SFB 649 Discussion Paper 2010-054

# Spatial Dependencies in German Matching Functions

Franziska Schulze*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

# Spatial Dependencies in German Matching Functions

Franziska Schulze*
Humboldt-Universität zu Berlin

This version: October 2010

**Abstract:**

This paper proposes a spatial panel model for German matching functions to avoid possibly biased and inefficient estimates due to spatial dependence. We provide empirical evidence for the presence of spatial dependencies in matching data. Based on an official data set containing monthly information for 176 local employment offices, we show that neglecting spatial dependencies in the data results in overestimated coefficients. For the incorporation of spatial information into our model, we use data on commuting relations between local employment offices. Furthermore, our results suggest that a dynamic modeling is more appropriate for matching functions.

**Keywords:** Empirical Matching, Geographic Labor Mobility, Spatial Dependence, Regional Unemployment

**JEL classification:** C21, C23, J64, J63, R12

*I would like to thank Nikolaus Hautsch, Melanie Schienle and Lada Kyj for helpful and instructive comments on this project. Financial support of the Deutsche Forschungsgemeinschaft through SFB 649 "Economic Risk" for the provision of data is gratefully acknowledged. Address: Humboldt-Universität zu Berlin, School of Economics and Business, Institute of Statistics and Econometrics, Chair of Econometrics. Spandauer Str. 1, 10178 Berlin, Germany. Email: franziska.schulze@wiwi.hu-berlin.de

# 1 Introduction

In 2009, there were about 9.25 million people that became unemployed in Germany. But, during the same time, about 9 million people left the state of inactivity while the average unemployment stock amounted to 3.42 million in 2009. These numbers illustrate that labor markets are characterized by large flows between the states of activity and inactivity. In labor market research, a standard tool to analyze these dynamics is the matching function which describes how the flow of new hires (matches) is related to the unemployment stock and to the stock of vacancies. With the help of the matching function the determinants of job creation and the structure of underlying search frictions in labor markets can be analyzed.

However, as shown in this paper, labor market activity is correlated over space. The presence of spatial (auto-)correlation implies that the extent of matching in one particular region is correlated with that in neighboring regions. Neglecting spatial correlation when modeling the matching process yields biased and inefficient estimates of the matching function. This is widely ignored in the empirical matching literature as matching functions are often specified according to models assuming cross-sectional independence among observations. This independence assumption is questionable in the labor market context due to commuting and migration between different regions.

The aim of this paper is the estimation of matching functions taking into account spatial dependencies in order to obtain unbiased and efficient estimates. For the estimation, we use an official data set that provides monthly information of 176 local employment offices (*Arbeitsagenturen*) for the period from 2000 until 2009. To exploit the panel structure of the data, we specify the matching function using a spatial panel model. As labor market data exhibits positive (temporal) autocorrelation, we apply not only a static but also a dynamic modeling.

Most of the contributions in the empirical matching literature estimate matching functions using aggregate time series (see, e.g., Blanchard and Diamond (1989), Van Ours

(1991) and Burda (1994)) as well as panel data sets (see, e.g., Burda (1993, 1994), Coles and Smith (1996) and Anderson and Burgess (2000)) without taking into account cross-sectional dependencies. The contributions by Fahr and Sunde (2001, 2005, 2006a, 2006b, 2009) also deal with data on German labor markets. The present paper updates and widens the range of their analysis by using data for the whole country of Germany covering a more recent period. To our best knowledge, only a few contributions deal with spatial dependencies in the empirical matching context as Burgess and Profit (2001), Hynninen (2005), Fahr and Sunde (2006a, b) and Dmitrijeva (2008). These authors introduce spatial interactions into their model using spatially lagged exogenous variables. This is a simple way of modeling a spatial process since there are no specific estimation techniques needed. As suggested by test results on cross-sectional dependence in the residuals of such a regression, this model does not capture the spatial autocorrelation in the data in a sufficient way. Therefore, we apply panel models including a spatial lag and a spatial error term to the matching function. Lee and Yu (2010b) propose a quasi-maximum likelihood approach for the static spatial autoregressive panel data model with fixed effects which we adopt here. For the estimation of the dynamic model, we employ the estimation methodology suggested by Lee and Yu (2010c). The application of spatial econometric methods to the context of empirical matching functions is novel in this literature.

An important component of spatial econometric modeling is the spatial weights matrix. As the amount of commuting reflects spatial relations on labor markets, we exploit a data set on commuting relations between local employment offices to construct both binary spatial weights matrices with entries zero and one and spatial weights matrices with general weights.

Our paper shows the following main results: Firstly, ignoring spatial dependencies in matching data when modeling the matching function results in overestimated matching elasticities. As the estimated matching elasticities reflect the structural features of the matching process, this finding is important. Secondly, the results suggest that compared to a static model, a dynamic approach results in a better fit of the data.

The structure of the paper is as follows: The second section presents the basic matching model while the third presents the data set and explains how the spatial weights matrix is defined. In order to motivate the spatial econometric approach, the fourth section provides test results of the (global) Moran $I$ test for spatial autocorrelation. Section five presents the econometric model and the sixth section is dedicated to the estimation results. Finally, the last section concludes.

## 2 Matching on labor markets

In macroeconomics, the matching function plays a central role for the analysis of labor market dynamics and labor market efficiency. The labor market is assumed to be a decentralized market where it takes time and resources for the unemployed persons and vacant jobs to find each other. Reasons for this complicated exchange process are trading frictions, incomplete information and heterogeneities. With the help of the matching function, this two-sided search process can be characterized. For a survey on the empirical matching literature, see Petrongolo and Pissarides (2001).

In the empirical matching literature, it is standard to use a Cobb-Douglas specification for the matching function.[1] Taking logs, the matching equation describing the flow of matches $m_{it}$ between time period $t$ and $t+1$ is given by

$$ln\, m_{it} = c_i + \alpha_t + \beta_1\, ln\, \mathcal{U}_{it} + \beta_2\, ln\, \mathcal{V}_{it} + \epsilon_{it}, \quad t = 1, \ldots, T, i = 1, \ldots, n \tag{1}$$

where $\mathcal{U}_{it}$ and $\mathcal{V}_{it}$ denote the stock of registered unemployment and the stock of regis-

[1] From a theoretical viewpoint, it is also possible to use a CES-type matching function. In this context, Burda (1994) explains that the assumption of this type of matching function does not entail additional explanatory power. Nevertheless, there are critical views concerning the Cobb-Douglas assumption for matching functions in the literature, see for example Stevens (2007).

tered vacancies at time point $t$, respectively.[2] $c_i$ is the time-invariant effect controlling for employment office-specific characteristics as, for example, its size, while $\alpha_t$ is a time effect controlling for aggregate shocks. $\epsilon_{it}$ describes the error term which is assumed to be homoskedastic and uncorrelated.

If the spatial dependence effects are ignored, standard OLS regression will provide biased parameter estimates in case of spatial lag dependence or spatiallly lagged exogenous variables whereas it provides unbiased and inefficient estimates for the spatial error model. Neglecting the spatial lag term is similar to an omitted variables bias (see Franzese and Hays (2007)). As the spatial lag term is always correlated with the errors, OLS estimation of the corresponding coefficient will be inconsistent (see Anselin and Bera (1998) or Franzese and Hays (2007)).

## 3 Data and spatial weights matrix

### 3.1 Measuring matches, unemployment and vacancies

We use the "outflows from unemployment into gainful employment" as measure for the matches which is provided by official labor statistics in Germany. The "outflows from unemployment" in general also include people entering into part-time employment, into labor-market policy measures or people leaving the labor force which we do not want to consider as successful matches. The data series of the matches as well as the data series of the unemployed persons and vacancies are available from the Federal Employment Office (*Bundesagentur für Arbeit*) on a monthly basis. Hence, we have panel data on 176 local employment offices (*Arbeitsagenturen*) for the time from 2000 until 2009. Table 1 shows the summary statistics of the labor market data. They show that there is a strong variation in the number of unemployed, vacant jobs and matches between the local employment offices. The maximum values are always attained in

[2] In order to ease notation for the spatial panel models, we differ between stocks and flows by using this notation in script for the stocks.

Table 1: Summary statistics of matches, unemployment and vacancy stock of German local employment offices (2000-2009)

| | unemployment stock | matches | vacancy stock |
|---|---|---|---|
| Min | 2,643 | 211 | 146 |
| 1st qu. | 10,920 | 805 | 1,253 |
| Median | 16,800 | 1,166 | 1,909 |
| Mean | 22,891 | 1,504 | 2,336 |
| 3rd qu. | 28,710 | 1760 | 3,022 |
| Max | 332,874 | 20,675 | 41,435 |

Source: Federal Employment Office (Bundesagentur für Arbeit)

Berlin.

Firms are not obliged to report their vacant jobs to the Federal Employment Office in Germany. Therefore, the registered vacancies represent only a fraction of the overall economic supply of vacant positions. In 2006, this fraction amounted to 44% only (see BA (2008)). The unemployment data is collected in accordance to the "concept of registered unemployment" which is regulated in the German Social Security Code. Hence, this analysis is limited to that part of the labor market which is officially registered at the Federal Employment Office. However, registered positions can also be filled with employed job searchers which are not covered in our data set. The registered unemployed and vacancies are possibly subject to a downward skill bias. On the one hand, highly-qualified persons mostly do not use the Federal Employment Office in order to find a new job. On the other, firms having vacant jobs for which a high qualification is needed prefer using web portals, national newspapers and internal channels to find suitable candidates (see Koppel (2008)). Christensen (2001) argues as well that the rate of reported vacancies is higher for jobs which require low skills.

We standardize the matches in every local employment office by corresponding unemployment stocks for the computations of the test statistics for spatial autocorrelation. Hence, the resulting standardized matches represent the fraction of unemployed persons leaving unemployment in order to start a job.

### 3.2 Time series properties

To test for the stationarity of the data, we apply panel unit root tests. The results of the Im et al. (2003) (IPS) test and the Fisher-type (ADF) test, that was proposed by Maddala and Wu (1999) and Choi (2001), clearly reject the hypothesis of a unit root in the unemployment and vacancy data as all p-values are zero. For the matches, the hypothesis of a unit root can only be rejected in case of the Fisher-type test. A more detailed description of these tests can be found in the appendix. Nevertheless, it has to be noted that the results of Baltagi et al. (2007) show that there can be considerable size distortions in panel unit root tests when the true model exhibits spatial error correlation. Therefore, the test results can only serve as an indication of possible nonstationarities in the data.

Figures 1 - 6 show some representative examples of autocorrelation function (ACF) plots for the three different variables. They show significant (temporal) autocorrelation in the data.

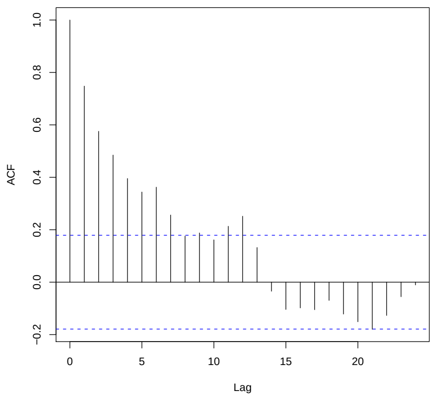



Figure 1: ACF plot of the matches in Bremen with a maximum lag length of 24 months (2000-2009)

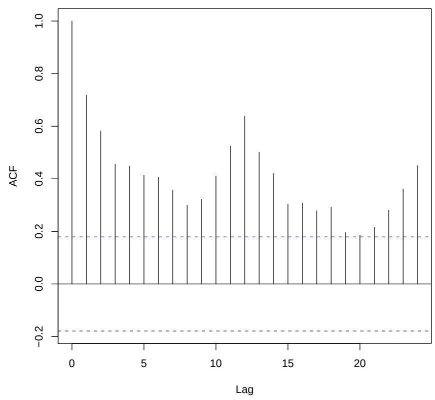



Figure 2: ACF plot of the matches in Hamburg with a maximum lag length of 24 months (2000-2009)

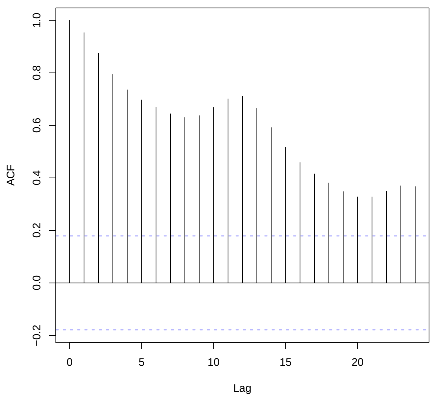



Figure 3: ACF plot of the unemployment stock in Mecklenburg-Western Pomerania with a maximum lag length of 24 months (2000-2009)

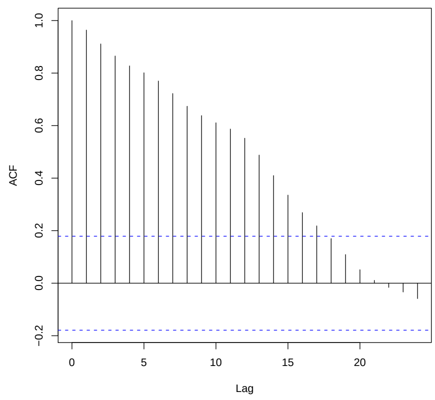



Figure 4: ACF plot of the unemployment stock in North-Rhine Westphalia with a maximum lag length of 24 months (2000-2009)

### 3.3 Specification of spatial influence

A fundamental building block of spatial econometric modeling is the spatial weights matrix. It is a nonstochastic matrix which defines exogenously the neighborhood of a certain location. Hence, the term ‚neighboring' in the present context addresses the neighborhood set which is defined by the corresponding spatial weights matrix. On the one hand, we use binary spatial weight matrices where the entries are either zero or one and, on the other, matrices with general weights.

The simplest version of a binary spatial weights matrix is a binary contiguity matrix. When two local employment offices are neighbors, i.e. they share a common border, the corresponding entry in the matrix is one and zero otherwise. The elements on the main diagonal are zero by construction. This matrix induces a simple spatial structure which might be not sufficient to capture the actual spatial relations on German labor markets. Commuting of people is not limited to the neighboring region and, additionally, the binary contiguity matrix weights all neighbors equally. The latter assumption might be critical for a region that is surrounded by both a big city and a rural area. In this case, one would suspect that more people commute to the big city than to the rural area.

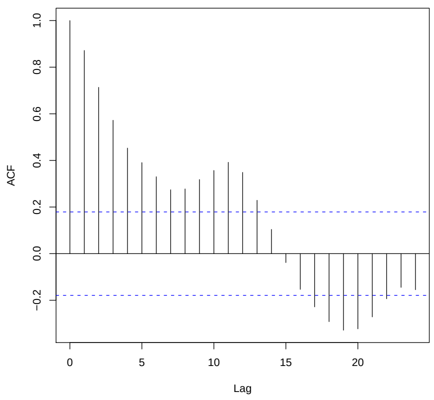



Figure 5: ACF plot of the vacancy stock in Mecklenburg-Western Pomerania with a maximum lag length of 24 months (2000-2009)

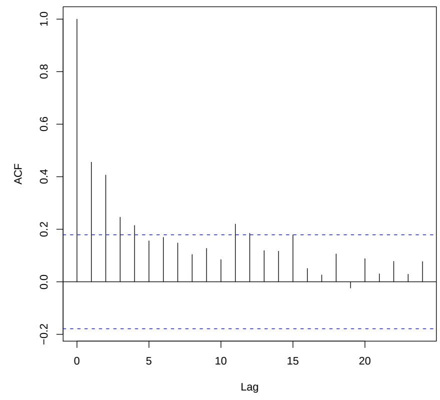



Figure 6: ACF plot of the vacancy stock in Saxony with a maximum lag length of 24 months (2000-2009)

To address these problems, we exploit a data set on commuting relations between different local employment offices. The amount of commuting reflects differences in labor market opportunities between local employment offices. Möller and Aldashev (2007), who also use commuter streams for constructing a spatial weights matrix, state that such a matrix captures the strength of interregional relationships among labor markets. The commuting data is also collected by the Federal Employment Office and it is part of the official statistic as well. It records all people who have a job that is subject to social insurance. The numbers of commuters between the local employment offices are recorded yearly at the appointed date June 30th. Our data set covers the period from 2000 until 2009. We use this data as a proxy for the interregional linkages between local employment offices. Therefore, we construct the average commuter matrix $P = (p_{ij})$, $i, j = 1, \dots, 176$. The element $p_{ij}$ of this matrix indicates the number of people that live in employment office j and work in employment office i. Hence, row $i$ of the average commuter matrix $P$ contains the incoming commuters to region $i$ while the elements of column $j$ represent the outgoing commuters from employment office $j$ to all other regions.[3] The summary statistics of matrix $P$ are found in Table 2. They show that there is commuting between most of the local employment offices, although it is not very strong between 75% of them. The highest numbers of incoming commuters are in big cities as Berlin, Düsseldorf, Cologne, Frankfurt, Munich and Hamburg.

Table 2: Summary statistics of the average commuter matrix $P$ (2000-2009)

| Min | 1st quartile | Median | Mean | 3rd. quartile | Max | Std. dev. |
|---|---|---|---|---|---|---|
| 0 | 6.2 | 16.3 | 216.02 | 45.8 | 60201.5 | 1583.24 |

Source: Federal Employment Office, author's calculations

We use the commuting information twofold to construct both binary weights matrices and weights matrices with general weights. First, we discretize the information and construct additional binary spatial weights matrices. We consider two local employment offices as neighbors when the commuting flow from region $j$ to $i$ exceeds a certain value $\delta$. Hence, the entries of the spatial weights matrix are defined by

$$w_{ij} = \begin{cases} 0, & p_{ij} < \delta \\ 1, & p_{ij} \geq \delta. \end{cases} \tag{2}$$

For the choice of the cutoff value $\delta$, we take higher values, namely 100, 250, 500 and 1000, in order to consider only those local employment offices as neighbors which exhibit strong commuting linkages. The resulting spatial weights matrix still weights all neighbors equally, yet it reflects the actual connections on the labor market in a better way by not restricting the analysis to physical neighbors.

Second, we exploit the full information contained in the average commuter matrix $P$ to construct a spatial weights matrix with general weights. Contrary to the applied spatial econometric literature where a distance decay function is often assumed, we need a monotonically increasing function as more intense commuting implies stronger spatial

[3]Due to limited data availability, we could not control for the different sizes of the local employment offices by standardizing the number of commuters by the corresponding working age population.

influence. We use the linear function for the specification of the weights, i.e. $w_{ij} = p_{ij}$.[4] This function implies that the marginal influence of one additional commuter is constant.

When computing spatially lagged matches, i.e. $Wln(M_{it})$, with the help of the general spatial weights matrix, the neighboring matches are weighted by the rows of the spatial weights matrix, i.e. by incoming commuters to region $i$. In general, this weighting scheme can be changed so that the neighboring matches are weighted by the outgoing commuters of region $i$. We also implemented this weighting scheme in our regressions and got virtually the same results.

# 4 Spatial dependencies in German labor markets

## 4.1 Empirical evidence on (global) spatial autocorrelation

A standard test for spatial autocorrelation is the Moran $I$ test, which was developed by Moran (1950). This test is not specified for a particular spatial process. Its null hypothesis is the absence of spatial autocorrelation whereas the alternative is not exactly specified. The test statistic can be expressed by

$$I = \frac{n}{S_0} \frac{e'We}{e'e}, \tag{3}$$

where $e = y - X\tilde{\beta}$ is a vector of standard OLS regression residuals, $\tilde{\beta} = (X'X)^{-1}X'y$, $W$ denotes the spatial weights matrix and $n$ is the number of observations (see Anselin and Bera (1998)). In our case $y$ are the matches and the matrix $X$ contains the unemployment and vacancy stock. $S_0$ is a standardization factor which is equal to the sum of the spatial weights, i.e. $S_0 = \sum_i \sum_j w_{ij}$. For a row-standardized spatial weights matrix, the statistic

[4] We also considered the logarithmic function to construct the weights. It produces results that give an indication for the spatial process to be nonstationary. This finding is supported by high values of the global Moran I statistic which, according to Fingleton (1999), can serve as an indicator for spatial nonstationarity.

becomes

$$I = \frac{e'We}{e'e}, \tag{4}$$

since $S_0 = n$. Cliff and Ord (1981) show that $I$ is asymptotically normally distributed for normally distributed regression residuals. Therefore, inference is based on the standard normal variate $z(I)$ which is yielded by the transformation $z(I) = \frac{I-E(I)}{\sqrt{V(I)}}$. The expectation $E(I)$ and the variance $V(I)$ are derived by Cliff and Ord (1972) under the null hypothesis of no spatial dependence.

Since Moran's $I$ test is designed to detect spatial autocorrelation from cross-section residuals, the test statistic is computed using standardized matches for each month within the period from 2000 until 2009. The values of the global Moran $I$ statistic are positive and significant on all reasonable significance levels for all months within the period. Hence, we conclude that the regional distribution of standardized matches in Germany is characterized by strong spatial dependencies. Figure 7 shows the evolution of the Moran $I$ values for standardized matches, its nine-month moving average and a linear trend line using the binary spatial weights matrix over the period from 2000 until 2009.

The linear trend is positive which means that spatial integration of German local employment offices becomes stronger during the period from 2000 until 2009. A reason for this is the increased mobility of people which is also supported by our commuting data. The mean relative change of (incoming) commuters between 2000 and 2009 amounts to 0.17, i.e. commuting increases significantly during this period.

Furthermore, Figure 7 shows a seasonal pattern in the (global) Moran values. Similarly to the seasonalities in the matching data, the Moran values are the highest during spring time. Hence, spatial dependencies seem to be stronger when the labor market is more active. Burgess and Profit (2001) analyze the cyclical variation of spatial dependence in the matching context using British data and find that the intensity of spatial

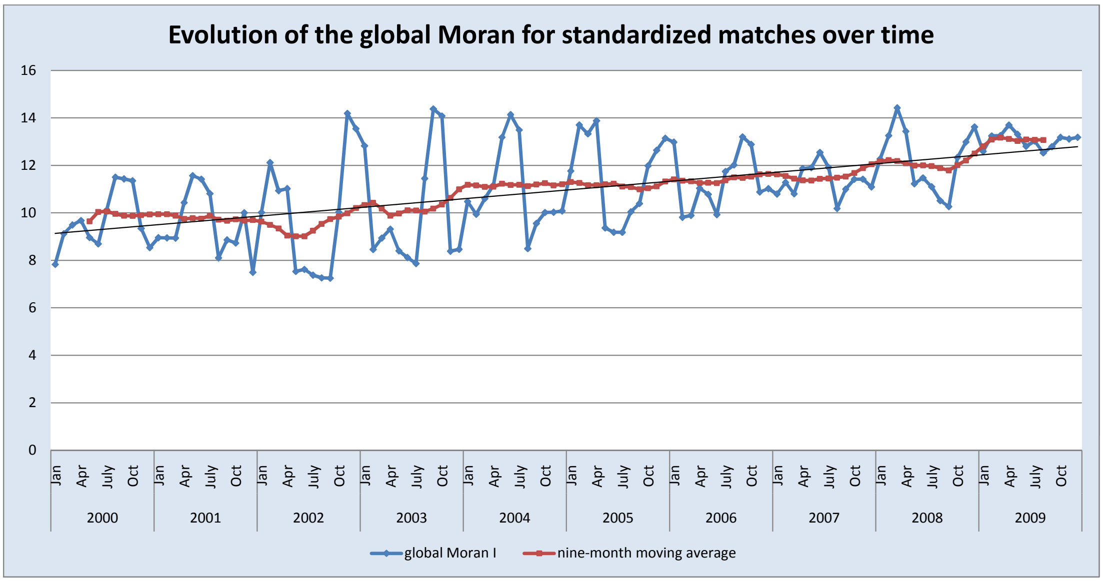



Figure 7: Evolution of the global Moran I for standardized matches for the period from 2000 until 2009

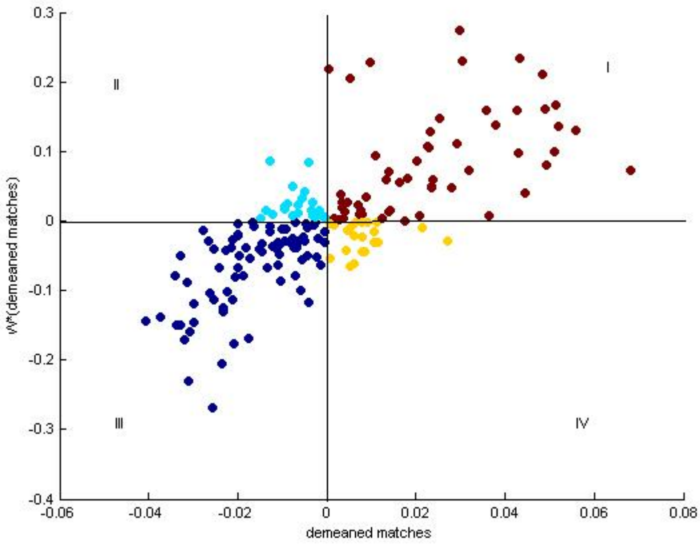
dependence for unemployment outflows moves counter-cyclically. Their explanation is that unemployed persons lower their search radius while firms have to search more widely in good times.

### 4.2 Local structure of spatial autocorrelation

The (global) Moran $I$ test only gives information about the global pattern of spatial dependence holding for all local employment offices in Germany. In order to analyze the local pattern of spatial autocorrelation, we compute Moran scatter plots (see Anselin (1996)). These are based on the interpretation of the Moran $I$ statistic as a regression coefficient in a regression of $Wy$ on $y$ where $y$ denotes demeaned matches in the present analysis. In order to show this, Anselin (1996) rearranges the Moran $I$ statistic to get the expression

$$I = \frac{y'Wy}{y'y} \tag{5}$$

which holds for a row-standardized spatial weights matrix, i.e. $S_0 = n$. Using the interpretation of the (global) Moran $I$ statistic as a regression coefficient, the linear relationship between $y$ and $Wy$ can be visualized by a bivariate scatter plot of $Wy$ against $y$.

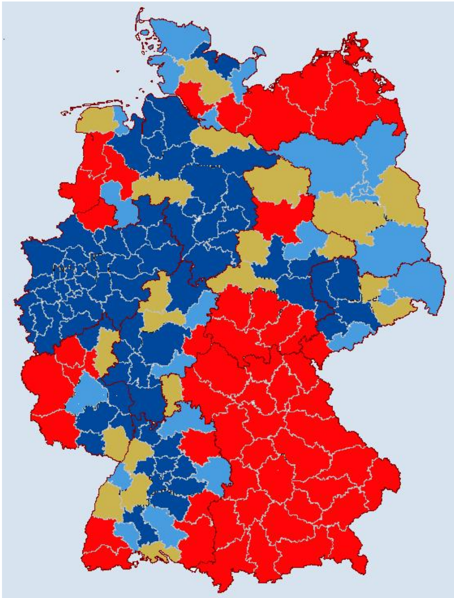



Note: This figure plots spatially lagged matches against demeaned matches. The colors identify the location of the points in the four quadrants. They are used to show where the points are located on the German map (see Figure 9).

Figure 8: Moran scatter plot using yearly averages of standardized matches in 2009

The Moran scatter plot for the standardized matches (yearly averages) of the year 2009 is shown in Figure 8. The figures for the remaining years can be found in the appendix (Figures 10 – 12).

The Moran scatter plots show that most of the local employment offices are positively spatially autocorrelated since most of the points lie in the first and third quadrant. This is in line with the results of the global Moran $I$ test. The position of local employment offices in the first and third quadrant (red and dark blue points) indicates that local employment offices with above-mean matches have neighbors with the same characteristic, while local employment offices with below-mean matches are more likely to be surrounded by local employment offices with low (standardized) matches. The re-



Note: The colors indicate the four quadrants of the Moran scatter plot as shown in Figure 8.

Figure 9: German map indicating the position of points in Moran scatter plot for yearly averages of standardized matches (2009)

maining points in the second and fourth quadrant represent local employment offices which are negatively spatially autocorrelated.

Figure 9 shows a map of Germany indicating the location of the points in the Moran scatter plot (figure 8). Maps for the other years are found in the appendix (Figures 13 – 15).

Interestingly, the Moran maps do not replicate the former border between East and West Germany. Furthermore, they show a band of local employment offices from Western (North-Rhine Westphalia) to Eastern Germany (Brandenburg) which seems to be positively spatially autocorrelated with standardized matches below the mean for the most part of the period. Most of the south German employment offices are positively spatially autocorrelated with above-mean matches. Finally, the Moran scatter plots again support the fact that German matching data exhibits spatial autocorrelation. Thus, we have to take into account this fact into our econometric analysis.

# 5 Econometric Modeling

To capture the spatial dependence and the panel structure of the data, we propose to model the matching function by a spatial panel model. Since we do not have a representative sample of German employment offices but data on all German local employment offices, a fixed effects model is preferred. In order to control for aggregate shocks, a model that takes into account time effects is used. Following most contributions in the empirical matching literature, we use a static specification of the matching function. Beyond that, we also specify the matching function in a dynamic way to capture the (temporal) autocorrelation of the data.

## 5.1 Static model specification

Our static model specification contains a spatial lag of the dependent variable as well as a spatial process for the error term. The corresponding matching equation is given by

$$\begin{aligned} ln\, M_t &= \lambda\, W\, ln\, M_t + \beta_1\, ln\, \mathcal{U}_t + \beta_2\, ln\, \mathcal{V}_t + c_n + \alpha_t \mathbb{1}_n + \Omega_t, \\ \Omega_t &= \rho W \Omega_t + \Xi_t, t = 1, \ldots, T, \end{aligned} \tag{6}$$

where $M_t = (m_{1t}, m_{2t}, \ldots, m_{nt})'$ is the $(n \times 1)$ vector of matches, $\mathcal{U}_t$ and $\mathcal{V}_t$ are the $(n \times 1)$ vectors of the unemployment and vacancy stocks, respectively. $c_n$ represents the $(n \times 1)$ vector of fixed individual effects and $\alpha_t$ is the fixed time effect. $W$ is the $(n \times n)$ nonstochastic spatial weights matrix and $\mathbb{1}_n$ is the $(n \times 1)$ vector of ones. $\Xi_t = (\xi_{1t}, \xi_{2t}, \ldots, \xi_{nt})'$ represents the $(n \times 1)$ vector of errors for which it is assumed that $\xi_{it}$ are i.i.d. across i and t with zero mean and constant variance $\sigma^2$.

A spatial error term implies that there are spatially correlated random components influencing a region of more than one local employment offices. Examples in the labor market context are regional shocks as changes in regional governments or the closure of a production site. The spatial lag structure implies that the matching process in a particular local employment office is influenced by matching in other locations. As matching theory suggests, the matches are determined by the unemployment and vacancy stock. Thus, the spatial influence of all variables is captured implicitly by using the spatial lag model.

The way of inserting spatial autocorrelation into the matching function goes beyond most matching specifications in the literature that control for spatial dependencies because we use a spatial lag and spatial error term in our static model. Contrary to this, in the empirical matching literature spatial dependencies are incorporated by spatially lagged exogenous variables into matching functions. In this way the external effect of unemployment and vacancies on the matching process in neighboring local employment offices can be estimated. As these additional regressors are exogenous and as the error term remains spherical, estimation by ordinary least squares is unbiased and consistent (see Klotz (2004) for the pooled case). We also adopt this modeling to our data and got insignificant spatial spillovers of both stock variables. Moreover, we apply Pesaran's CD test (see Pesaran (2004)) to test for cross-sectional dependence in the residuals. The results show that there is spatial correlation left in the residuals, i.e. the model with spatially lagged exogenous variables incompletely captures the spatial autocorrelation in the data.[5]

Lee and Yu (2010b) propose a quasi-maximum likelihood approach for the estimation of model (6). They show that (direct) maximum likelihood estimation yields inconsistent parameter estimates (unless $n$ is large). Even in the case when $n$ and $T$ are large, the asymptotic distribution of the estimates is not properly centered. Therefore, they propose a transformation approach to eliminate the individual and time effects. The transformations are the deviation from time mean, $J_T = I_T - \frac{1}{T}\mathbb{1}_T\mathbb{1}_T'$, and the devia-

[5]The full results of the estimation and the test can be obtained from the author upon request.

tion from cross section mean, $J_n = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n'$, operator as used in the literature on panel data analysis (see Baltagi (2005)). The disturbance terms in the resulting equation (after performing these operations) would be linearly dependent. For this reason, their proposition is to base the transformations on the orthonormal eigenvector matrices of $J_T$ and $J_n$. Let $[F_{T,T-1}, \frac{1}{\sqrt{T}}\mathbb{1}_T]$ be the orthonormal eigenvector matrix of $J_T$ where $F_{T,T-1}$ is the $(T \times (T-1))$ submatrix corresponding to eigenvalues of one. Furthermore, let $[F_{n,n-1}, \frac{1}{\sqrt{n}}\mathbb{1}_n]$ be the orthonormal eigenvector matrix of $J_n$ where $F_{n,n-1}$ is the $(n \times (n-1))$ submatrix corresponding to eigenvalues of one. The matching function (6) is firstly transformed by $F_{T,T-1}$ which yields

$$\begin{aligned} (ln\, M_t)^* &= \lambda W (ln\, M_t)^* + \beta_1 (ln\, \mathcal{U}_t)^* + \beta_2 (ln\, \mathcal{V}_t)^* + \alpha_t^* \mathbb{1}_n + \Omega_t^* \\ \Omega_t^* &= \rho W \Omega_t^* + \Xi_t^*, t = 1, \ldots, T-1, \end{aligned} \tag{7}$$

where $(ln\, M_t)^* = [ln\, M_{n1}, \ldots, ln\, M_{nT}]F_{T,T-1}$ and $[\alpha_1^*\mathbb{1}_n, \alpha_2^*\mathbb{1}_n, \ldots, \alpha_{T-1}^*\mathbb{1}_n] = [\alpha_1\mathbb{1}_n, \alpha_2\mathbb{1}_n, \ldots, \alpha_T\mathbb{1}_n]'F_{T,T-1}$ are transformed time effects. Secondly, in order to eliminate the time effects, the model is further transformed by $F_{n,n-1}$ yielding a $(n-1)$-dimensional vector $(ln\, M_t)^{**}$ such that $(ln\, M_t)^{**} = F_{n,n-1}'(ln\, M_t)^*$, i.e.

$$\begin{aligned} (ln\, M_t)^{**} &= \lambda (F_{n,n-1}' W F_{n,n-1}) (ln\, M_t)^{**} + \beta_1 (ln\, \mathcal{U}_t)^{**} + \beta_2 (ln\, \mathcal{V}_t)^{**} + \Omega_t^{**} \\ \Omega_t^{**} &= \rho (F_{n,n-1}' W F_{n,n-1}) \Omega_t^{**} + \Xi_t^{**}, t = 1, \ldots, T-1, \end{aligned} \tag{8}$$

where $(ln\, \mathcal{U}_t)^{**} = F_{n,n-1}'(ln\, \mathcal{U}_t)^*$ and $(ln\, \mathcal{V}_t)^{**} = F_{n,n-1}'(ln\, \mathcal{V}_t)^*$. Note that the effective sample size after both transformations is $(n-1)(T-1)$ and that the spatial weights matrix needs to be row-normalized for this transformation approach.

The transformed equation (8) can be estimated by quasi-maximum likelihood. After some rearrangements, Lee and Yu (2010b) derive the following log-likelihood function for the transformed model (8):

$$ln\, L_{n,T}(\theta) = -\frac{(n-1)(T-1)}{2} ln\, 2\pi\sigma^2 - (T-1)[ln(1-\lambda) + ln(1-\rho)] \\ + (T-1)[ln|S_n(\lambda)| + ln|R_n(\rho)|] - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \widetilde{\Xi}_t'(\phi) J_n \widetilde{\Xi}_t(\phi) \quad (9)$$

where $\theta = (\beta', \lambda, \rho, \sigma^2)$, $\beta' = (\beta_1, \beta_2)'$, $\phi = (\beta', \lambda, \rho)'$, $S_n(\lambda) = I_n - \lambda W$, $R_n(\rho) = I_n - \rho W$ and $\widetilde{\Xi}_t = R_n(\rho)[S_n(\lambda)\widetilde{ln\, M_t} - (\widetilde{ln\, \mathcal{U}_t}, \widetilde{ln\, \mathcal{V}_t})\beta]$. Note that $\widetilde{ln\, M_t} = ln\, M_t - \overline{ln\, M_t}$ for $t = 1, \ldots, T$, where $\overline{ln\, M_t} = \frac{1}{T}\sum_{t=1}^{T} ln\, M_t$. $\widetilde{ln\, \mathcal{U}_t}$, $\widetilde{ln\, \mathcal{V}_t}$ and $\widetilde{\Xi}_t$ are defined analogously.

Lee and Yu (2010b) show that the resulting quasi maximum-likelihood estimates for all parameters are consistent when either $n \to \infty$ or $T \to \infty$ and asymptotically normally distributed. Additionally, they derive explicitly the asymptotic distribution and show that it is properly centered.

### 5.2 Dynamic model specification

As shown in the data section, labor market data exhibits positive temporal autocorrelation. To capture these dynamics, we apply a spatial dynamic panel data model. In addition to a temporally lagged term, it contains a spatial lag term and a combined spatially and temporally lagged term of the dependent variable. Applying this model to our matching function, yields

$$ln\, M_t = \lambda\, W\, ln\, M_t + \gamma\, ln\, M_{t-1} + \delta\, W\, ln\, M_{t-1} + \beta_1\, ln\, \mathcal{U}_t + \beta_2\, ln\, \mathcal{V}_t \\ + c_n + \alpha_t \mathbb{1}_n + \Xi_t, t = 1, \ldots, T, \quad (10)$$

where $\gamma$ captures the pure time-dynamic effect and $\delta$ captures the combined spatial-time effect. The assumptions about the error term $\Xi_t$ are as before.

For the estimation of model (10) we adopt the methodology proposed in Lee and

Yu (2010a) and Lee and Yu (2010c). Lee and Yu (2010a) show that the (direct) maximum likelihood estimation method will yield a bias of order $O(max(1/n, 1/T))$ for the common parameters. Therefore, they propose two variants of a transformation approach. On the one hand, the transformation $J_n$ in combination with an eigenvalue and eigenvector decomposition is applied and, on the other, the model is transformed by $(I_n - W)$. Lee and Yu (2010a) show that the quasi-maximum likelihood estimates from the maximization of the log-likelihood function of the $J_n$-transformed model are free of $O(1/n)$ bias. Nevertheless, the resulting quasi-maximum likelihood estimates are biased and, therefore, Lee and Yu (2010c) propose a bias correction procedure which is applied here as well.

The $(I_n - W)$-transformation eliminates not only time effects but also possible unstable components. Thus, it can be applied to all possible data generating processes. We applied both transformations to our data. But as the results are fairly similar and in order to save space, we present only the results and theoretical foundations of the $(I_n - W)$-transformation.

Transforming the dynamic matching equation (10) by $(I_n - W)$, yields

$$(I_n - W)ln\, M_t = \lambda W(I_n - W)ln\, M_t + \gamma(I_n - W)ln\, M_{t-1} + \delta W(I_n - W)ln\, M_{t-1}$$
$$+ (I_n - W)\mathcal{X}_t\beta + (I_n - W)c_n + (I_n - W)\Xi_t,\ t = 1, \ldots, T, \quad (11)$$

where $\mathcal{X}_t = [ln\, \mathcal{U}_t, ln\, \mathcal{V}_t]$. The variance-covariance matrix of $(I_n - W)\Xi_t$ is given by

$$Var((I_n - W)\Xi_t) = \sigma^2\Sigma_n \quad (12)$$

with $\Sigma_n = (I_n - W)(I_n - W)'$. As the components of the error term in the transformed model (11) are linearly dependent, an eigenvalue-eigenvector decomposition is used again. For that, the matrix $[F_n, H_n]$ is defined to be the orthonormal matrix of eigenvectors and $\Lambda_n$ is defined to be the diagonal matrix of nonzero eigenvalues of $\Sigma_n$ such that $\Sigma_n F_n = F_n \Lambda_n$ and $\Sigma_n H_n = 0$. The columns of $F_n$ consist of eigenvectors corresponding to nonzero eigenvalues, and those of $H_n$ are for zero eigenvalues of $\Sigma_n$. According to Lee and Yu (2010c), the transformed spatial weights matrix is defined as $W^* = \Lambda_n^{-1/2} F_n' W F_n \Lambda_n^{-1/2}$. Then, the (further) transformed model is given by

$$(ln\, M_t)^* = \lambda W^* (ln\, M_t)^* + \gamma (ln\, M_{t-1})^* + \delta W^* (ln\, M_{t-1})^* + \mathcal{X}_t^* \beta + c_n^* + \Xi_t^*, \quad t = 1, \ldots, T, \quad (13)$$

where $(ln\, M_t)^* = \Lambda_n^{-1/2} F_n' (I_n - W) ln\, M_t$ and the other variable are defined accordingly. Note that the transformed vector $(ln\, M_t)^*$ is of dimension $n^*$ where $n^*$ is the rank of $\sigma^2 \Sigma_n$. The concentrated log-likelihood of equation (13) is

$$ln\, L_{n,T}(\theta) = -\frac{n^* T}{2} ln\, 2\pi - \frac{n^* T}{2} ln\, \sigma^2 - (n - n^*) T ln(1 - \lambda) + T ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \widetilde{\Xi}_t'(\theta) (I_n - W)' \Sigma_n^+ (I_n - W) \widetilde{\Xi}_t(\theta) \quad (14)$$

where $\widetilde{\Xi}_t(\theta) = S_n(\lambda) \widetilde{ln\, M_t} - \widetilde{Z}_t \vartheta$ with $Z_t = (ln\, M_{t-1}, W ln\, M_{t-1}, \mathcal{X}_t)$ and $\vartheta = (\gamma, \delta, \beta')$.

## 6 Estimation Results

In order to improve the success of the Federal Employment Office in placing unemployed persons in a job, the German government passed different laws to reform the German labor market during the period from 2002 until 2005 ("Hartz reforms"). Since one part of these reforms (becoming effective in 2004) entailed changes in the official definition of unemployment, we analyze the periods from 2000 until 2004 and from

2005 until 2009 separately.

Firstly, we estimate the basic matching model without any spatial terms. It is specified according to a two-way fixed effects model and it is estimated using the standard within-estimator.[6] Secondly, we estimate the static matching specification and, thirdly, the dynamic matching model, both using the different spatial weights matrices that we defined before. As we have six different specifications for the spatial weights matrix, we have 13 regressions for each period. The regression results are shown in Tables 3 and 5 for the period from 2000 until 2004 while the results for the second period are in Tables 4 and 6.

As suggested by matching theory, the estimated elasticities of matches on both stocks are positive and significant in all specifications. The elasticity of matches on unemployment for the basic model during the period from 2000 until 2009 amounts to 0.599. This means that an increase of the unemployment stock by 1% results in an increase of matching by 0.599 percent. The estimated elasticities with respect to vacancies are much smaller than those with respect to the unemployed in all specifications. This finding might be related to the underreporting of vacant jobs to the Federal Employment Office. Furthemore, it can be explained by the high vacancy turnover, i.e. vacant jobs are filled within a month, and thus, are not counted in the end-of-month stocks.

Ignoring spatial effects, the elasticities on both stock variables are overestimated. This result is particularly relevant since the matching elasticities give an indication of the relative importance of unemployment and vacancies in the matching process. The Bayesian information criterion also indicates a better fit of the spatial panel models. The existence of this bias is theoretically shown in Franzese and Hays (2007). They argue that neglecting a spatial lag process results in an omitted-variable bias. Fahr and Sunde (2006b) and Fahr and Sunde (2006a) also using German data get for the elasticity on unemployment 0.41 and 0.54 and for the elasticity on vacancies 0.513 and 0.34, respec-

[6] For more details on this subject, see for example Baltagi (2005).

Table 3: Estimates of matching functions using the basic and the static spatial panel model for the period from 2000 until 2004

| dependent variable: $ln\ M_{it}$ time period: 2000-2004 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | basic | static | | | | | |
| | | binary | $\delta = 100$ | $\delta = 250$ | $\delta = 500$ | $\delta = 1000$ | linear |
| $ln\ \mathcal{U}_{it}$ | 0.599 | 0.333 | 0.407 | 0.387 | 0.395 | 0.378 | 0.362 |
| | (30.9) | (16.8) | (21.43) | (19.9) | (20.17) | (19.2) | (18.41) |
| $ln\ \mathcal{V}_{it}$ | 0.092 | 0.062 | 0.071 | 0.066 | 0.063 | 0.064 | 0.067 |
| | (16.4) | (11.32) | (13.84) | (12.49) | (11.71) | (11.66) | (12.39) |
| $\lambda$ | — | 0.035 | 0.184 | 0.091 | 0.062 | 0.045 | 0.051 |
| | | (3.71) | (12.94) | (7.62) | (5.32) | (3.91) | (4.52) |
| $\rho$ | — | 0.054 | 0.206 | 0.114 | 0.083 | 0.064 | 0.07 |
| | | (5.4) | (13.78) | (9.1) | (6.8) | (5.31) | (5.9) |
| $\sigma^2$ | 0.021 | 0.019 | 0.017 | 0.018 | 0.019 | 0.019 | 0.019 |
| | | (84.64) | (74.91) | (87.51) | (82.9) | (80.95) | (82.59) |
| log-like | 5624.658 | 5737.146 | 6536.617 | 6113.52 | 5913.513 | 5783.096 | 5815.381 |
| BIC | -0.857 | -1.085 | -1.236 | -1.156 | -1.18 | -1.093 | -1.099 |
| observations | 10560 | 10560 | 10560 | 10560 | 10560 | 10560 | 10560 |

Notes: t-statistics in parentheses. t-statistics of the static spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010b). $\lambda$ is the spatial autoregressive coefficient and $\rho$ is the spatial autocorrelation coefficient.

tively. Our smaller matching elasticities can be explained by the fact that our models control for spatial, time-dynamic and combined space-time effects separately. Note that the data source and the period of their data differ from ours as well.

The spatial autoregressive ($\lambda$) and the spatial autocorrelation ($\rho$) coefficient measuring the spatial effects in our model are significant and positive. Hence, the number of matches in the neighborhood influences the matching process in a particular local employment office. The positive spatial autocorrelation coefficient indicates regional effects that affect the matching process in more than one local employment office. The effect of the spatial error term is stronger than that of the spatial lag term for the static model, i.e. spatially correlated random components play an important role on German labor markets.

Furthermore, the estimation results show that the matching elasticities are fairly robust with respect to the choice of the spatial weights matrix. Notably, this holds for the vacancies in all specifications. Likewise, the time-dynamic effect ($\gamma$) is not sensitive to different spatial regimes as well. However, this is not true for the estimates of the spa-

Table 4: Estimates of matching functions using the basic and the static spatial panel model for the period from 2005 until 2009

| dependent variable: $ln\ M_{it}$ time period: 2005-2009 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | basic | static | | | | | |
| | | binary | $\delta = 100$ | $\delta = 250$ | $\delta = 500$ | $\delta = 1000$ | linear |
| $ln\ \mathcal{U}_{it}$ | 0.697 | 0.519 | 0.575 | 0.545 | 0.539 | 0.536 | 0.533 |
| | (39.96) | (29.05) | (33.33) | (30.95) | (30.44) | (30.16) | (29.97) |
| $ln\ \mathcal{V}_{it}$ | 0.1 | 0.074 | 0.083 | 0.08 | 0.082 | 0.081 | 0.077 |
| | (17.5) | (13.49) | (15.98) | (15.06) | (15.11) | (14.86) | (14.06) |
| $\lambda$ | — | 0.032 | 0.182 | 0.086 | 0.058 | 0.043 | 0.047 |
| | | (2.8) | (10.7) | (6.24) | (4.4) | (3.31) | (3.57) |
| $\rho$ | — | 0.047 | 0.201 | 0.106 | 0.076 | 0.059 | 0.063 |
| | | (3.93) | (11.03) | (7.21) | (5.43) | (4.33) | (4.52) |
| $\sigma^2$ | 0.019 | 0.019 | 0.016 | 0.018 | 0.018 | 0.018 | 0.018 |
| | | (108.67) | (72.29) | (116.81) | (111.7) | (108.13) | (108.08) |
| log-like | 5936.595 | 5933.831 | 6558.874 | 6227.45 | 6084.772 | 5996.1 | 5997.738 |
| BIC | -0.916 | -1.122 | -1.24 | -1.178 | -1.151 | -1.133 | -1.134 |
| observations | 10560 | 10560 | 10560 | 10560 | 10560 | 10560 | 10560 |

Notes: t-statistics are in parentheses. t-statistics of the static spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010b). $\lambda$ is the spatial autoregressive coefficient and $\rho$ is the spatial autocorrelation coefficient.

Table 5: Estimates of matching functions using the spatial dynamic panel model for the period from 2000 until 2004

| dependent variable: $ln\ M_{it}$ period: 2000-2004 | | | | | | |
|---|---|---|---|---|---|---|
| | dynamic | | | | | |
| | (I-W)-transformation (after bias correction) | | | | | |
| | binary | $\delta = 100$ | $\delta = 250$ | $\delta = 500$ | $\delta = 1000$ | linear |
| $ln\ \mathcal{U}_{it}$ | 0.285 | 0.409 | 0.383 | 0.376 | 0.368 | 0.382 |
| | (22.13) | (26.61) | (24.4) | (23.9) | (23.37) | (24.19) |
| $ln\ \mathcal{V}_{it}$ | 0.047 | 0.055 | 0.051 | 0.049 | 0.048 | 0.05 |
| | (12.9) | (12.77) | (11.71) | (11.15)) | (10.83) | (11.37) |
| $\lambda$ | 0.638 | 0.214 | 0.119 | 0.086 | 0.065 | 0.071 |
| | (72.65) | (9.38) | (5.76) | (4.58) | (3.89) | (3.82) |
| $\gamma$ | 0.466 | 0.51 | 0.501 | 0.502 | 0.503 | 0.508 |
| | (53.63) | (52.96) | (49.95) | (50.03) | (50.56) | (49.31) |
| $\delta$ | -0.231 | -0.004* | 0.066 | 0.081 | 0.088 | 0.071 |
| | (-17.81) | (-017) | (3.17) | (4.17) | (4.92) | (3.63) |
| $\sigma^2$ | 0.009 | 0.012 | 0.012 | 0.012 | 0.013 | 0.013 |
| | (70.92) | (54.52) | (51.59) | (52.48) | (54.15) | (52.56) |
| log-like | 9618.98 | 8423.502 | 8223.439 | 8117.249 | 8046.945 | 8049.474 |
| BIC | -1.82 | -1.593 | -1.555 | -1.535 | -1.522 | -1.522 |
| observations | 10559 | 10559 | 10559 | 10559 | 10559 | 10559 |

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010c). $\lambda$ is the spatial autoregressive coefficient, $\gamma$ captures the time dynamic effect and $\delta$ captures the combined spatial-time effect.

Table 6: Estimates of matching functions using the spatial dynamic panel model for the period from 2005 until 2009

| dependent variable: $ln\ M_{it}$ period: 2005-2009 | | | | | | |
|---|---|---|---|---|---|---|
| | dynamic | | | | | |
| | (I-W)-transformation (after bias correction) | | | | | |
| | binary | $\delta = 100$ | $\delta = 250$ | $\delta = 500$ | $\delta = 1000$ | linear |
| $ln\,\mathcal{U}_{it}$ | 0.365 | 0.432 | 0.406 | 0.39 | 0.368 | 0.37 |
| | (27.96) | (29.88) | (27.67) | (26.43) | (27.9) | (28.2) |
| $ln\ \mathcal{V}_{it}$ | 0.039 | 0.041 | 0.038 | 0.037 | 0.04 | 0.039 |
| | (9.62) | (9.15) | (8.41) | (8.04) | (9.84) | (9.6) |
| $\lambda$ | 0.538 | 0.206 | 0.109 | 0.078 | 0.577 | 0.621 |
| | (53.54) | (9.34) | (5.71) | (4.5) | (52.06) | (52.53) |
| $\gamma$ | 0.507 | 0.513 | 0.51 | 0.515 | 0.514 | 0.512 |
| | (61.14) | (57.81) | (55.69) | (55.97) | (62.68) | (61.98) |
| $\delta$ | -0.246 | -0.019* | 0.045 | 0.051 | -0.285 | -0.316 |
| | (-18.2) | (-0.82) | (2.24) | (2.73) | (-19.87) | (-20.43) |
| $\sigma^2$ | 0.01 | 0.012 | 0.012 | 0.012 | 0.01 | 0.01 |
| | (71.16) | (61.04) | (58.77) | (58.94) | (71.42) | (71.32) |
| log-like | 9148.268 | 8494.757 | 8322.28 | 8234.828 | 9107.55 | 9114.678 |
| BIC | -1.731 | -1.607 | -1.574 | -1.557 | -1.722 | -1.724 |
| observations | 10559 | 10559 | 10559 | 10559 | 10559 | 10559 |

Notes: t-statistics are in parentheses. t-statistics of the dynamic spatial panel model are computed using the asymptotic distribution derived in Lee and Yu (2010c). $\lambda$ is the spatial autoregressive coefficient, $\gamma$ captures the time dynamic effect and $\delta$ captures the combined spatial-time effect.

tial coefficients ($\lambda$ and $\rho$) because they are sensitive to the choice of the spatial weights matrix.[7]

In the empirical matching literature, matching functions are mostly specified in a static way.[8] However, according to the Bayesian information criterion, the dynamic model fits the data better than the static model. This is in line with the positive temporal autocorrelation detected in the data. Thus, a dynamic approach is more appropriate for the modeling of matching functions.

Compared with the static model, the matching elasticities of the dynamic model are smaller. This can be explained by the strong time-dynamic effect in the data which is

[7] Hujer et al. (2009) also find in their study that the long-term effect of labor market policies is unaffected by changes in the spatial weights while the estimates of spatial coefficients differ with the choice of the spatial weights matrix.

[8] One exception is the contribution by Hujer et al. (2009) that specifies a spatial dynamic matching function for the analysis of the indirect and direct effects of active labor market policy at the regional level for Western Germany.

absorbed by the coefficients of the static model. Moreover, the coefficient of the spatial lag term ($\lambda$) is larger in the dynamic model. A reason for this might be that the dynamic model only contains a spatial lag term but not a spatial error term as the static model. The space-time effect is negative in some of the specifications. This means that an increase in the number of matches in neighboring local employment offices during the previous period results in lower matches during the present period. However, this result has to be taken with care. According to Ochsen (2009), the negative sign can arise from the perfect correlation of the space-time lagged variable with the time lagged and the spatially lagged variables.

Comparing both subperiods, the estimated elasticities of matches with respect to unemployment are larger during the period from 2005 until 2009 which holds for both the static and the dynamic model. Hence, the effect of additional unemployment on matching is stronger. However, the picture for the elasticities with respect to vacancies is different: They are larger during the time from 2000 until 2004 in case of the dynamic model while the opposite holds for the static model. Hence, we can conclude for the dynamic model that the relation between unemployed persons and vacant jobs has improved during the second period. The estimated spatial effects are similar for both subperiods in case of the static model while this is not true for the dynamic model. Only the pure time-dynamic effect is fairly similar in both subperiods.

## 7 Conclusion

In this paper, we estimate German matching functions taking into account spatial dependencies. We show that German matching data exhibit significant spatial autocorrelation. To avoid biased and inefficient estimates, we apply a spatial econometric modeling to the matching function. Our panel data set covers monthly information for 176 local employment offices in Germany for the period from 2000 until 2009. In order to capture the dynamics on labor markets, we use not only a static modeling but also a dynamic model specification. For the estimation, we follow the methodology proposed in Lee and Yu (2010b) and in Lee and Yu (2010c) for the static and the dynamic model, respectively. To incorporate the spatial information into the model, we construct different spatial weights matrices. As the amount of commuting reflects interregional relations on labor markets, we exploit commuting data for the construction of different spatial weights matrices. Our results suggest that neglecting spatial dependencies yields overestimated matching elasticities. Furthermore, they show that the dynamic model captures the structure in the data in a more appropriate way.

Regarding policy implications, our results suggest significant spatial spillovers. This means that regional policy activities have wider consequences. On the one hand, a local unemployment shock is not limited to one region but has also effects on neighboring regions. But on the other hand, regional activities aiming at a reduction of unemployment also have an impact on neighboring regions. Since we use numbers of commuters to measure the spatial impact in our model, neighboring regions are not limited to those that are a neighbors in the literally sense. Hence, the presence and the range of spatial spillovers has to be taken into account when regional policy measures are designed.

# Appendix

## Panel unit root tests

### Im, Pesaran and Shin (IPS) Test

Im et al. (2003) consider a sample of $N$ cross sections observed over $T$ time periods. They suppose that the stochastic process $y_{it}$ is generated by a first-order autoregressive process:

$$y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + \epsilon_{it}, \qquad i = 1,\ldots,N, t = 1,\ldots,T \tag{15}$$

where initial values $y_{i0}$ are given. To test the null hypothesis of unit roots, i.e. $\phi_i = 1$ for all $i$, Im et al. (2003) express equation (15) further:

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \epsilon_{it} \tag{16}$$

where $\alpha_i = (1-\phi_i)\mu_i$, $\beta_1 = -(1-\phi_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$. In this formulation, the null hypothesis that each series in the panel contains a unit root, and the alternative allowing for some (but not all) of the individual series to have a unit root, correspond to

$$H_0 : \beta_i = 0 \, \forall i \tag{17}$$

and

$$H_1 : \begin{cases} \beta_i < 0 & i = 1,\ldots,N_1 \\ \beta_i = 0 & i = N_1+1,\ldots,N. \end{cases} \tag{18}$$

This formulation of the alternative is more general than the homogeneous alternative, i.e. $\beta_i = \beta < 0$. Im et al. (2003) assume that under the alternative hypothesis the fraction of the individual processes that are stationary is nonzero, namely if $\lim_{N\to\infty}(N_1/N) = \delta$, $0 < \delta \leq 1$. This condition is necessary for the consistency of the test. Im et al. (2003) propose both unit root tests for heterogeneous panels with fixed $T$ and serially uncorrelated errors and unit root tests for heterogeneous panels with serially correlated errors. For the sake of brevity, we only consider the test for serially uncorrelated errors. The IPS $t$-bar statistic is defined as the average of the individual ADF statistics, i.e.

$$\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_{iT} \tag{19}$$

where $t_{iT}$ is the individual $t$-statistic for testing $H_0 : \beta_i = 0$ for all $i$ in equation (18). Im et al. (2003) show that for heterogeneous panels with serially uncorrelated errors the standardized $t$-bar statistic is distributed as standard normal as $N \to \infty$ for a fixed T, as long as $T > 5$ in the case of DF regressions with intercepts and $T > 6$ in the case of DF regressions with intercepts and linear time trends. Finally, in Monte Carlo experiments, Im et al. (2003) show that if a large enough lag order is selected for the underlying ADF regressions, then the small sample performance of the $t$-bar test is reasonably satisfactory and generally better than the Levin et al. (2002) (LLC) test.

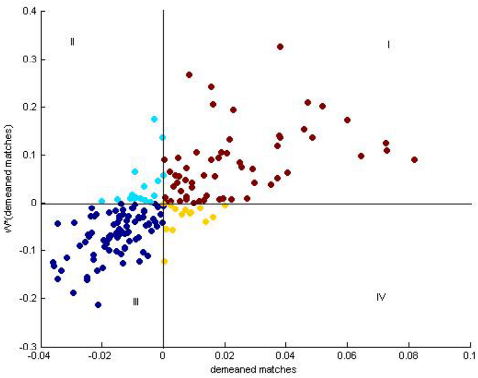
### Fisher-type Tests

Let $G_{iT_i}$ be a unit root test statistic for the $i$th group in a panel and assume that as the time series observations for the $i$th group $T_i \to \infty$, $G_{iT_i} \Rightarrow G_i$ where $G_i$ is a nondegenerate random variable. Let $p_i$ be the asymptotic $p$-value of a unit root test for cross-section $i$, i.e. $p_i = F(G_{iT_i})$, where $F(\bullet)$ is the distribution function of the random variable $G_i$ (see Baltagi (2005)). Maddala and Wu (1999) propose a Fisher-type test

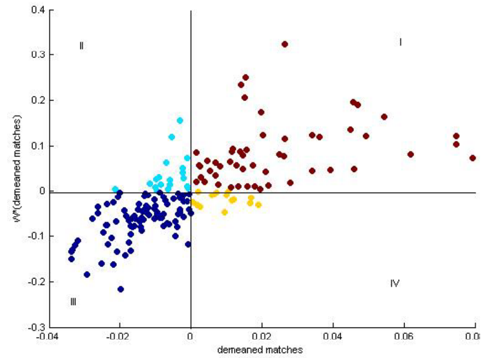
$$P = -2\sum_{i=1}^{N} lnp_i \tag{20}$$

which combines the $p$-values from unit root tests for each cross-section $i$ to test for unit roots in the panel data set. The statistic $P$ has a $\chi^2$ distribution with two degrees of freedom as $T_i \to \infty$ for finite N. Maddala and Wu (1999) argue that the advantage of this test is, firstly, that no balanced panel is required as it is the case for the IPS test. Secondly, it is possible to use different lag lengths in the individual ADF regressions and, thirdly, it can be carried out for any panel unit root test. However, the $p$-values have to be derived by Monte Carlo simulation which is a disadvantage of this test. Moreover, Maddala and Wu (1999) find that the Fisher-type test with bootstrap-based critical values performs the best and is the preferred choice for testing the null hypothesis of nonstationarity as well as in testing for cointegration in panels.

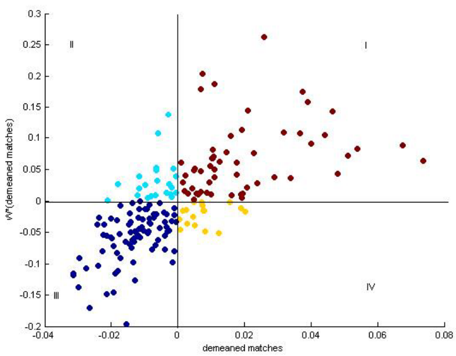
## Moran scatter plots (2000-2008)



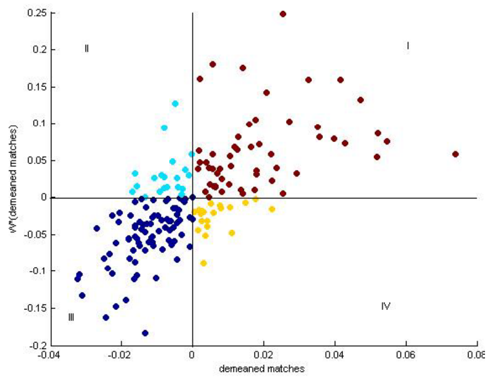

**2000**




**2001**



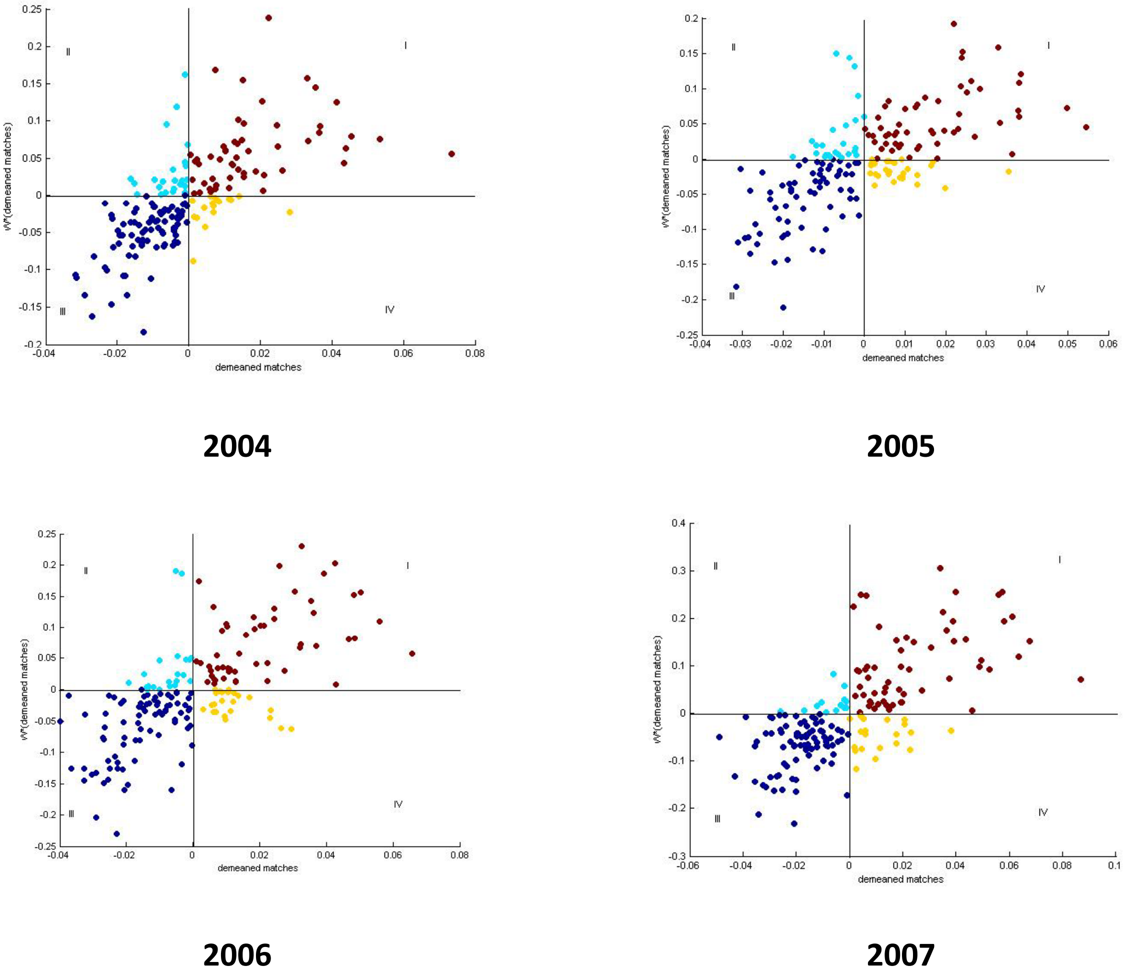

**2002**




**2003**

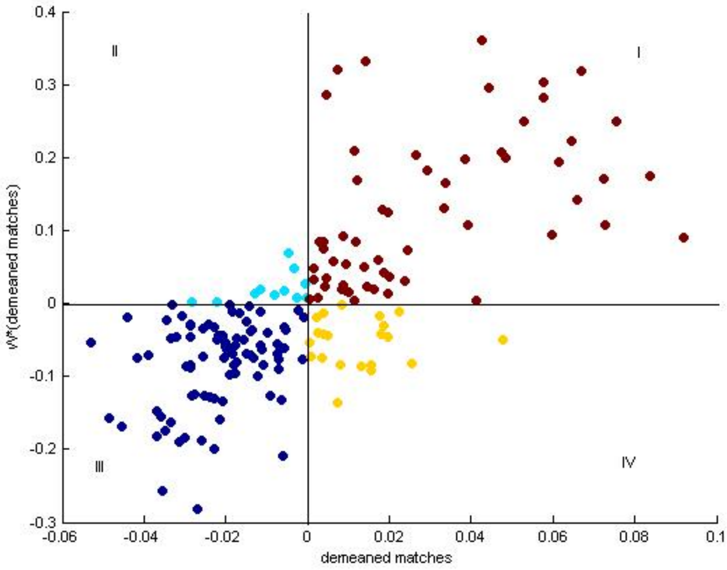
Note: This figure plots spatially lagged matches against demeaned matches. The colors identify the location of the points in the four quadrants. They are used to show where the points are located on the German map (see Figure 9).

Figure 10: Moran scatter plots using yearly averages of standardized matches for the period from 2000 until 2003




Note: This figure plots spatially lagged matches against demeaned matches. The colors identify the location of the points in the four quadrants. They are used to show where the points are located on the German map (see Figure 9).

Figure 11: Moran scatter plots using yearly averages of standardized matches for the period from 2004 until 2007

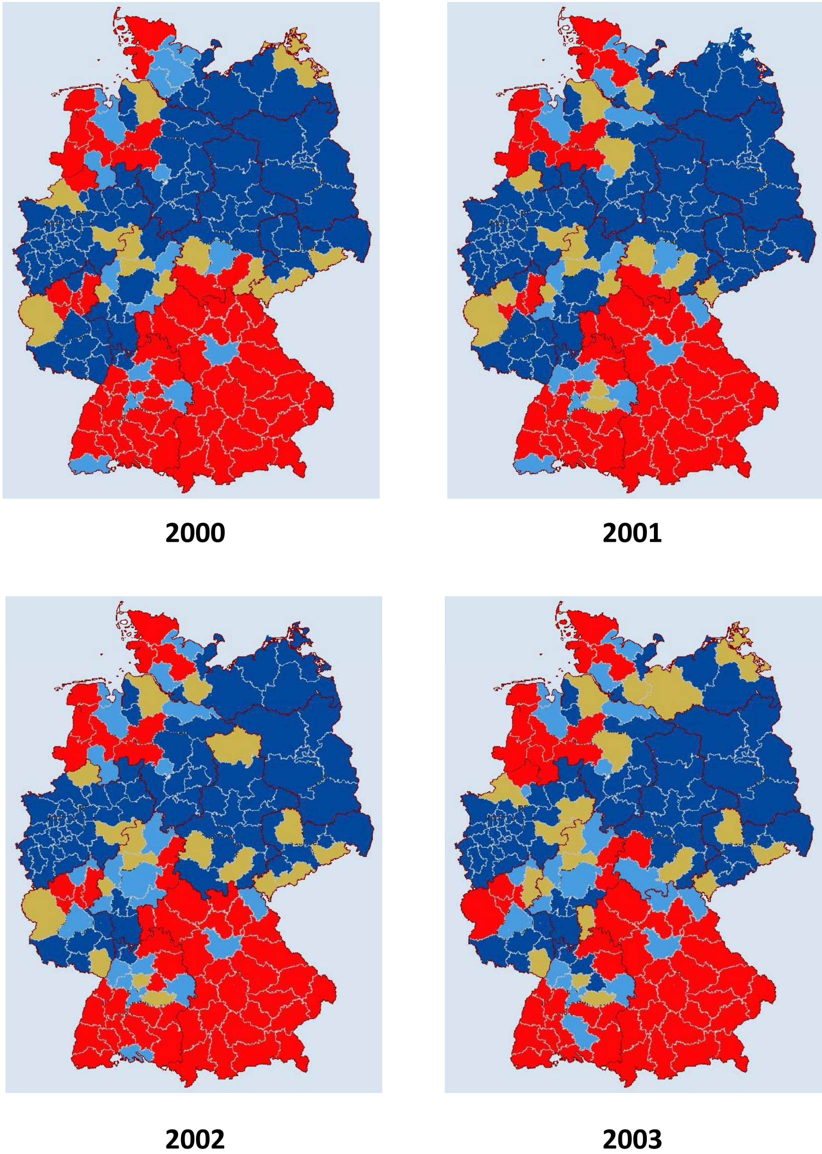



**2008**

Note: This figure plots spatially lagged matches against demeaned matches. The colors identify the location of the points in the four quadrants. They are used to show where the points are located on the German map (see Figure 9).

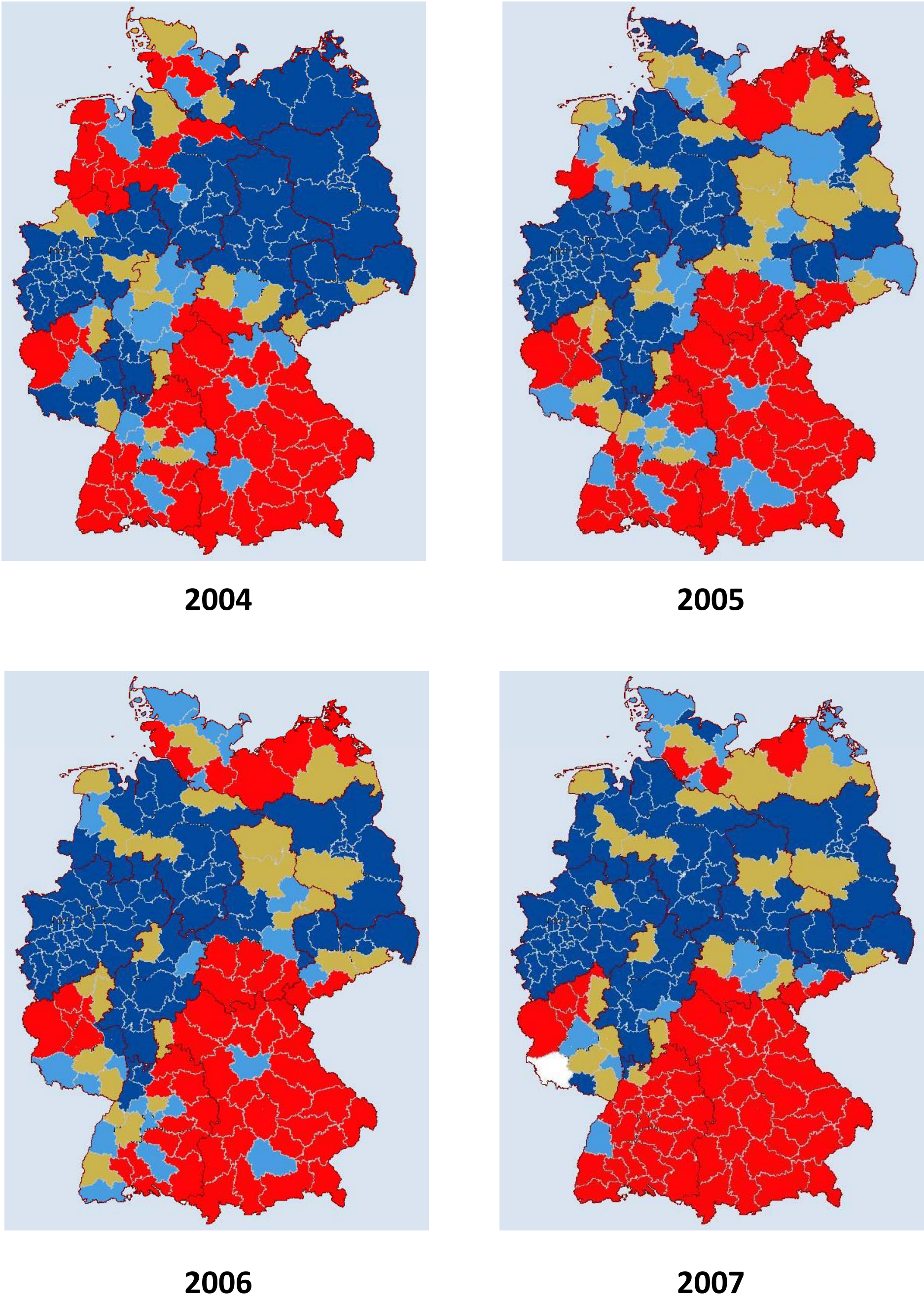
Figure 12: Moran scatter plots using yearly averages of standardized matches in 2008

**Moran maps (2000-2008)**



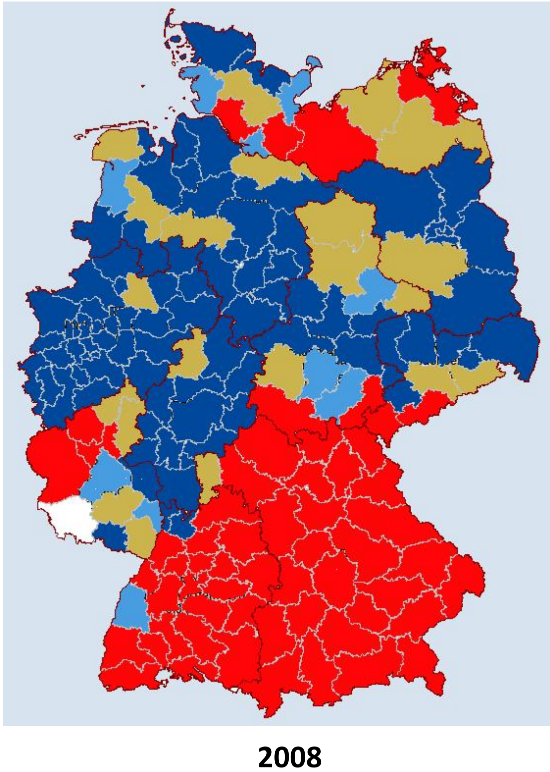

Note: The colors indicate the four quadrants of the Moran scatter plot as shown in Figure 10.

Figure 13: German maps indicating the position of points in Moran scatter plot for yearly averages of standardized matches for period from 2000 until 2003




Note: The colors indicate the four quadrants of the Moran scatter plot as shown in Figure 10.

Figure 14: German maps indicating the position of points in Moran scatter plot for yearly averages of standardized matches for period from 2004 until 2007




Note: The colors indicate the four quadrants of the Moran scatter plot as shown in Figure 10.

Figure 15: German map indicating the position in Moran scatter plot for yearly averages of standardized matches in 2008

## References

ANDERSON, P. M. AND S. M. BURGESS (2000): "Empirical Matching Functions: Estimation and Interpretation Using State-Level Data," *Review of Economics and Statistics*, 82, 93–102.

ANSELIN, L. (1996): "The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association," in *Spatial analytical perspectives on GIS*, ed. by M. Fischer, London: Taylor & Francis, 111–125.

ANSELIN, L. AND A. K. BERA (1998): "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics," in *Handbook of applied economic statistics*, ed. by A. Ullah and D. E. A. Giles, New York, NY: Dekker, 237–289.

BA (2008): "Qualitätsbericht der gemeldeten Stellen," *http://statistik. arbeitsagentur.de/Navigation/Statistik/Grundlagen/ Qualitaetsberichte/Qualitaetsberichte-Nav.html*, 10/21/2010.

BALTAGI, B. H. (2005): *Econometric analysis of panel data*, Chichester: Wiley, 3. ed.

BALTAGI, B. H., G. BRESSON, AND A. PIROTTE (2007): "Panel Unit Root Tests and Spatial Dependence," *Journal of Applied Econometrics*, 22, 339–360.

BLANCHARD, O. J. AND P. DIAMOND (1989): "The Beveridge Curve," *Brookings Papers on Economic Activity*, 1989, 1–76.

BURDA, M. C. (1993): "Labour Markets and Structural Change in Eastern Europe," *Economic Policy*, 8, 102–137.

——— (1994): "Modelling Exits from Unemployment in Eastern Germany: A Matching Function Approach," in *Arbeitsmarktdynamik und Unternehmensentwicklung in Osteuropa: Erfahrungen und Perspektiven des Transformationsprozesses: Beiträge eines Workshops des Zentrums für Europäische Wirtschaftsforschung am 4. und 5. März 1993 in Mannheim*, ed. by H. König and V. Steiner, Baden-Baden: Nomos-Verl.-Ges.

BURGESS, S. AND S. PROFIT (2001): "Externalities in the matching of workers and firms in Britain," *Labour Economics*, 8, 313–333.

CHOI, I. (2001): "Unit root tests for panel data," *Journal of International Money and Finance*, 20, 249–272.

CHRISTENSEN, B. (2001): “Mismatch-Arbeitslosigkeit unter Geringqualifizierten,” *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 34, 506–514.

CLIFF, A. D. AND J. K. ORD (1972): “Testing for spatial autocorrelation among regression residuals,” *Geographical Analysis*, 4, 267–284.

——— (1981): *Spatial processes: Models and applications*, London: Pion.

COLES, M. G. AND E. SMITH (1996): “Cross-Section Estimation of the Matching Function: Evidence from England and Wales,” *Economica*, 63, 589–597.

DMITRIJEVA, J. (2008): “Matching and labour market efficiency across space and through EU accesion: evidence from Latvia, Estonia and Slovenia,” *EPEE Working Paper, Université d’Evry*.

FAHR, R. AND U. SUNDE (2001): “Disaggregate Matching Functions,” *IZA Discussion Paper No. 335*.

——— (2006a): “Regional dependencies in job creation: an efficiency analysis for Western Germany,” *Applied Economics*, 38, 1193–1206.

——— (2006b): “Spatial mobility and competition for jobs: Some theory and evidence for Western Germany,” *Regional Science and Urban Economics*, 36, 803–825.

——— (2009): “Did the Hartz Reforms Speed-Up the Matching Process? A Macro-Evaluation Using Empirical Matching Functions,” *German Economic Review*, 10, 284–316.

FINGLETON, B. (1999): “Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration,” *Journal of Regional Science*, 39, 1–19.

FRANZESE, R. J. AND J. C. HAYS (2007): “Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data,” *Political Analysis*, 15, 140–164.

HUJER, R., P. J. M. RODRIGUES, AND K. WOLF (2009): “Estimating the macroeconomic effects of active labour market policies using spatial econometric methods,” *International Journal of Manpower*, 30, 648–671.

HYNNINEN, S.-M. (2005): “Matching across Space: Evidence from Finland,” *LABOUR*, 19, 749–765.

IM, K. S., M. H. PESARAN, AND Y. SHIN (2003): "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 115, 53–74.

KLOTZ, S. (2004): *Cross Sectional Dependence in Spatial Econometric Models: With an Application to German Start-up Activity Data*, Münster, Westf: LIT, 1st ed.

KOPPEL, O. (2008): "Ingenieurarbeitsmarkt in Deutschland - gesamtwirtschaftliches Stellenangebot und regionale Fachkräftelücken," *IW-Trends – Vierteljahresschrift zur empirischen Wirtschaftsforschung aus dem Institut der deutschen Wirtschaft Köln*, 35, 1–16.

LEE, L.-F. AND J. YU (2010a): "A Spatial Dynamic Panel Data Model with Both Time and Individual Fixed Effects," *Econometric Theory*, 26, 564–597.

——— (2010b): "Estimation of spatial autoregressive panel data models with fixed effects," *Journal of Econometrics*, 154, 165–185.

——— (2010c): "Some recent developments in spatial panel data models," *Regional Science and Urban Economics*, 40, 255–271.

LEVIN, A., C.-F. LIN, AND C.-S. J. CHU (2002): "Unit root tests in panel data: asymptotic and finite-sample properties," *Journal of Econometrics*, 108, 1–24.

MADDALA, G. S. AND S. WU (1999): "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics*, 61, 631–652.

MÖLLER, J. AND A. ALDASHEV (2007): "Wage Inequality, Reservation Wages and Labor Market Participation: Testing the Implications of a Search-Theoretical Model with Regional Data," *International Regional Science Review*, 30, 120–151.

MORAN, P. A. P. (1950): "Notes on continuous phenomena," *Biometrika*, 37, 17–23.

OCHSEN, C. (2009): "Regional Labor Markets and Aging in Germany," *Thünen-Series of Applied Economic Theory, No. 102, Universität Rostock*.

PESARAN, M. H. (2004): "General Diagnostic Tests for Cross Section Dependence in Panels," *CESifo Working Paper No. 1229*.

PETRONGOLO, B. AND C. PISSARIDES (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

STEVENS, M. (2007): "New Microfoundations for the Aggregate Matching Function," *International Economic Review*, 48, 847–868.

VAN OURS, J. C. (1991): "The Efficiency of the Dutch Labor Market in Matching Unemployment and Vacancies," *De Economist*, 139, 358–378.

# SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Volatility Investing with Variance Swaps" by Wolfgang Karl Härdle and Elena Silyakova, January 2010.
002 "Partial Linear Quantile Regression and Bootstrap Confidence Bands" by Wolfgang Karl Härdle, Ya'acov Ritov and Song Song, January 2010.
003 "Uniform confidence bands for pricing kernels" by Wolfgang Karl Härdle, Yarema Okhrin and Weining Wang, January 2010.
004 "Bayesian Inference in a Stochastic Volatility Nelson-Siegel Model" by Nikolaus Hautsch and Fuyu Yang, January 2010.
005 "The Impact of Macroeconomic News on Quote Adjustments, Noise, and Informational Volatility" by Nikolaus Hautsch, Dieter Hess and David Veredas, January 2010.
006 "Bayesian Estimation and Model Selection in the Generalised Stochastic Unit Root Model" by Fuyu Yang and Roberto Leon-Gonzalez, January 2010.
007 "Two-sided Certification: The market for Rating Agencies" by Erik R. Fasten and Dirk Hofmann, January 2010.
008 "Characterising Equilibrium Selection in Global Games with Strategic Complementarities" by Christian Basteck, Tijmen R. Daniels and Frank Heinemann, January 2010.
009 "Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory" by Julia Schaumburg, February 2010.
010 "On Securitization, Market Completion and Equilibrium Risk Transfer" by Ulrich Horst, Traian A. Pirvu and Gonçalo Dos Reis, February 2010.
011 "Illiquidity and Derivative Valuation" by Ulrich Horst and Felix Naujokat, February 2010.
012 "Dynamic Systems of Social Interactions" by Ulrich Horst, February 2010.
013 "The dynamics of hourly electricity prices" by Wolfgang Karl Härdle and Stefan Trück, February 2010.
014 "Crisis? What Crisis? Currency vs. Banking in the Financial Crisis of 1931" by Albrecht Ritschl and Samad Sarferaz, February 2010.
015 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappusl and Markus Reiß, February 2010.
016 "Honey, I'll Be Working Late Tonight. The Effect of Individual Work Routines on Leisure Time Synchronization of Couples" by Juliane Scheffel, February 2010.
017 "The Impact of ICT Investments on the Relative Demand for High-Medium-, and Low-Skilled Workers: Industry versus Country Analysis" by Dorothee Schneider, February 2010.
018 "Time varying Hierarchical Archimedean Copulae" by Wolfgang Karl Härdle, Ostap Okhrin and Yarema Okhrin, February 2010.
019 "Monetary Transmission Right from the Start: The (Dis)Connection Between the Money Market and the ECB's Main Refinancing Rates" by Puriya Abbassi and Dieter Nautz, March 2010.
020 "Aggregate Hazard Function in Price-Setting: A Bayesian Analysis Using Macro Data" by Fang Yao, March 2010.
021 "Nonparametric Estimation of Risk-Neutral Densities" by Maria Grith, Wolfgang Karl Härdle and Melanie Schienle, March 2010.

# SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

022 "Fitting high-dimensional Copulae to Data" by Ostap Okhrin, April 2010.
023 "The (In)stability of Money Demand in the Euro Area: Lessons from a Cross-Country Analysis" by Dieter Nautz and Ulrike Rondorf, April 2010.
024 "The optimal industry structure in a vertically related market" by Raffaele Fiocco, April 2010.
025 "Herding of Institutional Traders" by Stephanie Kremer, April 2010.
026 "Non-Gaussian Component Analysis: New Ideas, New Proofs, New Applications" by Vladimir Panov, May 2010.
027 "Liquidity and Capital Requirements and the Probability of Bank Failure" by Philipp Johann König, May 2010.
028 "Social Relationships and Trust" by Christine Binzel and Dietmar Fehr, May 2010.
029 "Adaptive Interest Rate Modelling" by Mengmeng Guo and Wolfgang Karl Härdle, May 2010.
030 "Can the New Keynesian Phillips Curve Explain Inflation Gap Persistence?" by Fang Yao, June 2010.
031 "Modeling Asset Prices" by James E. Gentle and Wolfgang Karl Härdle, June 2010.
032 "Learning Machines Supporting Bankruptcy Prediction" by Wolfgang Karl Härdle, Rouslan Moro and Linda Hoffmann, June 2010.
033 "Sensitivity of risk measures with respect to the normal approximation of total claim distributions" by Volker Krätschmer and Henryk Zähle, June 2010.
034 "Sociodemographic, Economic, and Psychological Drivers of the Demand for Life Insurance: Evidence from the German Retirement Income Act" by Carolin Hecht and Katja Hanewald, July 2010.
035 "Efficiency and Equilibria in Games of Optimal Derivative Design" by Ulrich Horst and Santiago Moreno-Bromberg, July 2010.
036 "Why Do Financial Market Experts Misperceive Future Monetary Policy Decisions?" by Sandra Schmidt and Dieter Nautz, July 2010.
037 "Dynamical systems forced by shot noise as a new paradigm in the interest rate modeling" by Alexander L. Baranovski, July 2010.
038 "Pre-Averaging Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence" by Nikolaus Hautsch and Mark Podolskij, July 2010.
039 "High Dimensional Nonstationary Time Series Modelling with Generalized Dynamic Semiparametric Factor Model" by Song Song, Wolfgang K. Härdle, and Ya'acov Ritov, July 2010.
040 "Stochastic Mortality, Subjective Survival Expectations, and Individual Saving Behavior" by Thomas Post and Katja Hanewald, July 2010.
041 "Prognose mit nichtparametrischen Verfahren" by Wolfgang Karl Härdle, Rainer Schulz, and Weining Wang, August 2010.
042 "Payroll Taxes, Social Insurance and Business Cycles" by Michael C. Burda and Mark Weder, August 2010.
043 "Meteorological forecasts and the pricing of weather derivatives" by Matthias Ritter, Oliver Mußhoff, and Martin Odening, September 2010.
044 "The High Sensitivity of Employment to Agency Costs: The Relevance of Wage Rigidity" by Atanas Hristov, September 2010.
045 "Parametric estimation of risk neutral density functions" by Maria Grith and Volker Krätschmer, September 2010.

## SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

046 "Mandatory IFRS adoption and accounting comparability" by Stefano Cascino and Joachim Gassen, October 2010.
047 "FX Smile in the Heston Model" by Agnieszka Janek, Tino Kluge, Rafał Weron, and Uwe Wystup, October 2010.
048 "Building Loss Models" by Krzysztof Burnecki, Joanna Janczura, and Rafał Weron, October 2010.
049 "Models for Heavy-tailed Asset Returns" by Szymon Borak, Adam Misiorek, and Rafał Weron, October 2010.
050 "Estimation of the signal subspace without estimation of the inverse covariance matrix" by Vladimir Panov, October 2010.
051 "Executive Compensation Regulation and the Dynamics of the Pay-Performance Sensitivity" by Ralf Sabiwalsky, October 2010.
052 "Central limit theorems for law-invariant coherent risk measures" by Denis Belomestny and Volker Krätschmer, October 2010.
053 "Systemic Weather Risk and Crop Insurance: The Case of China" by Wei Xu, Ostap Okhrin, Martin Odening, and Ji Cao, October 2010.
054 "Spatial Dependencies in German Matching Functions" by Franziska Schulze, November 2010.