Spatial Aspects of Job Creation: Evidence from Western Germany

René Fahr[*]
IZA, University of Bonn

Uwe Sunde
IZA, University of Bonn

January 2002

Approximate word count: 7900

Abstract

This paper provides detailed information about spatial interactions in the job creation process in West German regional labor markets. We investigate spatial (auto-) correlations in the matching process of vacancies and unemployed, examine regional hiring patterns, and identify clusters of regions with intense inter-regional matching. An extensive specification analysis illustrates the extent of regional dependencies. We investigate the impact of German re-unification on regional patterns of job creation, and compare regional matching efficiencies using a stochastic frontier approach.

JEL-classification: J61, J64, J21, R12

Keywords: Internal Migration, Regional Unemployment, Stochastic Frontiers

[*]Corresponding author. Address: IZA, Schaumburg-Lippe Str. 7, 53113 Bonn, Tel. +49-0228-3894-533, email: fahr@iza.org.

1 Introduction

In this paper, we investigate interactions in the job creation process in West German labor market regions. We deal with questions such as: Do geographic environment and spatial issues matter for the matching process? In particular, does the general job situation of neighboring regions matter for the job creation in a given region, even if regions are broadly defined as travel-to-work areas? Are there differences in the spatial structures of all matches, matches from unemployment and job-to-job transitions? Did German re-unification have a measurable impact on internal migration behavior in West Germany? We address these issues by providing evidence for spatial autocorrelation in labor market variables. We examine spatially augmented empirical matching functions, and we study the regional structure of labor market flows. Moreover, we isolate regions of particularly intense interregional dependencies of labor markets, so-called hot spots and clusters, and estimate the matching efficiencies of different regional labor markets.

Our findings extend the previous literature on spatial matching, including Gorter and Van Ours (1994), Burda and Profit (1996) and Burgess and Profit (2001) in several aspects. We present the first thorough analysis of this kind for Germany, and extend the spatial frameworks used before by an extensive specification analysis. Moreover, unlike any of the previous contributions, our data allow to decompose labor market flows simultaneously along the spatial dimension as well as by previous employment status of newly hired.

However, information about the structure and productivity of the matching process is not very informative, if, at the same time, the inefficiencies involved in the matching process are high, i.e. an increase in the stocks leads to the creation of fewer jobs than technically feasible. Similar to recent contributions by Ibourk, Maillard, Perelman, and Sneessens (2001) and Ilmakunnas and Pesola (2002), we estimate stochastic matching frontiers using regional data. This enables us to investigate the extent and determinants of the inefficiencies involved in the job creation process. For a policy maker deciding on labor market policies in certain regions this provides information about the appropriateness of different policy alternatives and can be useful for cost-benefit evaluations.

The paper proceeds as follows. Section 2 describes the data used throughout the analysis. In section 3 we investigate spatial dependencies in the labor market conditions and the job creation process across regional labor markets in West Germany. After providing evidence that conventional matching studies neglecting the spatial dimension of job creation are misspecified, section 4 presents results obtained using spatially augmented empirical matching functions. In section 5 we complement our analysis of the regional matching function with a stochastic production frontier analysis of regional matching efficiencies. Section 6 concludes.

2 Data issues

The data used for the analysis below are yearly data on unemployment, vacancies and hirings for the years 1980 until 1997 for 117 regions in Western Germany. The data on the stock of unemployed and vacanies are from official labor statistics and available for so called Employment Office Districts. The hirings are measured on the individual level and stem from an anonymized representative 1% sample of German social security records provided by the German Federal Institute for Employment Research (IAB). The database is supplemented by data on unemployment benefits recipients and by establishment information (see Bender, Haas, and Klose (2000) for details). The data allow to identify the precise date of a new hire, as well as the employment history and the geographical location (as well as changes in the location) of the respective individual. In particular, a change in the employment status of an individual indicates a transition from unemployment to employment or vice versa. No change in the employment status, but a change in the firm identifier indicates a job-to-job transition. Regions are identified by locations of employers, thus changes in (plant of) employer identifiers can imply changes in region identifier, and thereby regional mobility. We aggregate new matches into year-region cells, where regions correspond to labor market districts as defined by the Federal Office of Building and Regional Planning, and are designed so as to capture travel-to-work areas as good as possible. We merge the hirings data and the stock data to the respective coarser region definition, which is in most of the cases the one from official labor statistics defining regions as Employment Office Districts. A list with the labor market regions used in the empirical analysis, as well as a map indicating their location, are contained in the Appendix.

3 Spatial Dependencies in the Labor Market

This section attempts to shed some light on spatial dependencies in the labor market. In particular, we investigate whether new jobs, i. e. newly created employer-employee matches are spatially autocorrelated, and whether the labor market conditions (levels of new matches, vacancies and unemployment) in neighboring regions matter for the job creation process within a region. The section proceeds as follows. First, we employ tests for global spatial autocorrelation on the data for new matches. We then investigate this issue in some more detail and ask whether there is evidence for local spatial autocorrelation and clusters of regions affecting each other with respect to labor market outcomes.

The absence of evidence for spatial autocorrelation would indicate that considering geographic aspects is not crucial for modeling the labor market. However, we find indications for spatial effects. Therefore, we next estimate conventional U/V-matching functions and test for misspecification. In particular, we test the conventional model against alternatives like spatial autoregression in the dependent variable and spatially autoregressive error terms. Later, we also provide results from regressing spatial specifications of the matching function, including specifications instrumenting the (spatially) lagged dependent variable using (spatially) lagged explanatory variables.

3.1 Global and Local Spatial Autocorrelation

In order to reveal the spatial pattern of search and matching behavior on the labor market, we first test whether the variables of primary interest in the context of empirical labor market matching exhibit spatial autocorrelation. Spatial autocorrelation means that the spatial distribution of new successful matches during a certain defined period of time (in our case a year) exhibits a systematic pattern. In other words, if new matches are positively spatially autocorrelated, a high job creation activity in a certain region is associated with high job creation in nearby regions. Since the data we use consist of cells of 117 West-German labor market regions, we define contiguity between two regions as the regions sharing a common border. The corresponding spatial weights matrix W is therefore a symmetric 117×117 matrix with entries 0 and 1, where 1 indicates contiguity.[1] In order to test the null hypothesis of no spatial autocorrelation, we employ Moran's I-test for global spatial autocorrelation, see Anselin and Bera (1999) for details. Where appropriate, we also report results for alternative measures of global autocorrelation like Geary's c and Getis and Ord's G.

Since also the structure of the explanatory variables matters for the empirical matching context, we apply the testing procedures on new matches, the dependent variable, and on the stocks of unemployed and vacancies. Unfortunately, these three tests only utilize the cross-sectional dimension of the data. Therefore, we replicate the tests for each time period within the observation window 1980-1997. For reasons of space we only report the general findings. Detailed results are available from the authors upon request. The results can be summarized as follows. There is strong evidence for positive spatial autocorrelation of the explanatory variables, unemploy-

[1]The entries on the main diagonal of W are zeros, since a region cannot be contiguous to itself.

ment and vacancies, as measured by Moran's I and Geary's c. However, the null cannot be rejected for the dependent variable, hires. The analysis of Getis and Ord's G, leads to somewhat different conclusions.

According to this measure, new matches are spatially autocorrelated and characterized by strong high-valued global clustering. On the other hand, evidence for clustering in the explanatory variables, particularly regional unemployment, is weak. These results are interesting in the light of previous results in the literature. Burgess and Profit (2001) use data for the U.K. and test their two concepts of dependent variables, outflows from unemployment and filled vacancies, for spatial autocorrelation using Moran's I-test. They find strong evidence for spatial spillovers in matching, while our findings suggest that there is only very weak if any spatial autocorrelation in the dependent variables.

Next, we test for local spatial autocorrelation in the data. It turns out that spatial patterns exhibit substantial heterogeneity across regions. Moreover, this heterogeneity is fairly stable over time.

According to averages over the years 1980 to 1997 of Moran's I test statistics the Ruhr area around the cities Düsseldorf, Essen and Gelsenkirchen represents a huge common labor market (cluster), characterized by strong positive spatial autocorrelation.[2] On the other hand, agglomeration areas surrounded by less densely populated, rural regions, like Hamburg, Frankfurt, Stuttgart and Munich constitute hot spots characterized by strong negative spatial autocorrelation that attract many workers from surround-

[2] Other regions with high test scores for positive autocorrelation are Lübeck, Leer, Cologne, Mönchengladbach, Münster, Korbach.

ing areas during booms and set free many workers to surrounding areas during recessions.[3]

3.2 Spatial Misspecification of Conventional Matching Functions

In order to find out more about spatial dependencies in the matching process, we regress conventional matching functions of the Cobb-Douglas specification

$$\ln m_{it} = A + \alpha \ln U_{it} + \beta \ln V_{it} + \varepsilon_{it} \, , \tag{1}$$

where m_{it} denotes the new matches created in region i within a period of time, i. e. between t and $t+1$, U_{it} is the number of unemployed job seekers in region i at the observation period t, and V_{it} is the number of vacancies in i at t, while α and β are parameters. ε denotes a vector of normally distributed, homoskedastic and uncorrelated errors. In the presence of spatial dependencies among observations, this model might be misspecified. Therefore, we test this model against two alternative specifications taking spatial dependencies explicitly into account. The first of these is the *spatial error model*. Essentially, a model identical to (1) is estimated, but imposing a different error structure:

$$\varepsilon_{it} = \lambda W \varepsilon_{it} + \mu_{it} \, , \tag{2}$$

with W representing the spatial weights matrix mentioned above, λ denoting the spatial autoregressive parameter, and μ denoting a vector of homoskedastic and uncorrelated errors. The second spatial model we consider

[3] Other hot spots are Heide, Bremen, Hannover, Lüneburg, Düren, Nuremberg.

is the following mixed regressive spatial autoregressive model, the *spatial lag model*:

$$\ln m_{it} = A + \alpha \ln U_{it} + \beta \ln V_{it} + \rho W m_{it} + \varepsilon_{it} \ , \tag{3}$$

where ρ denotes the spatial autoregression parameter, and Wm_{it} denotes the spatially lagged dependent variable (the weighted sum of contemporary matches in neighboring regions). Since the tests make no use of the time dimension but only of the cross-sectional (regional) variation of the data, we estimate and test the three models for each year between 1980 and 1997.

The results of these tests reveal that the conventional specification of the matching function (1) can be rejected in favor of one or both alternative specifications for every year. For the years 1980, 1981, and 1989-1993, the conventional model is rejected in favor of the spatial lag model (3), that is the null that ρ equals zero cannot be rejected for these years, while the hypothesis that λ equals zero can be rejected at conventional levels. For the years 1984-1986, the opposite is true. During all the remaining years, both null hypotheses can be rejected, suggesting that both types of spatial dependencies play a role in the matching process. This casts doubts on the validity of results obtained from matching studies using regional data that neglect spatial dependencies in the variables, like Gorter and Van Ours (1994).

As a next step, we search for the most preferred specification of the matching process by estimating spatial error and spatial lag models by maximum likelihood separately for each year. Using the estimation results we test whether the hypotheses that $\lambda = 0$ in case of the spatial error model and $\rho = 0$ in case of the spatial lag model can be rejected. Generally, with the exception of four (out of 18) years, we cannot reject both hypotheses at the 5 percent level. However, $\rho = 0$ can be rejected at more generous significance levels (around, say, 10 to 12 percent) in most years, in favor of the spatial lag model. On the other hand, $\lambda = 0$ can be rejected only in three years, 1995-1997, and in the latter two significantly (at the 5 percent level) in favor of the spatial error model. The conclusion we draw from this is that there is evidence that spatial determinants play some role and therefore have to be contained in the correct specification of the matching function. The results point rather towards a spatial lag specification rather than a spatial error specification.

4 Spatial Structure of Job Creation

This section provides a detailed analysis of the composition of employment inflows with respect to the regional origin and destination of hirees, as well as their previous employment status.

The first part is devoted to checking the robustness of the conventional matching function specification as presented in the previous section with respect to the choice of the dependent variable. We then look closer into the migration behavior of workers by investigating the spatial decomposition of matches and its dependence on the spatial structure of explanatory variables. As a further issue, we examine whether the German re-unification, which occurs after about half of the observation periods covered by our data, had an impact on regional migration behavior and the spatial composition of new matches.

4.1 Spatially Augmented Matching Functions

Given the evidence for the importance of spatial issues presented in the previous section, the first question one has in mind is whether the results obtained by conventional matching functions neglecting the spatial dimension can still come up with unbiased estimates. To answer this question, we estimate matching functions of specification (1) for different concepts of flows. In particular, we compare the results obtained by taking all flows m as dependent variable with estimations for taking only individuals stemming from within the region (m_h), new matches of individuals who were previously employed in neighboring regions (m_n), or in other non-neighboring regions (m_f). Alternatively, we can decompose flows by the job status of the respective new employed: individuals who were unemployed before successfully matching (m_u), and previously employed job switchers (m_e).[4] Moreover, we have results for the same concepts, further decomposed as interactions, that is flows from unemployment decomposed by regional origin (m_{uh}, m_{un}, m_{uf}) and formerly employed job switchers decomposed by where they come from (m_{eh}, m_{en}, m_{ef}). Table 1 contains the sample averages of these different concepts of matches over all years and regions in order to give some information about the quantitative relevance of the different measures.

As a consequence of the results of the previous section, we estimate conventional matching functions augmented by spatially autoregressive compo-

[4]For workers who were unemployed immediately before being hired, we have information about the region of their last job.

		Shares of total matchings $\boldsymbol{m}$			
		from same region	from neighbor region	from other regions	from all regions
		m_h	m_n	m_f	m
Hirings:					
from employment:	m_e	0.218	0.058	0.043	0.319
from unemployment:	m_u	0.200	0.028	0.032	0.260
from employment and unemployment	m	0.418	0.086	0.496	(0.579)* 1.000

Note: All data are aggregated over all 117 regions and averages over the period 1980-1997. Table entries are shares of the respective group characterized by regional status prior to current match and employment status prior to current match, with respect to total shares (that is they add up to 100 % horizontally).

* The data cannot identify regional origin of new matches from out of the labor force. Therefore, only 57.9 % of the new matches can be decomposed regionally. New hires from out of the labor force, making up for 42.1 % of all hires, are contained in all hires $\boldsymbol{m}$, but we refrain from analyzing them separately. Hires with missing region identifier are coded as "from other regions" $\boldsymbol{m_f}$, hires with missing employment status identifier are coded as from out of the labor force.

Table 1: Composition of Hires with Respect to Regional Origin and Employment Status

nents, as suggested by results on the spatial lag model specification. These estimations are conducted first separately for each year, utilizing only the cross-sectional variation of the data. Then, we also estimate matching functions using the entire panel structure of the data.

For brevity, we only report the main results for the pooled sample in Table 2.[5] As is standard in empirical studies of the matching function (see e.g. Petrongolo and Pissarides, 2001), we consistently find for specifications by year as well for as pooled specifications that the stocks of unemployed and vacancies exhibit highly significant positive effects on the number of matches with coefficients of between 0.35 and 0.55 for the majority of years under study.

[5]Detailed estimation results are available from the authors upon request.

	Dependent Variables:					
$y = \ln($	m_{all}	m_{all}	m_u	m_u	m_e	m_e **)**
ln(*U*)	0.451	0.553	0.572	0.512	0.385	0.594
	(0.008)	(0.023)	(0.013)	(0.026)	(0.010)	(0.026)
ln(*V*)	0.467	0.387	0.138	0.238	0.660	0.471
	(0.009)	(0.022)	(0.013)	(0.025)	(0.011)	(0.025)
ln(*W*y*)	-0.082		0.056		0.049	
	(0.016)		(0.023)		(0.017)	
ln(*WU*)		-0.200		-0.048		-0.293
		(0.026)		(0.030)		(0.029)
ln(*WV*)		0.014		-0.267		0.215
		(0.024)		(0.028)		(0.028)
Linear time trend	-0.039	-0.032	-0.026	-0.012	-0.037	-0.035
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Constant	6.805	6.736	4.101	5.584	3.403	4.271
	(0.197)	(0.132)	(0.235)	(0.175)	(0.197)	(0.163)
R^2	0.816	0.827	0.617	0.654	0.805	0.817
Observations	2106	2106	2106	2106	2106	2106

Note: Robust standard errors are in parentheses. Data contain observations for 117 regions and 18 years (1980-1997). Legend: *U* unemployment level, *V* vacancy level, *WU* unemployment levels in neighboring regions, *WV* vacancy levels in neighboring regions; *y* denotes dependent variable, which is the log of the respective concept of new hires *m: all* all hirings, *u* hirings of formerly unemployed, *e* hirings of formerly employed. .

Table 2: Empirical Matching Functions with Spatial Dependence

More novel is that consistently for all yearly and pooled specifications we find a significant negative coefficient for the spatially lagged dependent variable if all hirings are used as dependent variable. In order to account for potential simultaneity bias, and for robustness, we instrument the spatially lagged dependent variable with spatially lagged observations of the explanatory variables. While the results for the local explanatory variables are virtually unaffected by this, the coefficient for lagged unemployment turns out to be significantly negative throughout all specifications, while the sign of the coefficient for lagged vacancies depends on the concept of the dependent variable used: the effect is significantly positive for all hires and hires from employment, significantly negative for unemployment outflows into employment. Burda and Profit (1996) also use regions sharing a common border as definition for spatial contiguity when estimating spatially augmented matching functions. They obtain, depending on the selection criterion for the dependent variable,[6] somewhat different results with the effect of spatially lagged vacancies significantly positive in the baseline specification, significantly negative for non-border regions only. The effect of spatially lagged unemployment is either insignificant or negative.

The negative effect of matches in neighboring regions on the number of successful matches in a given region hints at competition for matches between regions. It seems that regions seem to fare better if their neighboring regions experience low new job creation rates. This is not quite what one would expect. In particular, this finding means that there is negative spatial autocorrelation among regions with respect to matches. However, the picture becomes a bit more differentiated once one instruments spatially lagged matches using spatially lagged unemployment and vacancy levels. The negative effect of unemployment in neighboring regions seems to catch a cyclical effect: the higher the unemployment rates in other regions, the worse the economic situation, resulting in fewer matches. This argumentation seems validated by the fact that unemployment rates are spatially autocorrelated, as was reported before. Moreover, the finding hints at congestion effects, since, if a certain number of vacancies is to be filled, more non-resident unemployed job applicants crowd-out local applicants thereby decreasing the efficiency of the matching process. On the other hand, the positive effect of labor demand conditions in neighboring regions, as measured by vacancy rates, seems to express cyclical contingencies between regions: If firms are

[6]That is, whether district dummies or dummies for macro regions or only non-border regions are included.

willing to create more jobs and thus post more vacancies, this is positively correlated to the number of matches also in neighboring regions. This finding is corroborated by the positive spatial autocorrelation found for vacancies in the preceding section.

The results for the matching functions (including a constant and a linear time trend) exhibit highly significant, positive coefficients for both stocks, unemployed and vacancies for all concepts of flow data used as dependent variable. The time trend is significantly negative in all panel specifications. Overall, the significant effects of spatially lagged variables suggest that estimation results obtained with conventional matching functions neglecting spatial dependencies are biased.

Unlike previous studies of spatial matching functions, like Burda and Profit (1996) and Burgess and Profit (2001), we are able to distinguish labor market flows along several dimensions. When decomposing flows by source of origin, it turns out that while the elasticity of the respective concept of matches with respect to unemployment, $\hat{\alpha}$, is roughly the same as the elasticity with respect to vacancies, $\hat{\beta}$, or slightly smaller, $\hat{\alpha}$ is larger than $\hat{\beta}$ if flows out of unemployment into employment are considered. On the other hand, $\hat{\alpha}$ is smaller than $\hat{\beta}$ if job-to-job changes are regressed. These differences can be expected as a result of misspecification stemming from omitting relevant unobservable explanatory variables in the estimation. A discussion of the underlying mechanisms leading to these results is beyond the scope of this paper.[7] However, it is worth noting that spatially lagged

[7]See Sunde (2002) for a formal treatment of the bias resulting from an omission of unobservable endogenous search on both sides of the labor market. Another interpretation

unemployment has a consistently negative effect on matches regardless of the flow concept used as dependent variable. On the other hand, spatially lagged vacancies affect all hires, and hires from employment positively, but hires from unemployment significantly negatively. This can be interpreted as evidence that higher job creation activity elsewhere leads more unemployed to search elsewhere for jobs, and thus causes more regional emigration, indicating negative spatial autocorrelation in the reverse direction as discussed above.

The data allow us to investigate these issues further by checking whether this pattern remains once one considers regional heterogeneity among the new matches. Indeed, the differences are qualitatively the same, and quantitatively even slightly stronger when only matches of individuals who stayed within the same region (m_h) are considered.[8] The same is true for matches of individuals immigrating from neighboring regions. In contrast, results of coefficient estimates for flows from different labor market status do not differ for individuals immigrating from non-neighboring regions: $\hat{\alpha}$ is always slightly smaller than $\hat{\beta}$. These findings suggest that for intra-regional migration or migration between contiguous regions, labor market status has crucial effects on demand and supply elasticities, and therefore in some sense segments the labor market. On the other hand, status matters a lot less for far-distance migrants. These results are also broadly robust to estimations

of the data indicates the relevance of an adverse selection mechanism, see Kugler and Saint-Paul (2001).

[8] Regionally decomposed employment flows will be investigated in more detail in section 4.2. Results for regional decomposition of flows are presented in Table 3, albeit for a somewhat different specification.

which only use cross-sectional variation (year-by-year).

In order to infer more about the structure of inter-regional dependencies in job creation, we estimate spatially augmented matching functions separately for clusters and hot spots and confront them with the results obtained from the pooled sample. Of specific interest is the comparison of the results when instrumenting non-resident matches by non-resident unemployed job seekers and vacancies. While for the pooled sample, spatially lagged unemployment has a significantly negative effect on new hirings, the effect of spatially lagged vacancies is insignificant. If one concentrates on clusters, the effect of spatially lagged unemployment becomes significantly positive: unemployed from neighboring regions search all regions that form a cluster for new employment, and accept jobs they get offered. Vacancies in neighboring regions again play no significant role. The opposite is true for hot spots: spatially lagged unemployment decreases job creation in a given hot spot region significantly (and to a greater extent than in the regression for the pooled sample), likewise do spatially lagged vacancy levels. This result could be expected given the negative spatial autocorrelation of hot spots, and the fact that the pools of unemployment and vacancies in spatially contiguous regions both affect job creation in these contiguous regions positively.[9]

4.2 The Effect of German Re-unification

German re-unification has had a huge impact on German labor markets. When analyzing regional migration and job creation behavior, this event

[9]detailed results are available from the authors upon request.

cannot be neglected. The question is whether re-unification has had any impact on inter-regional migration, e.g. because individuals started migrating to Western Germany for jobs trying to avoid unemployment or increase their salary. We approach this issue by regressing regional matching functions of the form of Equation (1) with an additional dummy for the post-reunification period. Since the data cover the years 1980 until 1997, the dummy takes the value zero for the years 1980 to 1989, and one for the later years. Table 3 contains results for different specifications of the dependent variable. The results of these regressions are striking. The effect of re-unification on all hirings turns out insignificant for both specifications, with spatial lags defined as affecting contiguous regions (neighbors) and non-contiguous regions (other regions which share no common border with the region in question). However, the dummy is highly significant and positive for matching functions with matches from non-neighboring regions, m_f, as dependent variable, and significantly negative for matches from neighboring regions, m_n, as regressand.[10] This reflects the fact that migration from Eastern Germany indeed played an important role in the aftermath of re-unification. The negative effect on matches from contiguous regions originates from the fact that the source regions of flows from East German regions have by convention in the

[10]Note that matches from non-contiguous regions are regressed on spatially lagged explanatory variables. For obvious reasons, spatial lags apply to non-contiguous regions in this case. On the other hand, for the specification with hirings from neighboring regions as dependent variable, explanatory variables are spatially lagged with lags pertaining to contiguous regions.

creation of the data set no common borders with West-German regions.[11] Intensified flows from Eastern Germany therefore decreased the importance of 'neighboring migration'. Corroborating this is the finding that hirings of locals, m_h, have been negatively affected by re-unification. Further results not contained in the table suggest that if matches won by non-locals as a share of all matches or the ratio of non-local matches over local matches are taken to be the dependent variable, the re-unification dummy is highly significant and positive. This provides again strong evidence that overall regional mobility increased significantly as a consequence of the political process. Our results also confirm evidence provided by Hunt (2000) who also finds that there was substantial emigration from East to West Germany in the aftermath of re-unification. Interestingly, the coefficients of spatially lagged unemployment is not significantly different from zero if 'far-distance' migration m_f is concerned, while the coefficient for spatially lagged vacancies is significantly negative, but relatively small. We take this as evidence for economy-wide cyclical effects.

5 Efficiency of Regional Matching

A considerable number of empirical matching studies investigates the variation in the matching process across regions, see Petrongolo and Pissarides (2001) for an overview. However, evidence about the efficiency of the matching process and its determinants, in particular in the regional context is

[11] As a consequence, even if workers move from a contiguous Eastern region into a Western region, this would be recorded as a hiring from a non-contiguous region.

	Dependent Variables:				
***y* = ln(**	***m***	***m***	**m_h**	**m_n**	**m_f)**
ln(*U*)	0.553	0.473	0.565	0.488	0.467
	(0.023)	(0.018)	(0.025)	(0.033)	(0.019)
ln(*V*)	0.387	0.429	0.361	0.307	0.475
	(0.022)	(0.018)	(0.024)	(0.031)	(0.019)
ln(*WU*)	-0.200		-0.254	0.094	
	(0.026)		(0.029)	(0.038)	
ln(*WV*)	0.015		-0.084	0.396	
	(0.024)		(0.027)	(0.038)	
ln($\underline{W}U$)		-0.004			-0.004
		(0.003)			(0.003)
ln($\underline{W}V$)		0.008			0.009
		(0.004)			(0.004)
Linear time trend	-0.032	-0.040	-0.012	-0.028	-0.056
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)
re-unification	-0.006	0.001	-0.115	-0.188	0.125
dummy	(0.025)	(0.025)	(0.030)	(0.044)	(0.027)
Constant	6.700	5.987	5.380	-0.445	6.275
	(0.214)	(0.207)	(0.261)	(0.383)	(0.219)
R^2	0.827	0.815	0.756	0.590	0.813
Observations	2106	2106	2106	2105	2106

Note: Robust standard errors are in parentheses. Data contain observations for 117 regions and 18 years (1980-1997). Legend: *U* unemployment level, *V* vacancy level, *WU* unemployment levels in neighboring regions, *WV* vacancy levels in neighboring regions, $\underline{W}U$ unemployment levels in other, non-neighboring, regions, $\underline{W}V$ vacancy levels in other, non-neighboring. regions; *y* denotes dependent variable, which is the log of the respective concept of new hires: *m all* all hirings, m_h all hirings from same region, m_n from neighboring region, m_f from non-neighboring region.

Table 3: Empirical Matching Functions with Spatial Dependence by Regional Origin and Effects of Re-Unification

scarce. In this section, we analyze the the efficiency of the matching process in a stochastic frontier framework following the approach proposed by Battese and Coelli (1995).[12]

5.1 The Stochastic Matching Framework

Unlike in the conventional matching framework used before, we explicitly take account of inefficiency in the matching process. The inefficiency term can itself be a function of a set of explanatory variables Z. The first element of the Z-vector, Z_0, is a constant. Moreover, the composition of the regional labor force with respect to age and educational background seem relevant factors for explaining matching efficiencies. Therefore, we add the the shares of workforce younger than 25 years and older than 50 years in the respective region, as well as shares of workers with low education and high education, and the ratio of vacancies to unemployed in the respective region measuring labor market tightness as explanatory variables of inefficiency. As low education, we define individuals who neither successfully completed high school (*Abitur*), nor obtained a vocational degree. Individuals have high education, if they hold a degree from a university or an applied university (*Fachhochschule*). Furthermore, Z contains a deterministic time trend. The share of the total variance of the process explained by inefficiency is a measure of the importance of inefficiencies. A prediction of the matching

[12] To our knowledge, only two other contributions apply a stochastic frontier approach to regional data in a matching context. Ibourk *et al.* (2001) analyze the job creation patterns using French data, while Ilmakunnas and Pesola (2002) study matching efficiency for Finland.

efficiency of a particular regional labor market can be obtained by inserting the respective coefficient estimates into the stochastic matching frontier function.

Table 4 presents the main results for specifications with all hirings of non-employed individuals as dependent variable.[13] The data allow to identify matches with respect to the region of origin of the hiree. Column (1) presents results for a specification with all hirings of non-employed individuals in a given period and in a given region as dependent variable. The explanatory variables are the local stocks of unemployed and vacancies, the stocks of unemployed and vacancies in neighboring regions, and the respective stocks in non-neighboring regions. The stocks of job searchers and vacancies in non-neighboring regions do not affect matches significantly. Stocks in neighboring regions tend to significantly decrease matches, presumably because of competition effects between local and non-local job search. Local unemployment and vacancies enter significantly positive. The elasticity of matches from non-employment with respect to unemployed is, with a value of 52 percent, larger than that with respect to vacancies (34 percent). The time trend is significantly negative, indicating a decrease in total matching efficiency over time.

With regard to matching efficiency, the higher the fraction of young indi-

[13]We take this concept of flows as preferred specification since employment inflows of non-employed individuals reflects the relevant stock of unemployed job seekers in the matching function. At the same time, flows from nonemployment contains less potential for mismeasurement than inflows from unemployed, which miss e.g. unemployed individuals in active labor market programs.

	Dependent variable: logarithm of hirings from non-employment per region and year			
	(1) all	(2) from same region	(3) from neighbor region	(4) from non-neighbor region
	ln(m_X)	***ln(m_{Xh})***	***ln(m_{Xn})***	***ln(m_{Xf})***
ln*U*	0.522	0.799		
(Local UE)	(0.018)	(0.015)		
ln*V*	0.341	0.127	0.419	0.473
(Local reg. vacancies)	(0.017)	(0.016)	(0.013)	(0.006)
lnW*U*	-0.100		0.568	
(UE neighbor. regions)	(0.029)		(0.024)	
lnW*V*	-0.099			
(vacancies neighbor. regions)	(0.034)			
ln<u>W</u>*U*	-0.004			0.002
(UE non-neighbor. regions)	(0.004)			(0.001)
ln<u>W</u>*V*	0.005			
(vacancies non-neighbor. regions)	(0.005)			
time trend	-0.028	-0.016	-0.036	-0.046
	(0.002)	(0.006)	(0.003)	(0.006)
constant	4.513	1.581	-0.946	7.743
	(0.135)	(0.219)	(0.247)	(0.118)
Inefficiency term ***Z***:				
constant	2.482	3.313	-5.011	6.080
	(0.504)	(0.271)	(2.133)	(0.043)
fraction young (<25)	-6.035	-5.674	-13.099	-1.230
	(1.310)	(0.318)	(1.416)	(0.214)
fraction old (>50)	0.989	0.932	27.478	0.023
	(0.941)	(0.570)	(1.484)	(0.415)
fraction low education	-1.275	-1.058	1.975	-0.475
	(0.562)	(0.268)	(1.470)	(0.179)
fraction high education	-15.149	1.492	-48.463	-6.559
	(4.020)	(0.927)	(5.324)	(0.423)
tightness (ln(V/U))	0.035	0.002	0.859	0.002
	(0.022)	(0.013)	(0.154)	(0.008)
ln*U*			-6.220	-0.407
(Local UE)			(0.196)	(0.011)
time trend	-0.016	-0.026	-0.107	-0.003
	(0.008)	(0.007)	(0.051)	(0.004)
σ^2_{II}	0.161	0.168	10.298	0.085
	(0.033)	(0.006)	(0.654)	(0.003)
γ_{II}	0.806	0.999	0.993	0.999
	(0.035)	(0.000)	(0.001)	(0.000)
Log (likelihood)	-142.895	-1131.618	-2215.169	-384.954
N	2106	2106	2106	2106

Note: Standard errors are in parentheses. σ^2_I is defined as $\sigma^2_\varepsilon + \sigma^2_R$, γ is defined as $\sigma^2_R/(\sigma^2_\varepsilon + \sigma^2_R)$. A significant positive coefficient for γ indicates that a stochastic production frontier model is superior to simply estimating the model using ordinary least squares. Refer to the text for details. μ denotes the estimated mean of the distribution of the of the error for the technical inefficiency. η accounts for time variance in the efficiencies, specifically for η>0 technical efficiency improved over time while for η<0 the technical efficiency decreases over time.

Table 4: Stochastic Matching Frontier Estimates of Hiring Efficiency of Regions, 1980-1997

viduals in the labor force, the lower the inefficiencies in the regional matching process. The influence of the fraction of old individuals is not significant, however. Somewhat surprisingly, the higher the fraction of people with a low educational background, the lower the matching inefficiency. This might have to do with the fact that these individuals are hired for jobs without particular requirements, and therefore are not screened very carefully, which facilitates the matching. But also the more individuals with high education populate the labor market, the more efficient the matching of unemployed and vacancies. This seems contradictory, but might have to do with the fact that higher search efficiency of highly educated individuals, as well as more directed search on both sides of this segment of the labor market might overcompensate more stringent screening requirements. Note, that quantitatively the effect of the fraction of highly educated is almost twelve times as high as the one for the fraction with low education. Finally, the tighter the labor market, the more inefficient the matching process, presumably since search frictions in the form of coordination problems increase when firms obstruct each other in the search for new employees. Overall, variation in the inefficiency term explains about 80 percent of the total variation of all matches from non-employment.

In column (2), we present estimation results for matches of individuals who were non-employed before encountering the new match, but whose previous employer was located in the same region as their new one. Explanatory variables are local stocks of unemployed and vacancies. Again, both enter significantly positive, but the impact of unemployment is much larger than in the specification for non-employed matches from all regions, with a coefficient estimate of 0.8, while the vacancy elasticity of matches is only around 0.13. The time trend is negative. As for the inefficiency term, all effects are qualitatively the same as for specification (1) with one exception. The fraction of highly educated individuals now tends to increase inefficiency, but this effect is not significantly different from zero. Also in contrast to the results for all matches from non-employment, the variation in the inefficiency term explains virtually all the variation in matches of non-employed, local individuals.

The same result is found for matches of non-employed, who were previously employed in neighboring regions, column (3), and in non-neighboring regions, column (4), as dependent variable. In these two specifications, the flow of new matches is regressed on the stock of local vacancies, and the stock of unemployment in neighboring and non-neighboring regions, respectively. These are the relevant stocks, since employment inflows are recorded in the region under observation, such that only local vacancies can account for their creation. But since the inflows explicitly contain non-employed individuals with origin in neighboring or non-neighboring regions, they must have been contained in the unemployment pool of their respective region of origin, and not the destination region.

While for flows from neighboring regions (column 3) and for flows from non-neighboring regions (column 4) local vacancies significantly increase the number of matches, the relevant stock of unemployed searchers has only a significant positive impact on job creation for matches from neighboring regions.

The determinants of of inefficiencies in the matching process in the specifications to capture competition effects between unemployed in neighboring regions and local unemployed . The estimates show that the higher the local unemployment stock, the more efficient is the matching process with respect to applicants from neighboring regions.

For flows from neighboring regions (column 3) the determinants of matching inefficiencies exhibit some novel features. The positive impact of the number of young workers and of those workers with high education in the labor force in a region, as well as the negative impact of labor market tightness on matching efficiency is strongest for matches from neighboring regions. Only for this specification we find that the number of old workers in the labor market increases the inefficiency in the matching process. Also at odds with the findings for the other specifications is the result that a larger fraction of people with low education background lowers the matching efficiency, although this effect is not significant.

When matches of non-employed from non-neighboring regions are the dependent variable (column 4), the determinants of inefficiencies are qualitatively the same as for the benchmark specification (column 1) with the exception that the tightness of the local labor market has no effect.

5.2 Regional Matching Efficiencies

The regions can be ranked with respect to their matching efficiency estimates. The resulting rankings of the five regions with the highest and the five with the lowest matching efficiencies for the same specifications as displayed in Table 4 are presented in Table 5. Apparently, for all matches from non-employment (specification 5), southern regions around Munich (regions 112, 113, 114) exhibit particularly high matching efficiencies, while rural, thinly populated areas in Northern Germany exhibit the relatively lowest efficiency estimates. The picture changes when one looks at matches from the same region. The highest efficiencies are found in regions which are relatively remote from major urban areas, and which obviously recruit most of their hirees from within the same region. On the other hand, the lowest respective efficiency estimates are found for densely populated areas like in the Ruhr area, or in regions which neighbor major urban agglomeration areas. Interestingly, major cities like Frankfurt (region 51), Bremen (region 7), Cologne (region 39), Hamburg (region 2) and Düsseldorf (region 33) are the five areas exhibiting the highest efficiencies for matches from neighboring regions. Apparently, these cities attract people from surrounding regions, while remote areas close to borders or far away from agglomeration centers exhibit the lowest respective efficiencies. Finally, cities like Hamburg (region 2), Munich (region 112), Frankfurt (region 51), Stuttgart (region 89) and Cologne (region 39) are also those with the highest efficiency estimates for matches from non-neighboring regions, and successfully attract hirees from regions located further away. Again, remote, rural areas exhibit the opposite feature of extremely low matching efficiencies in this respect.

Regressing predicted efficiency estimates on turnover, squared turnover and employment delivers estimation results displayed in Table 6. These

Specification (cf. description in this chapter)	m_X (all) (5)		m_X (from home region) (6)		m_X (from neighbor region) (7)		m_X (from non-neighbor region) (8)	
	region	eff.	region	eff.	region	eff.	region	eff.
Rank 1	113	0.932	113	0.799	51	0.861	2	0.769
Rank 2	112	0.926	93	0.669	7	0.823	112	0.598
Rank 3	114	0.926	103	0.614	39	0.816	51	0.545
Rank 4	93	0.920	58	0.538	2	0.808	89	0.527
Rank 5	19	0.923	96	0.519	33	0.790	39	0.504
Rank 113	10	0.531	40	0.118	22	0.247	3	0.109
Rank 114	64	0.529	28	0.091	17	0.237	110	0.105
Rank 115	16	0.503	8	0.089	79	0.202	17	0.101
Rank 116	17	0.475	17	0.080	3	0.196	16	0.094
Rank 117	8	0.387	80	0.080	116	0.186	8	0.091

Note: Region numbers refer to the regions as listed in Appendix Table A2. Efficiency estimates refer to estimates of *TE*. Ranks performed on average efficiency over 1982-1997.

Table 5: Predicted Efficiencies with Stochastic Frontier Model. Regions with Highest and Lowest Technical Efficiency

results are considerably heterogeneous and not straightforward to interpret. While for all matches from non-employment as dependent variable the matching efficiency is higher the higher the turnover in the respective cell, the opposite is true when more spatial structure is added and matches are distinguished by geographic provenance of the hirees: Inefficiencies increase as turnover becomes higher. With the exception of all matches, the squared turnover effect is insignificant, which might explain some of the differences in the coefficient estimates for the linear term. The higher the employment level in a given region, the less efficient the matching process for non-employed, as well as for non-employed from neighboring and from non-neighboring regions. In contrast, higher employment levels increase the efficiency of the matching process of local job seekers. This result indicates that there might be something like a home field advantage with respect to search frictions and competition for vacant jobs.

Dependent variable in underlying specification: hirings from non-employment	Dependent variable: Average technical matching efficiency of region over the period 1980-1997			
	(9) (all)	(10) (from same region)	(11) (from neighboring region)	(12) (from non-neighboring region)
	TE	***TE***	***TE***	***TE***
Log turnover:	-0.187	0.447	0.178	0.162
ln(m/emp)	(0.052)	(0.188)	(0.107)	(0.050)
Log turnover squared:	-0.117	0.035	0.007	-0.000
ln(m/emp)^2	(0.013)	(0.042)	(0.025)	(0.011)
log employment	0.077	-0.012	0.169	0.178
	(0.003)	(0.004)	(0.007)	(0.003)
constant	0.021	1.188	-1.078	-1.495
	(0.054)	(0.180)	(0.118)	(0.056)
R^2	0.598	0.483	0.191	0.807
N	1872	1872	1872	1872

Note: Standard errors are in parentheses. See text in this chapter for details about the specification. ***m*** is the respective concept of hires in the given region per period, ***emp*** is the level of employment in the given region.

Table 6: Determinants of Regional Hiring Efficiency, 1980-1997

6 Conclusion

This paper investigates spatial dependencies across regional labor markets, in particular with regard to job creation. We find strong evidence for spatial autocorrelation in hirings for some labor market regions in West Germany. In particular, we isolate regions with significantly positive spatial autocorrelation in job creation (clusters), and regions where hirings are characterized by significantly negative spatial autocorrelation (hot spots). Furthermore, the results indicate that conventional empirical matching functions neglecting the spatial component are misspecified. We provide evidence that spatial lag models characterize the matching process better.

The estimation results for spatially augmented matching functions indicate that job creation is negatively affected by job creation in contiguous regions. Spatially lagged unemployment affects the hiring process in a given region negatively. This result is robust for several concepts of flows to employment. Once spatial matching functions are estimated separately for clusters and hot spots, these findings are put into perspective, with spatially lagged unemployment affecting hires positively in clusters, but negatively in hot spots. In general, the findings indicate that the concept of a matching function is empirically confirmed even in the presence of an explicit spatial dimension. German re-unification increased new hires from non-neighboring regions, which include among others also East German regions, significantly.

The efficiency of the matching process exhibits considerable heterogeneity across regions. The results for spatial disaggregation are *per se* interesting, and reveal the importance of considering regional labor markets and inefficiencies in their particular matching processes. The evidence presented in this paper indicates considerable differences between matching processes of individuals with different regional provenance and illustrates the importance of distinguishing flows by their respective source regions.

References

Anselin, L., and A. K. Bera (1999): *Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics*, vol. 155. Dekker.

Battese, G. E., and T. J. Coelli (1995): "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data," *Empirical Economics*, 20, 325–332.

Bender, S., A. Haas, and C. Klose (2000): "IAB Employment Subsample 1975-1995. Opportunities for Analysis Provided by the Anonymised Subsample," .

Burda, M. C., and S. Profit (1996): "Matching Across Space: Evidence on Mobility in the Czech Republic," *Labour Economics*, 3(3), 255–278.

Burgess, S., and S. Profit (2001): "Externalities in the Matching of Workers and Firms in Britain," *Labour Economics*, 8(3), 313–333.

Gorter, C., and J. Van Ours (1994): "Matching Unemployment and Vacancies in Regional Labor Markets: An Empirical Analysis for the Netherlands," *Papers in Regional Science*, 73(2), 153–167.

Hunt, J. (2000): "Why Do People Still Live in East Germany?," *NBER Working Paper*, 7564.

Ibourk, A., B. Maillard, S. Perelman, and H. R. Sneessens (2001): "The Matching Efficiency of Regional Labour Markets: A Stochastic Production Frontier Estimation, France 1990-1995," *IZA Discussion Paper*, 339, IZA Discussion Paper.

Ilmakunnas, P., and H. Pesola (2002): "Matching Functions and Efficiency Analysis," Helsinki School of Economics, Working Paper W-308.

Kugler, A. D., and G. Saint-Paul (2001): "How do Firing Costs affect Worker Flows in a World with Adverse Selection?," *mimeo*.

Petrongolo, B., and C. Pissarides (2001): "Looking Into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39(2), 390–431.

Sunde, U. (2002): "Unobserved Bilateral Search on the Labor Market: A Theory-Based Correction for a Common Flaw in Empirical Matching Studies," *IZA Discussion Paper*, 520.

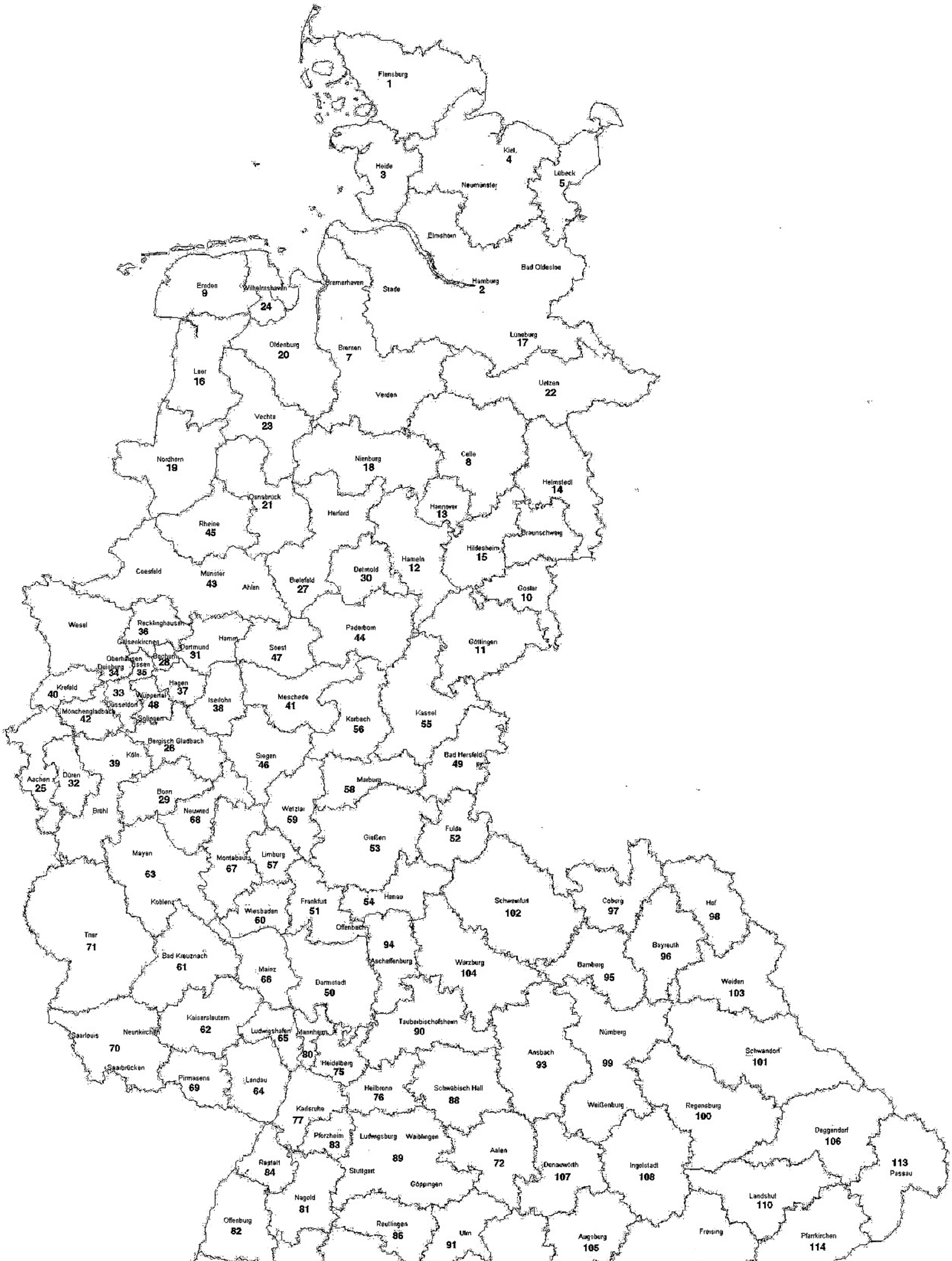

Figure 1: Regional Labor Markets in West Germany