren. Vollkommene regionale Mobilität könnte zu einem Anstieg der Matchings um 15% im Westen und um 5% im Osten führen. Diese Werte messen das Ausmaß des räumlichen Mismatches.

Der Artikel untersucht außerdem die durch partielle Aggregation verursachte Verzerrung der Schätzer. Es wird argumentiert, dass der Aggregationsfehler aufgrund der Verzerrung der Substituierbarkeits-Schätzer entsteht.

Occupational and locational substitution: measuring the effect of occupational and regional mobility.

Alisher Aldashev*

ZEW, Mannheim

March 24, 2009

Abstract

The paper analyzes effects of occupational and regional mobility on the matching rate using the monthly panel disaggregated on regional and occupational level. The main contribution of the paper is measuring the effect of substitutability between vacancies for different occupations and vacancies in different regions on matchings. The estimates indicate higher regional mobility in West Germany but higher occupational mobility in East Germany. The results show that if occupations were perfect substitutes, the number of matches could increase by 5-9%. Perfect regional mobility could increase matchings by 5-15%. It is also shown that partial aggregation causes a downward bias in substitutability estimates.

Keywords: Matching function, constant elasticity of substitution, spatial correlation, occupational and regional mobility, nonlinear least squares, GMM

JEL Classification: J62, J63, R23, J61

1 Introduction

In the empirical context mismatch "...measures the degree of heterogeneity in the labor market across a number of dimensions, usually restricted to skills, industrial sector, and location." (Petrongolo and Pissarides, 2001; p. 399) Spatial mismatch happens because moving across locations is costly for unemployed workers. Skill mismatch occurs because unemployed workers cannot instantaneously change their occupation/profession at no cost depending on what is demanded in the market.¹

*Alisher Aldashev is Research Fellow at the Centre for European Economic Research (ZEW), L 7,1, D-68161 Mannheim, e-mail: aldashev@zew.de.

¹ Definitions of skill mismatch may vary. Skill mismatch could be also understood as differences between skills possessed by workers and requirements of firms. In the context of this paper, skill mismatch is understood as occupational mismatch.

The literature on aggregate matching function is abundant (Layard et al., 1991, Burda and Wyplosz, 1994, Petrongolo and Pissarides, 2001; and citations therein). With the availability of disaggregated regional data spatial matching function has also been estimated by a number of authors (Coles and Smith, 1996, Anderson and Burgess, 2000, Burgess and Profit, 2001, Fahr and Sunde, 2006; among others). Mismatch across occupations is analyzed by Fahr and Sunde (2004) and in the recent paper by Dolado and Felgueroso (2008).

All those studies mainly try to identify factors which affect the matching rate. However, quantification of the mismatch (or what percentage of the matching rate is due to occupational mismatch and what is due to spatial immobility) has not been done so far. Put it in other words, what would the matching rate have been if workers could move across locations at less (more) cost? What would the matching rate have been if workers could change occupations at less (more) cost?² The studies on spatial mismatch suffer from aggregation bias across occupations and studies on occupational mismatch suffer from aggregation bias across spatial units. Hence, having separate data on locations and occupations does not enable us to split the mismatch rate into occupational and spatial mismatch components.

In this paper I use the panel data set disaggregated both on regional as well as on occupational level. This is to the best of my knowledge the first attempt to decompose the matching function into occupational and locational mismatch components. Empirical estimates show that if occupations were perfect complements (no substitutability) the number of matches would have dropped by about 3-5% for unemployed men in the West and 6-10% in the East on average. If regions were isolated units, so no spatial mobility took place, the number of matches would have decreased by about 10-25% on average. If workers could immediately re-qualify for occupations demanded by employers the number of matches could increase by about 5-9%.

It is argued that the aggregation bias of substitutability estimates arises because of aggregation of inputs in the matching function with different levels of substitutability. Empirical exercise show that partial aggregation downward biases the estimates of substitution.

The paper is organized as follows: Section 2 gives a theoretical overview of the matching function and constant elasticity of substitution, Section 3 describes the dataset used for estimation, Section 4 describes the method of correction for spatial correlation in the nonlinear model, Section 5 presents estimation results, Section 6 discusses the aggregation bias problem, Section 7 discusses spatial weighting, and Section 8 concludes.

2 Matching Function

The matching function in its simplest form can be expressed as $M = m(U, \check{V})$, with U denoting the unemployed stock and \check{V} the vacancy stock. Mostly, the matching function is treated as a “black-box” in the literature, usually approximated by the Cobb-Douglas function with

² This is equivalent to saying that workers supplied occupations which were more (less) demanded by firms.

constant returns to scale:³ $M = U^\alpha \check{V}^{1-\alpha}$. However, a recent study by Stevens (2007) provides new microfoundations for the matching function. She shows that the matching process based on a “telephone-line” Poisson queuing process with both sides (unemployed and firms) engaged in the search implies a general CES matching function.

To analyze matching across locations and occupations I consider two levels of heterogeneity in vacancies: regions and occupations. Consider then \check{V} as a set of three factors: V , V_S , and V_R . Variable V denotes vacancies located in the unemployed’s region of residence with occupational requirements corresponding to the one possessed by an unemployed worker. Variable V_S denotes vacancies for an occupation different to the one possessed by an unemployed worker⁴ but located in the same spatial unit. Variable V_R denotes vacancies for the same occupation as the one possessed by an unemployed but located in a different spatial unit.⁵ An unemployed can substitute V_S and V_R for V given the degree of substitutability between them.⁶

To allow for different substitutability between V and V_S and V and V_R I apply the Sato (1967) three-level constant elasticity of substitution (CES) production function. The matching function can be expressed as a mixture of Cobb-Douglas and CES components assuming constant returns to scale (CRS)⁷ (Krusell et al., 2000, Duffy et al., 2004; among others):

$$M(U, V, V_S, V_R) = \gamma U^\alpha \left(a \left[bV^\theta + (1-b)V_R^\theta \right]^{\rho/\theta} + (1-a)V_S^\rho \right)^{(1-\alpha)/\rho}, \quad (1)$$

where $\{0 \leq a \leq 1\}$, $\{0 \leq b \leq 1\}$, $\{-\infty < \rho \leq 1\}$, and $\{-\infty < \theta \leq 1\}$.

Parameters θ and ρ measure elasticity between V and V_R and V and V_S respectively. If $\theta = 1$ then V_R is perfectly substitutable for V and if $\rho = 1$ then V_S is a perfect substitute for V . Evaluation of the function in Equation 1 at $\theta = 1$ or $\rho = 1$ is trivial. Another interesting example is perfect complementarity between factors. Values $\theta = -\infty$ or $\rho = -\infty$ imply no substitution between inputs. One can show that:

$$\lim_{\rho \rightarrow -\infty} M(U, V, V_S, V_R) = \gamma U^\alpha \min \left(V_S, (bV^\theta + (1-b)V_R^\theta)^{1/\theta} \right)^{(1-\alpha)} \quad (2)$$

and

$$\lim_{\theta \rightarrow -\infty} M(U, V, V_S, V_R) = \gamma U^\alpha \left[a \min(V, V_R)^\rho + (1-a)V_S^\rho \right]^{(1-\alpha)/\rho}. \quad (3)$$

Equation 2 shows the case where vacancies for the same occupation but in a distant region cannot be substituted for vacancies for the same occupation in the region of residence. This

³ There is a large number of studies which indicate existence of constant returns to scale in the matching function (Petrongolo and Pissarides, 2001).

⁴ I would call it *unfamiliar occupation* throughout the paper.

⁵ I would call it *distant region* throughout the paper.

⁶ One could argue that an unemployed could also consider a vacancy for an unfamiliar occupation in a distant region. However, an unemployed would strictly prefer V_S and V_R to this option. So the substitutability of this option for V should be low. It is in my view reasonable to leave this option in the residual rather than complicate the model even further making it intractable.

⁷ Relaxing the CRS assumption resulted in non-convergence of the NLS estimation.

could be if an unemployed with occupation i in region A will not accept a job for occupation i in region B . So for convenience I would assume that

$$\lim_{\rho \rightarrow -\infty} M(U, V, V_S, V_R) \equiv \gamma U^\alpha \left(bV^\theta + (1-b)V_R^\theta \right)^{(1-\alpha)/\theta}, \quad (4)$$

so that the number of matchings is independent of vacancies in distant regions. This is the case of no substitution between factors. It is somewhat similar to perfect complementarity. There is, however, conceptual difference. Perfect complementarity implies that both factors are needed to produce an output. Here, by no substitution I mean that only V is needed to produce M and V_S and/or V_R have no effect on M whatsoever. This is equivalent to setting $a = 1$ in equation 1.

Equation 3 shows the case where vacancies for an unfamiliar occupation in the same location cannot be substituted for vacancies for the same occupation in the region of residence. This could be if an unemployed with occupation i in region A will not accept a job for occupation j in region A . So for practical reasons I would assume that

$$\lim_{\theta \rightarrow -\infty} M(U, V, V_S, V_R) \equiv \gamma U^\alpha \left[aV^\rho + (1-a)V_S^\rho \right]^{(1-\alpha)/\rho}, \quad (5)$$

so that the number of matchings is independent of vacancies for other occupations. This is equivalent to setting $b = 1$ in equation 1.

3 Data and Model Specification

In the spatial matching literature the specification of the matching function usually depends on data availability and the way data are recorded. For example, Ibourk et al. (2004) and Ilmakunnas and Pesola (2003) do not include spatially lagged exogenous variables in their specification. However, Burgess and Profit (2001) include both spatially lagged vacancies and spatially lagged unemployment as regressors. The authors find positive effect of vacancies in distant regions on matches, which present additional job opportunities to local unemployed. The effect of the unemployment stocks in distant regions is negative, which can be interpreted as congestion or competition. Fahr and Sunde (2002) used both spatially lagged unemployment and vacancies as regressors and found that both effects were negative. However, there are no theoretical arguments as to why additional job opportunities in surrounding regions should cause congestion in the local job market. In their later work Fahr and Sunde (2006) include only spatially lagged unemployment in their specification and find a negative effect, which might be interpreted as competition among the unemployed for jobs. The exclusion of spatially lagged vacancies is motivated by the data construction. Namely, the location of the match is related to the employer.

Schematically this can be shown as:

	V	A	B
U			
A		match in A	match in B
B		match in A	match in B

This implies that job creation in region A is recorded if a vacancy in A is matched with an unemployed from A or B. Hence, matches in region A are directly dependent on vacancy stock in A and unemployment stocks in A and B.

For estimation I use the data provided by the Federal Labor Office (Bundesagentur für Arbeit) which contains the information on reported vacancies and registered unemployed as well as on flows from unemployment into employment. Vacancies which are reported to the Federal Labor Office contain the information on occupation required for the job. Unemployed workers who register at the Federal Labor Office report occupation in which they would like to work. Using this information the Federal Labor Office constructed a monthly panel covering years 2000-2004 disaggregated at regional and occupational level. The regional dimension comprises 440 German NUTS3 regions (Kreise). Occupations are summarized in 18 broad categories.⁸ The analysis is restricted to unemployed men. The dataset

The dependent variable is the flow out of unemployment into employment (without professional training at workplace) in a given region for a given occupational group. One should note that the variable is recorded in such a way that if an unemployed registered in region A finds a job in region B, he will be counted in the flow into employment of region A.

This can be schematically summarized as:

	V	A	B
U			
A		match in A	match in A
B		match in B	match in B

The first column represents the region in which an unemployed worker resides and the first row represents the region in which the vacancy is posted. So if an unemployed living in region A finds a job which is located in region B, this will be recorded as a match in region A. This implies that job creation in region A is recorded if an unemployed worker in A is matched with a vacancy in region A or B. Hence, matches in region A are directly dependent on vacancy stocks in A and B and unemployment stock in A only.

The main independent variables are the stock of unemployed and vacancies (V , V_S , V_R), registered at the Federal Labor Office. Only vacancies which foresee an employment spell of at least seven calendar days are recorded. V is the number of reported vacancies for the same occupation as the one specified by the unemployed and located in the same NUTS3 region. V_S

⁸ See Appendix for details.

is the number of vacancies for all other occupations than the one specified by the unemployed and located in the same NUTS3 region. V_R is the number of vacancies for the same occupation as the one specified by the unemployed but located in all other regions. The variable V_R was then weighted using the inverse travel time spatial weighting matrix. The information on travel time between the administrative centers of NUTS3 regions were provided by the Federal Statistical Office (Statistisches Bundesamt). The elements on the main diagonal were set to zero. The spatial weighting matrix was then row-normalized. Hence, V_R can be interpreted as the weighted average of vacancies in all other regions for the same occupation.

One could argue that officially registered vacancies do not correspond to the number of vacancies available in the economy. In fact, the coverage rate might change with economic cycle. Authors dealing with time-series data usually adjust the reported vacancy variable to “clean” it from the cyclical component. In case of longitudinal data one could capture the changes of the coverage rate over time by using time dummies as additional regressors without having to adjust the vacancy variable.⁹

One should keep in mind, however, that this is not a panacea as it would still bias the results if the change in the coverage rate in a business cycle is unequal across locations or occupations. In fact, Abraham (1987) provides some evidence at least for the US Help-Wanted Index that the composition of vacancies changes over time (blue vs. white-collar occupations). To partially mitigate for that I also control for occupational dummies.

Since in my data location of the match is related to the residence of the unemployed worker, I do not include spatially lagged unemployment stock as a regressor in my specification (see argumentation above). Another argument is that with more than three factors the CES function becomes numerically intractable unless one assumes equal substitutability across factors, which is preferable to avoid.

To control for possible differences in search intensity across different demographic groups I also use the share of long-term unemployed (unemployment duration of at least one year), the share of unemployed under 25 years of age as well as over 55, the share of unemployed with foreign citizenship as auxiliary regressors (see Petrongolo and Pissarides, 2001, and citations therein).

4 Econometric methods

The model given in 1 can be estimated by the generalized method of moments. In the nonlinear model, the moment condition is $X'(\beta)(y - x(\beta)) = 0$, with X being a matrix containing the first derivatives of the regressor matrix, x , with respect to the parameter vector β evaluated at x . The parameter vector estimated given this moment restriction is the nonlinear least squares (NLS)

⁹ For example Fahr and Sunde (2004), Fahr and Sunde (2006), Burgess and Profit (2001), Ibourk et al. (2004) also use unadjusted officially reported vacancies as regressors.

estimator and is close to the method of moments estimator (see Davidson and MacKinnon, 1993).

Since unemployed workers from different regions are free to move across locations and hence compete with each other for existing vacancies, this could cause spatial correlation in the regression residual which is typical for spatial data (see Anselin, 2000, for an overview on spatial interaction). The standard errors could be corrected for spatial correlation using the method of Driscoll and Kraay (1998). The method is in principle an extension of GMM estimator of Newey and West (1987). In a simple univariate model $y_{it} = x_{it}\beta + \epsilon_{it}$ with spatial but no time dependence, the identifying moment restriction is: $E(x_{it}\epsilon_{it}) = 0$. Driscoll and Kraay (1998) show that the variance matrix is given by:

$$V = (X'X)^{-1}S_T(X'X)^{-1}, \quad (6)$$

where $S_T = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^N E(x_{it}\epsilon_{it}x_{jt}\epsilon_{jt})$. The method works, however, in the linear case. To correct standard errors in the nonlinear model one could run a linear artificial regression:

$$r(\hat{\beta}) = X(\hat{\beta})b + res, \quad (7)$$

where $r(\hat{\beta})$ are residuals from the NLS regression evaluated at the estimated parameter value $\hat{\beta}$, $X(\hat{\beta})$ is the matrix of first derivatives of x evaluated at the estimated parameter value $\hat{\beta}$ and b is the coefficient vector and res is the residual which have no further interpretation. The estimated covariance matrix of b is an estimator of the covariance matrix of β . Applying the method of Driscoll and Kraay (1998) on the linear regression in equation 7 gives the consistent estimator of the covariance matrix of β .

5 Estimation results

5.1 Specification Issues

The functional form of the spatial or occupational matching function is usually taken as Cobb-Douglas in the literature (see for example Fahr and Sunde, 2006; 2004, Burgess and Profit, 2001, Ilmakunnas and Pesola, 2003).¹⁰ The constant returns to scale is a usual assumption, albeit some authors do not find support for it (Fahr and Sunde, 2006, Burgess and Profit, 2001; are few examples).

In case of spatially and occupationally disaggregated data the Cobb-Douglas specification is problematic as it would necessitate that the substitutability across all factors is equal. For this reason the three-level CES specification was chosen. The constant returns to scale condition was

¹⁰ For an example of a translog specification see Ibourk et al. (2004).

imposed for computational reasons: while relaxing the CRS assumption numerical convergence could not be achieved.

In section 2 the specification of the matching function was given as:

$M(U, V, V_S, V_R) = \gamma U^\alpha \left(a [bV^\theta + (1-b)V_R^\theta]^{\rho/\theta} + (1-a)V_S^\rho \right)^{(1-\alpha)/\rho}$. However, one could consider an alternative specification:

$$M(U, V, V_S, V_R) = \gamma U^\alpha \left(a [bV^\theta + (1-b)V_S^\theta]^{\rho/\theta} + (1-a)V_R^\rho \right)^{(1-\alpha)/\rho}. \quad (8)$$

The difference between the two is in differences in elasticities of substitution between the inputs. In the first specification the elasticities of substitution between V and V_S and between V_R and V_S are the same. In the second, the elasticities of substitution between V and V_R and V_S and V_R are the same. Both specifications are theoretically plausible and one cannot prefer one over the other *a priori*. For regression models based on the second specification I obtained the substitutability estimates implausible in magnitude, namely values of θ around two, whereas theoretically they cannot exceed unity. It is not unusual that one of the two specifications produces implausible estimates (Krusell et al., 2000; among others). Duffy et al. (2004) also obtained substitutability estimates in excess of unity in one of the specifications and rejected it on these grounds. For this reason specification 1 is preferred throughout the paper. Estimated parameters for a specification in equation 8 can be found in table A.2.

5.2 Counterfactuals

The estimation results of the model in equation 1 are given in table 1. In the light of equation 1 large θ indicate high substitutability between V and V_R . One should, however, not interpret large θ as perfect regional mobility or low cost of spatial mobility. One should remember that V_R was weighted by the inverse distance and therefore implicitly accounts for travel cost by construction. Hence, large θ should not be surprising.

Comparing the results between East and West Germany one could notice that the elasticity of the unemployment stock is almost identical. However, one could see important differences. Similar values of b but larger θ for West German males indicate higher regional mobility in the western part of the country. On the other hand, smaller value of a and larger value of ρ for East German males indicate higher occupational mobility in the East. Put it in other words, the unemployed in the West are more likely to move to a different region to get a job than the unemployed in Eastern Germany. On the other hand, the unemployed in the East are more likely to accept a job for an occupation different to the one they possess than their western counterparts.

An interesting exercise would be to estimate the number of matchings if there were no substitutability between V and V_S or V and V_R . To do so, I plug the parameter estimates from

Tab. 1: Three-level CES model estimates, preferred specification^a

Parameter ^b	Men West		Men East	
	Coefficient	D-K St. Error ^c	Coefficient	D-K St. Error ^c
a	0.92	0.03	0.87	0.03
b	0.58	0.02	0.57	0.02
α	0.77	0.01	0.76	0.01
θ	0.93	0.08	0.81	0.05
ρ	0.22	0.13	0.38	0.09
R^2	0.86		0.90	
N	306,373		107,767	

^a variables are based on the Model in Equation 1

^b the model also contains the shares of long-term unemployed, unemployed with foreign nationality, youth and elderly unemployed, log area, as well as monthly and yearly dummies and occupation dummies.

^c D-K st. err. stand for Driscoll-Kraay standard errors

equation 1 into equations 4 and 5:

$$\begin{aligned}
 M_1(U, V, V_S, V_R) &= \hat{\gamma} U^{\hat{\alpha}} \left(\hat{b} V^{\hat{\theta}} + (1 - \hat{b}) V_R^{\hat{\theta}} \right)^{(1-\hat{\alpha})/\hat{\theta}} \\
 M_2(U, V, V_S, V_R) &= \hat{\gamma} U^{\hat{\alpha}} \left[\hat{a} V^{\hat{\rho}} + (1 - \hat{a}) V_S^{\hat{\rho}} \right]^{(1-\hat{\alpha})/\hat{\rho}}.
 \end{aligned} \tag{9}$$

Alternatively, one could estimate the number of matchings if there were perfect substitutability between V and V_R or V and V_S :

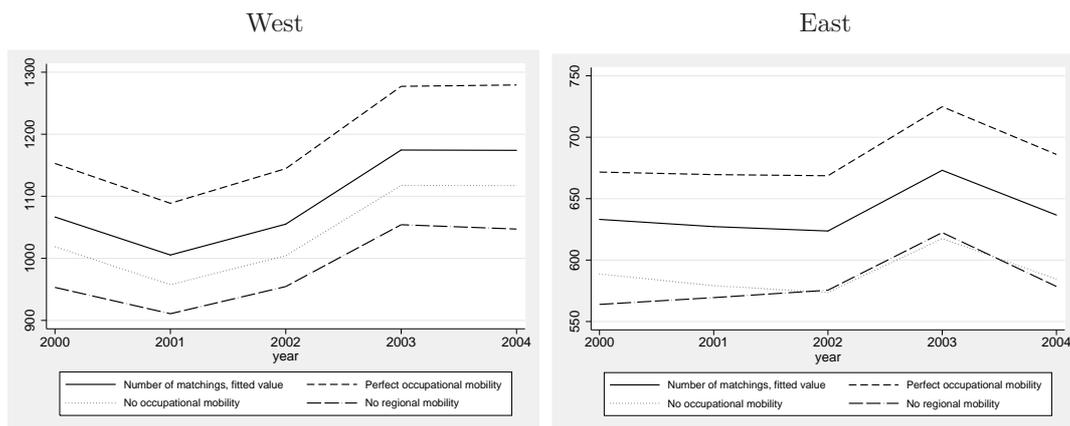
$$\begin{aligned}
 M_3(U, V, V_S, V_R) &= \hat{\gamma} U^{\hat{\alpha}} \left(\hat{a} \left[\hat{b} V + (1 - \hat{b}) V_R \right]^{\hat{\rho}} + (1 - \hat{a}) V_S^{\hat{\rho}} \right)^{(1-\hat{\alpha})/\hat{\rho}} \\
 M_4(U, V, V_S, V_R) &= \hat{\gamma} U^{\hat{\alpha}} \left(\hat{a} \left[\hat{b} V^{\hat{\theta}} + (1 - \hat{b}) V_R^{\hat{\theta}} \right]^{1/\hat{\theta}} + (1 - \hat{a}) V_S \right)^{(1-\hat{\alpha})}.
 \end{aligned} \tag{10}$$

The counterfactuals M_1 , M_2 , and M_4 are shown in figure 1.¹¹ The dotted line depicts the counterfactual situation if no occupational mobility were possible. Hence, the difference between the solid and the dotted line measures the net gain from occupational mobility. Thus, the interpretation is that if occupations were not substitutable the outflow from unemployment into employment would have dropped by 5% in the West and 10% in the East.

The dash-dotted line corresponds to the counterfactual situation of no regional mobility. Hence, it implies that the difference between the solid and dash-dotted line constitutes the net gain

¹¹ M_3 is not shown as it is very close to the fitted value \hat{M} due to a large value of θ , M_3 is only about 0.3-0.9% larger than the fitted value of M .

Fig. 1: Nonlinear matching estimates, fitted values (in thousands)



from regional mobility. If regions were isolated units, so no spatial mobility took place, the number of matches would have decreased by about 10% on average.

The difference between the solid and dashed line shows the potential gain from perfect occupational mobility. This shows room for improvement of the matching rate. The matching rate could move closer to that limit if, for example, costs of changing a profession would decrease (for example by providing retraining programs) or if, for instance, more jobs are created corresponding to the professions supplied by the unemployed. This difference can be interpreted as the maximum level that can be achieved by increasing occupational mobility, which is roughly 10%. This is the part of mismatch caused by imperfect occupational mobility.

An attempt to conduct similar counterfactual experiments was undertaken by Börsch-Supan (1990). The author, however, does not estimate the matching function but rather concentrates on numerical exercises of changing the distribution of unemployed or vacancies across regions and occupations. The author uses a cross section for the year 1986 for 142 local labor markets and 86 occupations, thereby having 142×86 submarkets. The first counterfactual situation considered by the author is what would have been if all vacancies in each submarket were filled by the unemployed in case of excess supply or all unemployed were matched with a vacancy in case of excess demand. This would be the case of perfect match within each submarket but no mobility across submarkets. Aggregating submarkets over occupations and redoing the same mechanical exercise would correspond to perfect occupational mobility. The perfect regional mobility counterfactual is constructed in a likewise manner. The author comes to the conclusion that perfect occupational and regional mobility can bring 10% in the matchings at maximum.

One could see the differences between the approach presented in this paper and the one of Börsch-Supan (1990). The latter compares mobility versus no mobility counterfactuals under assumption of perfect search, that is given that all vacancies are filled by the unemployed in case of excess supply of labor and all unemployed find jobs in case of excess demand for labor. In other words, the counterfactuals are compared under assumption of no *frictional unemployment*.

In my paper, the matching function is estimated and the counterfactuals are constructed given the existing level of frictions in the labor market.

6 Aggregation Bias and Functional Form

The aggregation bias of the matching function has been scarcely addressed in the literature mainly comparing the spatially aggregated matching function and spatially disaggregated one. Coles and Smith (1996) argue that increasing returns to scale in the matching function can be expected because of spatial interaction and hence spatial aggregation might downward bias the elasticity estimates. Anderson and Burgess (2000) also report the point estimates of the spatially disaggregated matching function to be higher than those obtained using aggregate national data. However, the authors find it difficult to explain the discrepancy: "... with this difference perhaps due to the approximations necessary to use the national data."¹²

The aim of this section is to look at the aggregation bias from another perspective: namely, the biasedness of the substitutability estimates. Assume that the economy consists of two regions A and B with only two occupations i and j . Consider an unemployed worker of occupation i residing in region A . He has four vacancy types: V_{iA} , V_{iB} , V_{jA} , and V_{jB} with V_{iA} strictly superior and V_{jB} strictly inferior to any other (as it involves a cost of locational move and a cost of changing an occupation).¹³

If V_{jB} is strictly inferior to other options and hence is a poor substitute for V_{iA} then estimation of the CES model using partially aggregated data would bias the substitutability estimates downward. For example, aggregation across regions would produce substitutability estimates between V_A and V_B lower than between V_{iA} and V_{iB} because V_B would contain V_{jB} , i.e. a job searcher is less likely to move to a distant region for an unfamiliar occupation. In the same fashion, aggregation across occupations would produce substitutability estimates between V_i and V_j lower than between V_{iA} and V_{jA} because V_j would contain V_{jB} .

Tab. 2: Estimates comparison

variables	Men West			Men East		
	occup. and reg. data	occupational data only	regional data only	occup. and reg. data	occupational data only	regional data only
a	0.92	1.00		0.87	0.99	
ρ	0.22	0.83		0.38	1.52	
b	0.58		0.63	0.57		0.71
θ	0.93		0.14	0.81		0.64

¹² Anderson and Burgess (2000; p.99)

¹³ It has been already mentioned in the footnote in section 2.

As an exercise I estimated equation 4 using regional data only (all other variables within a region are aggregated across occupations) and equation 5 using occupational data only (all other variables within an occupation are aggregated across regions). These results are compared with estimates using two levels of disaggregation in table 2. Aggregation across regions produces a larger estimate of b and smaller estimate of θ , which implies less substitutability between vacancies in different regions than if using disaggregated data. These results indicate that difference in substitutability between occupations and regions (or to be precise poor substitutability between vacancies in a certain region for a certain occupation and vacancies in other regions for other occupations) might bias the estimates downward when using aggregate data. What is most surprising is that estimates using regional data only produces larger elasticity of substitution for East Germany whereas estimates using regional and occupational data showed that substitution across space was larger in the West. One explanation could be that the estimate of the elasticity of substitution using regional data only is not a correct estimate of regional mobility. In other words, regional data implies aggregation of occupational groups within regions and hence the estimate of the elasticity of substitution measures both mobility across occupations as well as regions. Since occupational substitution in the West is lower than in the East (see table 1) the estimate of the elasticity of substitution using regional data appears to be lower in West Germany.

The results based on occupationally aggregated data are difficult to interpret as the estimates have theoretically implausible magnitudes (a larger than 1 for Males West and ρ greater than unity for Males East).

I also compare the estimates using regional data only using the CES specification with the Cobb-Douglas specification used by Fahr and Sunde (2006). The function to be estimated (suppressing auxiliary variables to save notation) is:

$$M = U^\alpha V^f V_R^g. \quad (11)$$

This is unrestricted Cobb-Douglas function estimated by Fahr and Sunde (2006).¹⁴ Since I assumed constant returns to scale in the CES model estimation I also estimate the Cobb-Douglas matching function under CRS assumption, which would be of the form:

$$M = U^\alpha V^f V_R^{1-\alpha-f} \quad (12)$$

Remembering that the CES function is of the form: $M = U^\alpha (bV^\theta + (1-b)V_R^\theta)^{\frac{1-\alpha}{\theta}}$, if one imposes the Cobb-Douglas restriction to the CES ($\theta = 0$), this would imply that $f = b(1-\alpha)$ and $g = (1-b)(1-\alpha)$ in case of constant returns to scale. Table 3 compares estimates of the CES model with restricted and unrestricted Cobb-Douglas models.

The last row in table 3 shows the measure of substitutability between regional vacancies and therefore is equal to zero for Cobb-Douglas specification by definition. One could see that

¹⁴ As noted before due to differences in data construction Fahr and Sunde (2006) use spatially lagged unemployment instead of vacancies in their specification.

Tab. 3: CES and Cobb-Douglas Estimates, comparison

variables	Men West			Men East		
	CES-CRS ^a	CD-CRS ^b	CD-unrestr. ^c	CES-CRS	CD-CRS	CD-unrestr.
α	0.89	0.89	0.87	0.86	0.80	0.81
b	0.63			0.71		
f	-	0.07	0.06	-	0.10	0.10
g	-	0.04	-0.12	-	0.10	0.12
θ	0.14	0	0	0.64	0	0

^a CES-CRS: Constant elasticity of substitution with constant returns to scale restriction imposed

^b CD-CRS: Cobb-Douglas specification with constant returns to scale restriction imposed

^c CD-CRS: unrestricted Cobb-Douglas specification

the estimates of α and b (given that $f = b(1 - \alpha)$) are very similar across all specifications for males in West Germany. This could mean that Cobb-Douglas function is a reasonable approximation for this group of unemployed workers (in fact θ is statistically indistinguishable from zero). However, the effect of vacancies in surrounding regions is negative in the unrestricted Cobb-Douglas specification which is counterintuitive. For males in East Germany Cobb-Douglas specification supports the CRS assumption. However, θ is substantially larger than zero and hence Cobb-Douglas specification is not an adequate approximation of the CES function.

7 Spatially Weighted vs. Unweighted Vacancies

Consider for illustration purpose that vacancies and unemployed were homogenous with respect to occupations. There is however regional heterogeneity. Namely, there are three types of vacancies: located in the region of residence (denoted V), within an acceptable commuting distance (V_{R1}), lying beyond an acceptable commuting distance (V_{R2}). The matching function has the following form:

$$M(U, V, V_{R1}, V_{R2}) = \gamma U^\alpha \left(\tilde{a} \left[\tilde{b} V^{\tilde{\theta}} + (1 - \tilde{b}) V_{R1}^{\tilde{\theta}} \right]^{\tilde{\rho}/\tilde{\theta}} + (1 - \tilde{a}) V_{R2}^{\tilde{\rho}} \right)^{(1-\alpha)/\tilde{\rho}}. \quad (13)$$

Vacancies which lie beyond an acceptable commuting distance cannot be accepted¹⁵ and therefore V_{R2} cannot be substituted for V , and hence, $\tilde{a} = 1$ and/or $\tilde{\rho} = -\infty$. Consequently in estimation one could drop V_{R2} from the model.

Whether a vacancy lies within an acceptable commuting distance depends on various factors like the distance, the wage paid, acceptance probability or value of leisure. For example, for a high

¹⁵ One could argue that a person can still migrate to the region beyond an acceptable commuting distance and accept the job there. However, the migration probability also declines with the distance.

enough wage a far away region might still be an acceptable commuting destination. However, the likelihood for a region to be an acceptable commuting destination goes down with distance. Hence, weighting the variable using inverse distances is an approximation to calculating the number of *acceptable* vacancies from the total. On these grounds it seems reasonable to weight vacancies using inverse distances and use this variable as a proxy for the number of *acceptable* vacancies in other regions, which was done in the empirical section of the paper. Albeit, it is not known how close is the approximation to the true value.

Another problem arises with interpretation of counterfactual case of perfect regional mobility. Dropping V_{R2} from the model and setting $\tilde{\theta} = 1$ would imply perfect mobility between *acceptable* vacancies but not all vacancies. Perfect regional mobility implies that both $\tilde{\theta}$ and $\tilde{\rho}$ in equation 13 equal to 1. To alleviate the problem, I estimate the matching function in equation 1 using unweighted vacancies, i.e. V_R is the sum of all vacancies in other regions without weighting.

Tab. 4: Three-level CES model estimates, alternative specifications

Parameter ^a	Men West		Men East	
	Model 1 ^b	Model 2 ^c	Model 1 ^b	Model 2 ^c
a	0.92	0.93	0.87	0.84
b	0.58	0.99	0.57	0.99
α	0.77	0.76	0.76	0.76
θ	0.93	0.83	0.81	0.72
ρ	0.22	0.21	0.38	0.45
R^2	0.86	0.86	0.90	0.90
N	306,373	306,373	107,767	107,767

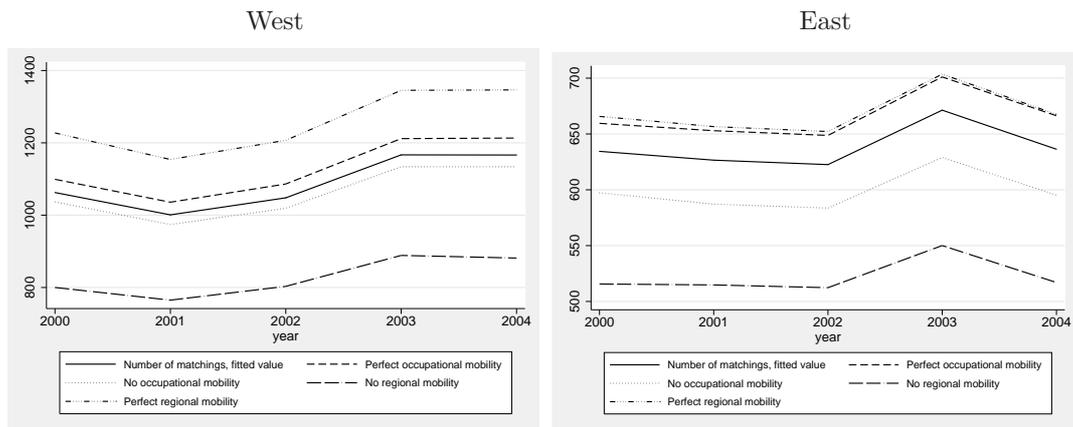
^a the model also contains the shares of long-term unemployed, unemployed with foreign nationality, youth and elderly unemployed, log area, as well as monthly and yearly dummies and occupation dummies.

^b regional vacancies are spatially weighted

^c regional vacancies are NOT spatially weighted

The estimation results using unweighted vacancies can be found in columns 3 and 5 in table 4. Since a weighted average number of vacancies in other region is smaller than the sum of all vacancies in other region, the parameter b is larger in magnitude in this alternative specification. Moreover, since the sum of all vacancies in other regions contains also unacceptable vacancies V_{R2} the parameter θ is smaller in the alternative specification. Remarkably, other main parameters are unaltered for males in the West. For eastern males the parameters slightly changed. Increase in ρ and decrease in a imply that the alternative specification using unweighted vacancies produces a higher estimate of occupational mobility for East German males.

Fig. 2: Nonlinear matching estimates (unweighted regional vacancies), fitted values (in thousands)



The counterfactuals for different levels of elasticity of substitution are plotted in figure 2. The graph shows that in case of perfect regional mobility 15% improvement of the number of matchings were possible in the West and about 5% in the East. One should also note differences in the results between figure 1 and 2. The estimates based on spatially weighted vacancies imply about 9% gain in matchings if occupations were perfectly substitutable. The estimates based on unweighted vacancies suggest on the contrary some 5% improvement. Hence, the occupational component of the mismatch is reduced by about half by using unweighted vacancies.

If occupations were not substitutable the number of matchings would have dropped by 5% in the West and 10% in the East according to estimates using spatially weighted vacancies and by 3% and 6% respectively according to estimates using unweighted vacancies. In case of no regional mobility the number of matchings would have dropped by 10% according to estimates using spatially weighted vacancies and by 20-25% according to estimates using unweighted vacancies.

The comparison shows that using unweighted vacancies in the estimation produces a smaller estimate of the elasticity of regional substitution and could serve as a lower bound. On the other hand, using weighted vacancies might potentially overestimate the effect if the strict inverse relationship between the number of acceptable vacancies and travel time is violated. For that reason the loss in the number of matchings in case of no regional mobility would be in the interval between 10% (using weighted vacancies) and 20-25% (using unweighted vacancies).

8 Conclusion

The paper analyzes regional and occupational mobility in the matching function framework. The matching function was built around the Sato (1967) three-level constant elasticity of substitution model. Within this model unemployed workers of a certain occupation and residing in a certain region could choose between vacancies for a familiar occupation located in the same region,

vacancies for an unfamiliar occupation in the same region, and vacancies for a familiar occupation in a distant region.

The results show that the unemployed in East Germany are more likely to accept a job for an unfamiliar occupation than their western counterparts. The unemployed in West Germany are more (spatially) mobile than unemployed in the East.

Specifications using spatially weighted and unweighted regional vacancies have been tried. The results using spatially weighted regional vacancies show that the loss in matchings would have been 10 percent in the absence of regional mobility. When using unweighted vacancies the loss would have been 20% for males in West Germany and 25% for males in the East.

In case of perfect regional mobility, the number of matchings could improve by 15% in the West and 5% in the East. If occupations were perfect substitutes, the number of matches could increase by 5-9%. If occupations were not substitutable the number of matching would have dropped by 3-5% in the West and 6-10% in the East.

It is also argued that the aggregation bias arises because of aggregation of inputs of the matching function with different levels of substitutability.

References

- C. Abraham. Help-wanted advertising, job vacancies, and unemployment. *Brooking Papers on Economic Activity*, 18:207–248, 1987.
- P. Anderson and S. Burgess. Empirical matching functions: Estimation and interpretation using state-level data. *The Review of Economics and Statistics*, 82(1):93–102, 2000.
- L. Anselin. Spatial econometrics. In B. Baltagi, editor, *A Companion to Theoretical Econometrics*, chapter 14, pages 310–330. Blackwell Publishing Ltd, 2000.
- A. Börsch-Supan. Regionale und sektorale Arbeitslosigkeit: durch höhere Mobilität reduzierbar? *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 110:55–82, 1990.
- A. Börsch-Supan. Panel data analysis of the beveridge curve: Is there a macroeconomic relation between the rate of unemployment and the vacancy rate? *Economica*, 58:279–297, 1991.
- M. Burda and C. Wyplosz. Gross worker and job flows in Europe. *European Economic Review*, 38:1287–1315, 1994.
- S. Burgess and S. Profit. Externalities in the matching of workers and firms in britain. *Labour Economics*, 8:313–333, 2001.
- M. Coles and E. Smith. Cross-section estimation of the matching function: Evidence from England and Wales. *Economica*, 63(252):589–597, 1996.
- R. Davidson and J. G. MacKinnon. Artificial regressions. In B. Baltagi, editor, *A Companion to Theoretical Econometrics*, chapter 1, pages 16–37. Blackwell Publishing Ltd, 2000.
- R. Davidson and J. G. MacKinnon. *Estimation and Inference in Econometrics*. Oxford University Press, New York, 1993.
- J. Dolado and F. Felgueroso. Occupational mismatch and moonlighting among spanish physicians: Do couples matter? Discussion Paper no.3419, IZA, 2008.
- J. Driscoll and A. Kraay. Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4):549–560, 1998.
- J. Duffy, C. Papageorgiou, and F. Perez-Sebastian. Capital-skill complementarity? Evidence from a panel of countries. *The Review of Economics and Statistics*, 86(1):327–344, 2004.
- R. Fahr and U. Sunde. Estimation of occupational and regional matching efficiencies using stochastic production frontier models. Discussion Paper 552, IZA, 2002.
- R. Fahr and U. Sunde. Occupational job creation: Patterns and implications. *Oxford Economic Papers*, 56:407–435, 2004.

- R. Fahr and U. Sunde. Regional dependencies in job creation: An efficiency analysis for Western Germany. *Applied Economics*, 38:1193–1206, 2006.
- W. Franz. Match and mismatch on the German labour market. In F. Padoa Schioppa, editor, *Mismatch and Labour Mobility*, chapter 3, pages 105–135. Cambridge University Press, Cambridge, 1990.
- A. Ibourk, B. Maillard, S. Perelman, and H. R. Sneessens. Aggregate matching efficiency: A stochastic production frontier approach, France 1990-1994. *Empirica*, 31:1–25, 2004.
- P. Ilmakunnas and H. Pesola. Regional labour market matching function and efficiency analysis. *Labour*, 17:413–437, 2003.
- P. Krusell, L. Ohanian, J.-V. Rios-Rull, and G. Violante. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68:1029–53, 2000.
- R. Layard, S. Nickell, and R. Jackman. *Unemployment, Macroeconomic Performance and the Labour Market*. Oxford University Press, Oxford, 1991.
- W. Newey and K. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708, 1987.
- B. Petrongolo and C. Pissarides. Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, 39(2):390–431, 2001.
- K. Sato. A two-level constant-elasticity-of-substitution production function. *Review of Economic Studies*, 34:201–218, 1967.
- M. Stevens. New microfoundations for the aggregate matching function. *International Economic Review*, 48(3):847–868, 2007.

A Appendix

A.1 Occupation Groups

Occupations are summarized in 18 broad categories:

- Agriculture
- Mining, basic manufacturing
- Textile, leather
- Food
- Construction

- Engineers, technical occupations
- Sales/trade
- Private business services
- Transport and communications
- Logistics
- Clerks
- Security, law
- Health, body care
- Social care
- Teaching, social and natural sciences
- Hotels, Restaurants and Catering
- Housekeeping services
- Cleaning and unskilled occupations

A.2 Alternative Specification

Tab. A.2: Three-level CES model estimates

Parameter ^a	Men West		Men East	
	Preferred Specification ^b	Alternative Specification ^c	Preferred Specification ^b	Alternative Specification ^c
a	0.92	0.51	0.87	0.56
b	0.58	1.00	0.57	1.00
α	0.77	0.77	0.76	0.77
θ	0.93	2.35	0.81	2.04
ρ	0.22	0.41	0.38	0.35
R^2	0.86		0.90	
N	306,373		107,767	

^a the model also contains the shares of long-term unemployed, unemployed with foreign nationality, youth and elderly unemployed, log area, as well as monthly and yearly dummies and occupation dummies.

^b preferred specification (Equation 1)

^c alternative specification (based on the Model in Equation 8)