

Schanne, Norbert

**Conference Paper**

## Forecasting Regional Labour Markets with GVAR Models and Indicators

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Schanne, Norbert (2010) : Forecasting Regional Labour Markets with GVAR Models and Indicators, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/119123>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Forecasting Regional Labour Markets with GVAR Models and Indicators

Norbert Schanne\*

14th April 2010

## Abstract

The development of employment and unemployment in regional labour markets is known to be spatially interdependent. Global Vector-Autoregressive (GVAR) models account for the link between the local and the surrounding labour markets and thus might be useful when analysing and forecasting employment and unemployment. Furthermore, GVARs have the advantage to allow for both strong cross-sectional dependence on “leader regions” and weak cross-sectional, spatial dependence.

For the recent and further development of labour markets the economic situation (described e.g. by business-cycle indicators), politics and environmental impacts (e.g. climate) may be relevant. Information on these impacts can be integrated in addition to the joint development of employment and unemployment and the spatial link in a way that allows on the one hand to carry out economic plausibility checks easily and on the other hand to directly receive measures regarding the statistical properties and the precision of the forecasts. Then, the forecasting accuracy is demonstrated for German regional labour-market data in simulated forecasts at different horizons and for several periods.

Business-cycle indicators seem to have no information regarding labour-market prediction, climate indicators little. In contrast, including information about simultaneous labour-market policies and vacancies, and accounting for lagged and contemporaneous spatial dependence can improve the forecasts relative to a simple bivariate model.

KEYWORDS: Global VAR; Labour-market forecasting; Leading indicators; Regional forecasting; Space-time dynamic model

JEL: C 31 ; E 24 ; E 27 ; R 11

---

\*Institute for Employment Research (IAB), Regensburger Straße 104, 90478 Nuremberg, phone +49 (911) 179-5904, E-Mail: norbert.schanne@iab.de

# 1 Introduction

Making predictions on aggregate development of quantities and prices in the markets – GDP, inflation, liquidity demand, or as in this paper, unemployment and job growth – is one of the most important tasks of the economic profession. And much of the recent critique the discipline has to face is due to the fact that the 2008/2009 crisis has not been foreseen by economists, and that many of the forecast revisions made throughout the crises turned out to be wrong as well. This “falsification” of traditional forecasting wisdom by reality made us re-think some of the economic relations we previously thought to be plausible, e.g. labour market’s dependence on expected production and the business cycle.

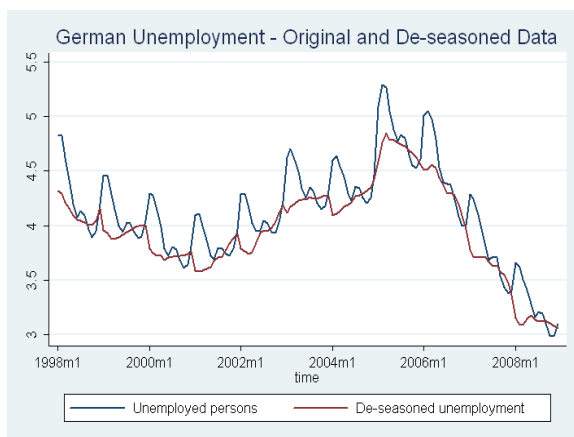


Figure 1: Unemployed Persons in Germany (in millions)

Data: Federal Employment Agency (*Bundesagentur für Arbeit*);

De-seasoned unemployment calculated as  $\tilde{U}_{\text{month,year}} = U_{\text{month,year}} - \bar{U}_{\text{month}}$

When looking at unemployment data for Germany at monthly frequency (shown in Fig. 1) we observe two interesting phenomena:

- The strange development of the German labour market in the current economic situation which is sometimes denoted as the German *job miracle*: the crush of the real-estate bubble and first bank bankruptcies in the US in 2007, the world-wide collapse of financial markets in summer 2008 and the resulting economic crises which is – at the moment – seen as the worst crises since the great depression 70 years ago, a GDP decline of 6 % within one year. And at the same time, unemployment remains at the same level as in 2008, showing the lowest unemployment rate since the early 1990s. Moreover, when the perspective is on a sub-national level, e.g. the federal states, we find (in particular in southern and western Germany) regions which seem to be affected by the crisis whereas others (in northern and eastern Germany) are not.

- The decline in German unemployment from 5 million persons in early 2005 to 3 million people in summer 2008 (with deseasoned data  $\sim 4.8$  mio to 3.3 mio) which seemed to be related to the cycle's upturn from 2004 to 2007. On the first glance, this decline is continuous. However, if one filters the usual seasonal figure out of the development and then takes the de-seasoned unemployment (the red line in Fig. 1) into consideration, two jumps become visible, in the winter 2006/2007 and 2007/2008 – which were distinct from previous and following winters regarding warmth: most parts of Germany didn't see much snow, and the temperature hardly dropped below  $0^\circ$  C. Once again, there is a spatial pattern. Labour-markets improved the most in central, eastern and south-eastern Germany, and only little in the states along the river Rhine.

Previous studies found that considering spatial dependence either contemporaneously (see e.g. Longhi and Nijkamp, 2007) or serially lagged (Hampel et al., 2008 and Schanne et al., 2009) may improve the accuracy of short-term labour-market forecasts. Nevertheless, recently techniques became available that allow a more distinct view on the interrelations between regions by discussing the conditions for either strong or weak cross-sectional dependence (Pesaran and Tosetti, 2007). In addition to determining the dominance of a single region, the setting allows for a certain variability of other variables' impacts across the regions. A similar technique has proven when modeling a multinational monetary system in the so-called Global VAR (Pesaran et al., 2004). As well, this framework is employed to model the spatio-temporal diffusion of shocks on housing prices across regions and to forecast real-estate markets in the US and the UK (Holly et al., 2010*a,b*). Thus, establishing a multivariate model of regional labour market development in a similar manner seems promising.

In this paper, we focus on the development of regional labour market quantities, (log) employment and (log) unemployment at a monthly frequency. As sketched in Section 3, both depend on the share of newly created job matchings (the flows from unemployment into employment, related to past unemployment) and the rate of destructed jobs (flows from employment to unemployment, related to past employment) in the own region, as well as in all other regions (see as well Patacchini and Zenou, 2007)<sup>1</sup>. This can be formalized as a two-variables multi-regional VAR/VEC which is made computationally tractable as a GVAR model, see Section 4. It can be augmented by indicator variables, in order to test their information content with regard to forecasting. Simulated out-of-sample forecasts for the ten regional

---

<sup>1</sup>Because there is commuting between the regions, the place of work does not need to coincide with the place of residence, and the place of a separated job may diverge from the location where the person is registered as unemployed. Thus, the change of employment in location  $i$  can differ from the change of unemployment in the same region.

subdivisions of the Federal Employment Agency (in size roughly equivalent to NUTS-1 regions) are evaluated in Section 5. Indeed, in our setting the prospective information regarding the labour market which is provided by business-cycle indicators turns out to be extremely limited, even smaller than the contribution some climate series can make to labour-market forecasting.

## 2 The Data and Their Statistical Properties

Information on labour market quantities is provided at various regionally disaggregated levels by the German Federal Employment Agency (FEA, *Bundesagentur für Arbeit*). Our series on unemployment and employment stem from register data. Unemployment covers all persons officially considered as unemployed: they receive unemployment benefits from the FEA, look for a job and are up to take on a job. Our metric for employment covers all employees in full- and part-time jobs liable to social security contributions. The analysis is carried out at the level of the Federal Employment Agencies Regional Subdivisions, which we abbreviate with RD. These are equivalent or slightly larger than the German federal states, often fusing two smaller states. Some descriptive statistics and a stationarity analysis is provided in Table 1. Here, in the average monthly change in employment a spatial pattern arises again, with declining employment in the eastern regions (BB, S, SAT), only little change in the north and west, and employment growth in the south (BW, BY). Furthermore, the reported ADF tests and HEGY tests always reject stationarity at the first lag (the zero frequency), whereas stationarity at the seasonal frequencies is not rejected; this finding is supported by other (not reported) unit-root tests. Thus, all series should be treated as non-stationary. Shocks in the regional labour markets can be considered as persistent.

To approximate the business cycle expectations, we use a set of publicly available national indicators: the Stock Exchange Index (DAX), Consumer Price Index and Wholesale Index (provided by the German Federal Bank<sup>2</sup>), survey based indicators regarding business situation and sentiments (two indicators gained from a survey amongst financial experts, provided by the ZEW Mannheim<sup>3</sup>; two from a management survey, provided by the ifo institute Munich<sup>4</sup>); there are no production related regional indicators available at monthly frequency. To measure the effect of climate, we use a set of publicly available metrics on Temperature, Sun-shine, Wind force and Precipitation collected by the German Climate Service (*Deutscher Wetterdienst*)

---

<sup>2</sup>See [http://www.bundesbank.de/statistik/statistik\\_zeitreihen.php](http://www.bundesbank.de/statistik/statistik_zeitreihen.php).

<sup>3</sup>See <ftp://ftp.zew.de/pub/zew-docs/div/konjunktur.xls>.

<sup>4</sup>In april 2010, the series Geschäftsbeurteilungen and Geschäftserwartungen are available at [http://www.cesifo-group.de/link/ifo\\_geschaeftsklima\\_lr\\_03\\_2010.xls](http://www.cesifo-group.de/link/ifo_geschaeftsklima_lr_03_2010.xls).

Table 1: Descriptive Statistics and Unit root tests

RD	Trend (first differences)			HEGY-tests		ADF-tests	
	Mean	Std.Dev.	Mean $\Delta Y$	$t_{(\pi_1)}$	$F_{(\pi_2-\pi_{12})}$	$t_{(DF),Y}$	$t_{(DF),\Delta Y}$
Log employment, 1/1996-7/2008							
BB	14.44	.0592	-.00091	-1.528	783.878	-1.731	-7.282
BW	15.14	.0228	.00033	-2.828	1098.426	-1.123	-10.584
BY	15.27	.0295	.00062	-2.742	867.914	-1.656	-7.852
H	14.57	.0241	.00018	-2.670	550.046	-1.456	-8.774
Nord	15.56	.0237	.00002	-2.522	677.383	-1.351	-8.958
NRW	14.78	.0224	.00017	-2.284	304.276	-2.146	-8.250
NSB	14.56	.0288	-.00010	-2.624	1019.210	-1.674	-6.469
RPS	14.23	.0212	.00023	-2.871	517.828	-1.803	-8.089
S	14.18	.0672	-.00098	-1.691	466.909	-1.521	-6.471
SAT	14.26	.0762	-.00124	-1.400	375.732	-1.446	-6.623
Log unemployment, 1/1991-9/2008							
BB	13.03	.1633	.00140	-1.298	782.452	-2.842	-10.563
BW	12.60	.2355	.00110	-3.126	1852.208	-1.967	-10.258
BY	12.80	.2418	-.00077	-2.122	976.624	-1.980	-7.888
H	12.28	.2244	.00173	-2.884	1089.350	-2.151	-10.080
Nord	13.59	.1583	.00088	-2.854	1198.573	-1.822	-9.452
NRW	12.88	.1636	.00007	-2.723	680.781	-1.807	-9.163
NSB	12.77	.1362	.00006	-1.993	632.703	-2.089	-9.114
RPS	12.12	.1797	.00036	-2.611	932.466	-1.826	-9.259
S	12.76	.1670	.00101	-1.246	696.161	-3.193	-10.748
SAT	12.96	.1491	.00026	-1.012	706.236	-2.752	-10.172

BB: Berlin & Brandenburg – BW: Baden-Wurttemberg – BY: Bavaria – H: Hesse – Nord: City of Hamburg, Mecklenburg-Western Pommerania, Schleswig-Holstein – NRW: Northrhine-Westfalia – NSB: City of Bremen & Lower Saxony – RPS: Rhineland-Palatinate & Saarland – S: Saxony – SAT: Saxony-Anhalt & Thuringia

HEGY-tests are carried out with seasonal dummies and a constant (without deterministic trend), see Beaulieu and Miron (1993).

at 40 stations all over Germany<sup>5</sup>. In addition, we test also the information content of some labour market series on vacancies and participants in active labour-market policies (ALMP); the series are even register data collected by the German Federal Employment Agency at the same regional level as the employment and unemployment series. A complete list of the tested indicators follows from Table 4. Information on certain features of the data will be provided throughout the following sections, in subsequence to the corresponding description of the estimation technique; here we will report just some descriptions and the stationarity analysis of our two variables of interest.

### 3 A sketch of cross-regional labour market dynamics

The standard dynamics in a search-matching framework (see, for example, the textbook version of Cahuc and Zylberberg, 2004, Ch. 9.3) can be adapted in a multi region model such that the unemployment change equation is

$$(1) \quad \Delta U_{i,t} = \Delta N_{i,t} - [m(\theta_i)(1 - c_i) + m(\theta_i^*)c_i] U_{i,t-1} \\ + \delta_i(1 - c_i)L_{i,t-1} + \delta_i^*c_iL_{i,t-1}^*$$

where  $U_{i,t}$  denotes unemployment in region  $i$  at time  $t$ .  $N_{i,t}$  is the labour force (for simplicity, all entering persons are assumed to start as unemployed job-searchers, all retiring persons are supposed to leave from unemployment).  $m(\theta)$  is the matching (job-creation) function depending on labour-market tightness at home  $\theta_i$  or abroad  $\theta_i^*$  weighted with the probability to work at home  $1 - c_i$  or to commute  $c_i$ . Likewise,  $L_{i,t}$  is employment and the parameters  $\delta_i$  the job-separation rate; an asterisc marks variables abroad. Analogously, the employment change equation is

$$(2) \quad \Delta L_{i,t} = -\delta_iL_{i,t-1} + m(\theta_i)(1 - c_i)U_{i,t-1} + m(\theta_i)c_i^*U_{i,t-1}^* \quad .$$

Putting both equations together, this can be written as

$$(3) \quad \begin{pmatrix} \Delta U_{i,t} \\ \Delta L_{i,t} \end{pmatrix} = \begin{bmatrix} -m(\theta_i) - [m(\theta_i^*) - m(\theta_i)]c_i & \delta_i(1 - c_i) \\ m(\theta_i)(1 - c_i) & -\delta_i c_i \end{bmatrix} \begin{pmatrix} U_{i,t-1} \\ L_{i,t-1} \end{pmatrix} \\ + \begin{bmatrix} 0 & \delta_i^*c_i \\ m(\theta_i)c_i^* & 0 \end{bmatrix} \begin{pmatrix} U_{i,t-1}^* \\ L_{i,t-1}^* \end{pmatrix} + \begin{pmatrix} \Delta N_{i,t} \\ 0 \end{pmatrix} \quad .$$

Note that the structure of this model is similar to a first order VEC model. Postmultiplying (3) with  $\begin{pmatrix} U_{i,t-1}^{-1} & 0 \\ 0 & L_{i,t-1}^{-1} \end{pmatrix}$  results in a model

---

<sup>5</sup>See [http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?\\_nfpb=true&\\_pageLabel=\\_dwdwww\\_klima\\_umwelt\\_klimadaten\\_deutschland](http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_klima_umwelt_klimadaten_deutschland).

for the growth rates (or the difference of the logs) depending on the ratio of previous unemployment at home and abroad, the matching rate and the job-separation rate which can be assumed to depend on the business cycle, labour market policy etc.

## 4 Estimating a system of regional labour markets

### 4.1 The Global VAR formulation

Vector Autoregressions (VAR) are known to perform exceptionally well in forecasting multiple interdependent time-series, and thus seem adequate in our case. Let  $y_{it} = (\ln U_{it}, \ln L_{it})'$  denote the vector of interest, and  $\xi_{it} = B_i x_{it}$  be the information on unemployment and employment provided by the indicators available at time  $t$ ; for notational simplicity,  $\xi_{it}$  contains the deterministic mean (modeled by a constant plus seasonal dummies) as well. Then, with  $\mathbf{Y}_t = (y'_{0t}, \dots, y'_{Nt})'$  and  $\Xi_t = (\xi'_{0t}, \dots, \xi'_{Nt})'$ , a VAR over all regions can be written as

$$(4) \quad \mathbf{Y}_t = \Phi_1 \mathbf{Y}_{t-1} + \Phi_2 \mathbf{Y}_{t-2} + \Xi_t + \Upsilon_t,$$

or, rewritten in VEC form, as

$$(5) \quad \Delta \mathbf{Y}_t = \Pi \mathbf{Y}_{t-1} + \Gamma \Delta \mathbf{Y}_{t-1} + \Xi_t + \Upsilon_t \quad .$$

These equation systems are not estimable unrestrictedly if the number of regions is not extremely small. To impose restrictions, we assume that for most regions their cross-sectional impact on other regions is not dominant, i.e. they contribute only little to the explanation of another region (relative to the joint influence of all other regions). The labour market in region  $i$  can be considered to depend on a weighted average over these non-dominant region instead of the particular development of each region  $j = \{0, 1, \dots, i-1, i+1, \dots, N\}$ . Variation in the strength of dependence across regions can be modeled by various predetermined or exogenous metrics for proximity between regions  $i$  and  $j$ . These weights  $w_{ij,k}$  may reflect geographical, cultural, social or economical distance. Different weights can be used for different variables, although typically the same weights are applied for all  $m$  elements of the vector  $y_{it}$ ; thus we define the  $m \times m$  matrix block  $w_{ij} = w_{ij,k} I_m$ . The weights need to satisfy the ‘‘smallness’’ or ‘‘granularity’’ conditions (see Pesaran and Tosetti, 2007), i.e.  $\sum_{i=1}^N w_{ij,k}^2 = O(N^{-1})$ . The matrix containing these weights has bounded row and column sums, i.e.  $\sum_{i=0}^N w_{ij,k} \leq c$  and  $\sum_{j=0}^N w_{ij,k} \leq c$ ; often the weights matrix is row-standardized (the row sums are always unity). Let the *local average* in region  $i$  be denoted with the star variables  $y_{i,t}^* = \sum_{j=0}^N w_{ij} y_{j,t}$ . On the other hand, there may be also dominant regions,



i.e. units in space which are sufficiently large/important such that their impact is notable in all (or, at least most) other regions – e.g. London for the UK or Paris for France. For these regions it might be preferable to consider their impact explicitly.

Then, if there is no dominant unit, a system for a single region can be extracted from the VEC form (which is used here for convenience) as

$$(6) \quad \Delta y_{i,t} = h_i y_{i,t-1} + h_i^* y_{i,t-1}^* + a_i \Delta y_{i,t-1} + g_i \Delta y_{i,t-1}^* + \xi_{i,t} + \varepsilon_{i,t} \quad .$$

In contrast, if there is a dominant region (e.g.  $i = 0$ ), the vector  $y_{0,t}$  is added as a variable to all other regions, and the region-specific equations become

$$(7) \quad \Delta y_{i,t} = h_i y_{i,t-1} + h_i^* y_{i,t-1}^* + h_i^0 y_{0,t-1} + a_i \Delta y_{i,t-1} + g_i \Delta y_{i,t-1}^* + \xi_{i,t} + c_i^0 \Delta y_{0,t} + \varepsilon_{i,t} \quad .$$

For the dominant region itself, eq. (6) remains valid. It can be seen that the region specific systems give (for the case of one dominant region) the VEC by defining  $\mathbf{w}_i = (w_{i0}, \dots, w_{iN})'$ ,

$$A_1 = \begin{pmatrix} a_0 & 0 & \cdots & 0 & 0 \\ 0 & a_1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & a_N \end{pmatrix}, C_0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ c_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{N-1} & 0 & \cdots & 0 \\ c_N & 0 & \cdots & 0 \end{pmatrix},$$

$$G_1 = \begin{pmatrix} g_0 \mathbf{w}'_0 \\ g_1 \mathbf{w}'_1 \\ \vdots \\ g_{N-1} \mathbf{w}'_{N-1} \\ g_N \mathbf{w}'_N \end{pmatrix}, H = \begin{pmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1^0 & h_1 & 0 & \cdots & 0 \\ h_2^0 & 0 & h_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ h_N^0 & 0 & \cdots & 0 & h_N \end{pmatrix} + \begin{pmatrix} h_0^* \mathbf{w}'_0 \\ h_1^* \mathbf{w}'_1 \\ h_2^* \mathbf{w}'_2 \\ \vdots \\ h_N^* \mathbf{w}'_N \end{pmatrix} \quad .$$

It easily can be seen that system (5) follows from setting  $\Upsilon_t = R\varepsilon_t$  with  $R = (I_{m(N+1)} - C_0)^{-1}$ ,  $\Pi = RH$  and  $\Gamma = R(A_1 + G_1)$ ; to get the VAR described in system (4), write  $\Phi_1 = (\Pi + \Gamma - I_{m(N+1)})$  and  $\Phi_2 = -\Gamma$ . For the error covariance matrix we assume that

$$(8) \quad \Sigma_{\Upsilon_t} = R' \begin{pmatrix} \sigma_0^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,N} \\ 0 & \sigma_{2,1} & \sigma_2^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \sigma_{N-1,N} \\ 0 & \sigma_{N,1} & \cdots & \sigma_{N,N-1} & \sigma_N^2 \end{pmatrix} R \quad ,$$

with  $\sigma_{ij}$  the system cross-covariance between the regions and  $\sigma_i^2$  the variance-covariance of the system within a region. I.e. the dominant region is considered to be stochastically independent from the other regions, whereas for all non-dominant regions the errors may be correlated. The unit-specific

systems given by (6) and (7) can be estimated region-by-region. Since the number of parameters per unit is limited, the estimations are computationally tractable. What needs to be answered ex-ante is whether there are one or more dominating regions, and if so, which regions are dominant.

## 4.2 Regional dominance in cross-sectional dependence

It can be tested for regional dominance – in terms of a joint factor dominating cross-sectional dependence – by analysing the variance-covariance matrix of the unobserved heterogeneity (the residuals) across the regions. If all row and column norms of the above matrix  $R$  are bounded in  $N$ , i.e. if they satisfy the granularity condition, there is weak cross-sectional dependence. If a (column) norm is unbounded, there must be a dominating region on which all others strongly depend – and the unbounded column in the matrix  $R$  can be determined easily. A necessary condition for the existence of an unbounded column is that the largest eigenvalue of the residuals variance-covariance matrix  $\lambda_{max}(\hat{\Sigma}_{\gamma_t|\mathcal{I}_{t-1}}) \geq O(N)$ , see Pesaran and Tosetti (2007); thus if the eigenvalues are sufficiently small, the columns are bounded and there is only weak cross-sectional dependence.

Table 2: Maximum eigenvalues of  $\hat{\Sigma}_{\gamma_t|\mathcal{I}_{t-1}}$

Information set	Employment	Unemployment
no information	5.6397	8.3161
SURE-AR(1)	0.0001	0.1334
VAR(1)	0.0001	0.0160

The maximum eigenvalues of the variance-covariance estimates for some estimations are provided in Table 2. Analysis is carried out in a univariate setting, separately for (log) employment and unemployment. The first line gives, just for comparison, the unconditional covariance of the variables across the regions. The second line reports the maximum eigenvalues of  $\hat{\Sigma}_{\gamma|\mathcal{I}}$  in a system of seemingly unrelated regional AR(1) equations (a SURE), the third line the maximum eigenvalues of the residuals covariance matrix in a VAR(1) (with lagged interdependence across the regions). For both conditional estimations on the series of the 10 regional FEA subdivisions, the eigenvalue is smaller than 1; this indicates that there is no dominant region. Hence, we can estimate a model without the rows and columns referring to region 0.

### 4.3 The matter of cointegrating relations

We have seen in Table 1 that regional employment and unemployment (in logs) can be considered as non-stationary series, integrated of order 1. Likely, two series for a region, or a series and its corresponding local average follow a joint stochastic trend; cointegration between these series may arise (see e.g. Lütkepohl, 2001, Lütkepohl, 2005). I.e., the matrices  $\Pi$  and  $H$  in system (5) would be rank-deficient, and they were determined by the errors of the long-run relations between the series and the “loading vector” of these errors. However, to make use of these relations their number has to be known, and it has to be ascertained which are the cointegrating variables. For this, the rank of the system’s coefficient matrix  $\Pi$  and the rank of its subsystems are relevant.

Table 3: Cointegration rank tests

RD	$U, L, U^*, L^*$	$U, L$	$U, U^*$	$L, L^*$
Johansen trace statistics				
BB	0	0	0	1
BW	3	1	1	1
BY	1	0	1	0
H	1	0	1	0
Nord	2	0	0	0
NRW	1	0	1	1
NSB	3	0	1	1
RPS	2	0	1	1
S	1	0	1	0
SAT	2	0	1	0

Reported results refer to the period 1996/01 to 2008/07. Test results regarding the number of cointegrating relations may vary if the sampling period changes.

Table 3 reports the commonly used Johansen trace statistics for the number of cointegrating relations per region. Here, for the full sampling period, the number of cointegrating relations varies across the regions from 0 to 3. Almost certain is cointegration between unemployment and its spatial lag; on the other hand, a joint trend of unemployment and employment can be rejected mostly. However, the findings are somewhat contradictory. First there seem to be cointegrating relations in a subsystem without cointegration in the whole system, or cointegration in the complete system with a smaller number of cointegrating subsystems. As well, the tests for the whole system (reported in the first column) are not too robust when changing the sampling period. Thus, because of the ambiguous evidence regarding cointegration and the fact that the forecasts from a VAR and a VEC are formally equivalent (see Lütkepohl, 2005, Ch. 6.5), we do not pursue the aspect of joint trends

further on<sup>6</sup>; the results presented latter stem from models which do not consider error correction.

#### 4.4 Selection and inclusion of appropriate indicators

This section focuses on the appropriate determination of the component  $\xi_{it}$  in equations 6 for  $i = \{1, \dots, N\}$ . Often it is argued that the inclusion of a small number of indicators with a high information content performs better in forecasting than a larger number with less information (see e.g. Stock, 2001?, or, as an application for Germany Gaggermeier, 2006). A good indicator is, on the one hand, highly correlated with the variable of interest. On the other hand, it should have a certain temporal lead to the variables of interest, such that the relevant observations of the indicator have realized (or, to be more precise, are known) already in the period when the forecast is made. It does not need to be known necessarily, but the forecast variance will be smaller if it is. Nevertheless, the time delay between indicator and target variable should not be too large to be economically reasonable.

To restrict the number of relations tested out, we determine for each indicator the (in absolute value) highest correlation with log employment and, respectively, log unemployment shown at any lead between zero and thirty-six months: The maximum pairwise correlation coefficients are shown in Table 4. The last column in this table reports the sum over the correlation coefficients (absolute values). In the further analysis we report only results for those indicators which show the highest sum per group: The judgements of experts (**zew-lage**) and managers (**ifo-lage**) on the current situation (which are higher correlated with future labour market development than their expectations about the future), the stock-market index (**dax**), Minimum Temperature within a month (**tmm**), cloud amount (**nmm**), total monthly precipitation (**rss**) and average windforce (**fmm**).

For these indicators we then determine the optimum lead, i.e. the time delay for which the highest correlation with regional employment or unemployment can be found. In general, there is no clear timing for the peak in the degree of correlation between two variables. Thus, we determine for each indicator-target variable-region combination the three leads with the highest correlation, which are shown in Table 5.

---

<sup>6</sup>At the problem of uncertain cointegration when forecasting, it is argued by Stock (2001, p. 578): “However, even if cointegration is correctly imposed, it remains to estimate the parameters of the cointegrating vector, which are, to first-order, estimated consistently (and at the same rate) if cointegration is not imposed. If cointegration is imposed incorrectly, however, asymptotically biased forecasts with large risks can be produced.”

Table 4: Maximum correlation (in absolute value) of lagged indicator variables

Indicator	BB	BW	BY	H	Nord	NRW	NSB	RPS	S	SAT	Sum
Correlation of log employment with ...											
log unemployment ( $U$ )	-0.850	-0.896	-0.924	-0.925	-0.890	-0.795	-0.793	-0.782	-0.673	-0.632	
$L^*$	0.961	0.886	-0.642	0.891	0.948	0.902	0.926	0.908	0.932	0.882	
$U^*$	-0.602	-0.844	-0.813	-0.592	-0.666	-0.770	-0.679	-0.857	-0.651	-0.686	
ALMP	0.709	0.658	0.576	0.338	0.405	0.555	0.733	0.628	0.797	0.849	6.248
Vacancies	0.326	0.867	0.755	0.880	0.909	0.908	0.801	0.898	0.735	0.630	7.710
ZEW sentiments	0.304	-0.502	-0.569	-0.295	0.203	-0.235	0.261	-0.362	0.342	0.384	3.457
ifo expectation	0.293	-0.314	0.284	-0.356	-0.381	-0.336	0.372	-0.321	0.273	0.245	3.175
ZEW situation	0.369	0.514	0.530	0.525	0.615	0.626	0.648	0.583	0.279	0.276	4.967
ifo situation	-0.321	0.508	0.535	0.463	0.491	0.558	0.467	0.521	-0.279	-0.342	4.484
DAX (stock market)	-0.342	0.845	0.832	0.853	0.702	0.756	0.376	0.855	-0.422	-0.387	6.367
CPI (consumer prices)	-0.838	0.563	0.688	0.211	-0.281	0.066	-0.541	0.276	-0.864	-0.900	5.227
Whole sales Index	-0.197	0.160	0.181	0.089	-0.053	0.055	-0.124	0.094	-0.202	-0.210	1.364
Temperature (tnn)	-0.053	0.081	0.093	0.089	0.077	0.124	-0.073	0.095	-0.051	-0.062	.798
Temperature (tnm)	-0.039	0.066	0.067	0.081	0.078	0.117	-0.037	0.086	-0.029	-0.047	.646
Temperature (txm)	-0.055	0.067	0.073	0.043	-0.030	0.050	-0.043	0.047	-0.059	-0.067	.535
Temperature (txx)	0.026	0.023	0.041	-0.017	-0.026	0.019	-0.035	-0.035	-0.036	-0.042	.300
Sunshine hours (sos)	-0.094	0.049	0.072	-0.041	-0.109	-0.060	-0.079	-0.061	-0.118	-0.118	.802
Cloud amount (nmm)	0.152	-0.046	-0.064	0.101	0.170	0.149	0.102	0.094	0.131	0.196	1.206
Precipitation (rss)	-0.104	0.128	0.139	0.262	0.179	0.151	0.107	0.268	-0.054	0.082	1.474
Precipitation (rsx)	0.098	-0.091	0.112	0.135	0.072	-0.154	0.062	0.144	0.064	-0.138	1.071
Wind force (fmm)	-0.057	0.308	0.335	0.246	0.193	0.199	-0.068	0.302	-0.163	-0.217	2.090
Wind force (fxx)	0.068	0.089	0.126	0.186	0.079	-0.090	0.063	-0.238	-0.124	0.115	1.177
Correlation of log unemployment with ...											
log employment ( $L$ )	0.880	-0.896	-0.924	-0.925	-0.889	-0.795	0.779	-0.782	0.799	0.810	
$L^*$	0.845	-0.829	0.752	-0.905	-0.866	-0.746	0.755	-0.737	0.826	0.834	
$U^*$	0.926	0.933	0.878	0.792	0.751	0.925	0.926	0.935	0.850	-0.743	
ALMP	-0.448	-0.740	-0.563	-0.441	-0.499	-0.665	-0.692	-0.770	0.685	0.690	6.193
Vacancies	-0.733	-0.836	-0.736	-0.956	-0.897	-0.803	-0.848	0.745	0.801	0.841	8.196
ZEW sentiments	0.406	0.612	0.532	0.353	0.263	0.477	0.450	0.608	0.527	0.696	4.923
ifo expectation	-0.598	-0.415	-0.459	0.461	0.404	0.357	-0.570	-0.416	-0.543	-0.545	4.767
ZEW situation	-0.765	-0.799	-0.874	-0.749	-0.725	-0.720	-0.870	-0.796	-0.657	-0.687	7.641
ifo situation	-0.734	-0.741	-0.786	-0.618	-0.605	-0.663	-0.801	-0.747	-0.704	-0.776	7.173
DAX (stock market)	0.349	-0.833	-0.756	-0.696	-0.565	-0.661	-0.533	-0.711	0.197	-0.402	5.703
CPI (consumer prices)	-0.255	-0.473	-0.257	-0.060	0.179	-0.317	-0.250	-0.452	-0.554	-0.701	3.497
Whole sales Index	-0.087	-0.190	-0.127	-0.081	-0.042	-0.141	-0.082	-0.165	-0.143	-0.204	1.262
Temperature (tnn)	0.133	-0.147	-0.141	-0.121	-0.099	-0.156	-0.109	-0.148	0.161	-0.166	1.382
Temperature (tnm)	0.115	-0.122	-0.125	-0.117	-0.107	-0.146	-0.117	-0.139	0.168	-0.166	1.321
Temperature (txm)	0.106	-0.119	-0.116	-0.085	-0.073	-0.112	-0.104	-0.113	-0.167	-0.178	1.173
Temperature (txx)	0.098	-0.068	0.083	0.079	0.058	-0.077	-0.087	0.085	-0.151	-0.148	.9335
Sunshine hours (sos)	-0.111	-0.097	-0.083	0.097	0.123	0.098	0.077	0.106	-0.194	-0.189	1.174
Cloud amount (nmm)	-0.131	0.068	-0.112	-0.139	-0.188	-0.149	-0.086	-0.119	0.157	0.212	1.360
Precipitation (rss)	-0.137	0.147	0.126	-0.214	-0.133	-0.178	-0.180	-0.185	-0.125	0.173	1.597
Precipitation (rsx)	-0.121	0.108	0.122	-0.101	-0.071	0.208	0.094	0.166	-0.117	-0.186	1.296
Wind force (fmm)	0.190	-0.185	-0.157	-0.228	-0.185	-0.158	-0.114	-0.247	-0.189	0.123	1.776
Wind force (fxx)	0.124	-0.045	0.085	-0.109	-0.094	0.092	0.093	0.181	0.184	0.125	1.132

Table 5: Lead of the indicator variables, in months

	BB	BW	BY	H	Nord	NRW	NSB	RPS	S	SAT
	Leads related with unemployment									
ALMP	25, 24, 26	25, 26, 24	25, 24, 26	0, 1, 2	25, 26, 31	19, 20, 18	24, 23, 22	36, 35, 34	31, 30, 32	32, 31, 33
Vacancies	14, 15, 13	4, 3, 5	4, 5, 3	15, 16, 14	14, 15, 13	13, 12, 9	8, 9, 13	36, 35, 34	36, 35, 34	35, 34, 36
DAX	36, 35, 34	24, 25, 26	0, 23, 22	0, 1, 2	0, 1, 2	0, 1, 2	0, 1, 2	0, 1, 2	36, 35, 0	36, 0, 35
ifo sit.	19, 18, 20	15, 14, 16	15, 16, 14	23, 22, 24	20, 19, 21	20, 19, 21	19, 20, 18	16, 17, 15	9, 8, 7	5, 7, 6
ZEW sit.	11, 10, 12	12, 11, 13	12, 11, 13	19, 18, 20	17, 16, 18	16, 17, 15	12, 13, 11	12, 13, 11	9, 6, 10	6, 7, 5
fmm	36, 35, 34	1, 0, 2	1, 0, 2	23, 19, 22	9, 8, 10	36, 31, 35	36, 35, 34	0, 1, 2	20, 21, 19	26, 27, 0
nmm	36, 27, 26	15, 4, 5	15, 14, 16	3, 2, 4	24, 3, 2	23, 22, 2	36, 27, 35	3, 14, 15	27, 26, 36	27, 26, 28
rss	16, 15, 17	34, 35, 33	34, 33, 35	35, 24, 34	23, 24, 25	33, 35, 34	16, 17, 15	36, 35, 34	29, 30, 32	33, 32, 34
tnn	32, 33, 31	33, 32, 34	33, 34, 32	33, 32, 34	33, 32, 34	26, 25, 27	31, 32, 33	33, 32, 34	32, 31, 33	32, 31, 14
	Leads related with employment									
ALMP	31, 30, 32	33, 34, 32	24, 23, 25	0, 1, 2	32, 31, 34	17, 18, 16	24, 25, 26	36, 35, 34	0, 1, 2	0, 1, 2
Vacancies	25, 24, 26	6, 7, 8	6, 5, 7	16, 15, 17	14, 15, 13	12, 13, 11	15, 14, 16	1, 2, 0	15, 14, 16	14, 15, 13
DAX	36, 35, 34	15, 19, 16	13, 14, 12	19, 20, 18	14, 15, 13	17, 16, 15	11, 12, 10	14, 15, 13	36, 35, 34	36, 35, 34
ifo sit.	0, 1, 2	20, 21, 19	16, 15, 14	34, 35, 33	25, 23, 24	23, 24, 25	22, 21, 23	20, 21, 22	0, 1, 2	0, 1, 2
ZEW sit.	18, 19, 17	13, 12, 14	10, 9, 11	16, 17, 15	18, 19, 17	18, 17, 16	16, 17, 18	13, 14, 12	18, 19, 17	19, 20, 18
fmm	35, 14, 16	8, 9, 7	8, 9, 7	21, 17, 16	11, 9, 10	20, 19, 21	35, 34, 33	9, 10, 8	8, 7, 6	35, 36, 34
nmm	4, 5, 3	36, 35, 34	10, 16, 17	22, 25, 21	16, 14, 17	22, 23, 21	4, 9, 16	22, 23, 21	4, 1, 5	4, 5, 3
rss	6, 3, 4	5, 4, 6	8, 4, 5	8, 9, 11	11, 12, 13	17, 8, 15	17, 16, 22	8, 9, 10	3, 0, 1	25, 26, 24
tnn	36, 5, 4	6, 5, 7	10, 9, 8	11, 10, 17	7, 8, 11	17, 16, 18	36, 33, 32	10, 11, 9	8, 5, 7	8, 5, 36

## 5 Forecast evaluation

We evaluate simulated out-of-sample forecasts at the three-months, six-months and twelve-months horizon, with the end of the sampling window rolling from July 2005 to June 2008. I.e., the first three-month forecast is made for October 2005, the last twelve-month forecast predicts the labour market quantities in June 2009. In each estimation we include just a single indicator at one lead. The coefficients are estimated region by region; the coefficients are then inserted in the full model (containing all regions) for forecasting.

Thus, with 36 different sampling periods and 6 leads per indicator, we carry out 216 estimations (and predictions) per forecast horizon and indicator. The average root mean squared forecast errors (RMSFE) over these 288 different predictions are reported separately per regional FEA division in Tables 6 and 7. In addition to the RMSFE per indicator which are always determined from a model containing the local average (the spatial lag, or the star variables), we report also the RMSFE of a simple bivariate VAR of the target variables without spatial interdependencies (the row denoted with “ $U, L$ ”), a model containing the target variables and their local averages, a model containing the target variables plus the (simultaneous) series of log ALMP participants and log vacancies (“ $U, L, A, V$ ”) and a 6-variable VAR with the latter two series, the two target variables and their local averages (“ $U, L, U^*, L^*, A, V$ ”). The latter two models can be interpreted as systems containing more information on the labour market.

The RMSFE reported in Tables 6 and 7 refer to log unemployment and log employment. The numbers are roughly equivalent to the percentage error when forecasting the corresponding levels. I.e., a RMSFE of 0.0045 for the model ( $U, L, U^*, L^*$ ) in Baden-Wurtemberg (BW) at the twelve months horizon (Table 7, second line of the third block, the element in the second column of numbers) means that using this model, employment one year later can be predicted with an error of, on average, less than half a percent.

For most models and regions, the errors of the 12 month forecast are three to four times higher than those at the 3 month horizon. Forecast errors are almost twice as high when we double the forecast horizon as we expect for a non-stationary process: shocks are persistent, and thus the uncertainty accumulates.

When we look at the regions, we find some notable spatial pattern: We seem to forecast unemployment for the regional subdivisions located in eastern Germany (BB, S, SAT) with higher accuracy (with an RMSFE of .06 to .08 at the twelve month horizon) than most parts of western Germany (often with an RMSFE higher than .1), only Lower-Saxony/Bremen (NSB) is in a similar range as eastern Germany. The pattern for the forecast precision in employment is somewhat reverse, with slightly more accurate predictions in western Germany.

Table 6: Root Mean Squared Forecast Error, Unemployment

	BB	BW	BY	H	Nord	NRW	NSB	RPS	S	SAT
Forecast horizon 3										
(U,L)	0.0215	0.0339	0.0436	0.0385	0.0346	0.0417	0.0280	0.0349	0.0252	0.0254
(U,L,U*,L*)	0.0217	0.0317	0.0415	0.0256	0.0219	0.0372	0.0282	0.0278	0.0213	0.0242
(U,L,A,V)	0.0221	0.0353	0.0480	0.0226	0.0315	0.0433	0.0243	0.0349	0.0231	0.0256
(U,L,U*,L*,A,V)	0.0240	0.0327	0.0375	0.0215	0.0164	0.0348	0.0237	0.0247	0.0236	0.0258
ALMP	0.0230	0.0360	0.0433	0.0318	0.0371	0.0471	0.0314	0.0384	0.0236	0.0263
vacancies	0.0215	0.0289	0.0434	0.0346	0.0327	0.0444	0.0281	0.0358	0.0297	0.0276
dax	0.0214	0.0287	0.0437	0.0354	0.0321	0.0439	0.0287	0.0373	0.0298	0.0283
ifo-lage	0.0204	0.0287	0.0433	0.0340	0.0314	0.0446	0.0292	0.0365	0.0283	0.0269
zew-lage	0.0212	0.0287	0.0429	0.0346	0.0317	0.0434	0.0280	0.0362	0.0294	0.0282
fmm	0.0213	0.0285	0.0435	0.0351	0.0318	0.0440	0.0286	0.0364	0.0300	0.0280
nmm	0.0213	0.0287	0.0433	0.0353	0.0316	0.0438	0.0286	0.0360	0.0291	0.0280
rss	0.0200	0.0285	0.0391	0.0341	0.0290	0.0449	0.0236	0.0301	0.0249	0.0274
tnn	0.0217	0.0286	0.0432	0.0357	0.0325	0.0442	0.0289	0.0372	0.0291	0.0262
Forecast horizon 6										
(U,L)	0.0331	0.0614	0.0685	0.0656	0.0582	0.0672	0.0446	0.0581	0.0356	0.0362
(U,L,U*,L*)	0.0343	0.0537	0.0656	0.0424	0.0362	0.0607	0.0485	0.0457	0.0282	0.0314
(U,L,A,V)	0.0332	0.0624	0.0767	0.0358	0.0529	0.0705	0.0368	0.0588	0.0302	0.0355
(U,L,U*,L*,A,V)	0.0409	0.0585	0.0684	0.0343	0.0276	0.0586	0.0404	0.0403	0.0355	0.0355
ALMP	0.0366	0.0660	0.0755	0.0562	0.0634	0.0740	0.0520	0.0715	0.0369	0.0385
vacancies	0.0366	0.0524	0.0688	0.0630	0.0612	0.0742	0.0472	0.0626	0.0465	0.0430
dax	0.0365	0.0521	0.0687	0.0653	0.0605	0.0733	0.0475	0.0653	0.0468	0.0442
ifo-lage	0.0330	0.0521	0.0686	0.0609	0.0594	0.0726	0.0497	0.0635	0.0421	0.0398
zew-lage	0.0363	0.0517	0.0677	0.0641	0.0600	0.0729	0.0474	0.0640	0.0462	0.0439
fmm	0.0366	0.0514	0.0681	0.0653	0.0600	0.0736	0.0478	0.0643	0.0468	0.0437
nmm	0.0363	0.0520	0.0683	0.0653	0.0596	0.0735	0.0479	0.0640	0.0461	0.0434
rss	0.0326	0.0502	0.0629	0.0573	0.0537	0.0694	0.0375	0.0511	0.0419	0.0422
tnn	0.0380	0.0514	0.0688	0.0622	0.0605	0.0728	0.0480	0.0644	0.0460	0.0383
Forecast horizon 12										
(U,L)	0.0583	0.1354	0.1296	0.1192	0.1028	0.1130	0.0792	0.1089	0.0630	0.0610
(U,L,U*,L*)	0.0619	0.1037	0.1209	0.0729	0.0584	0.0962	0.0886	0.0830	0.0470	0.0529
(U,L,A,V)	0.0589	0.1314	0.1492	0.0565	0.0957	0.1163	0.0664	0.1122	0.0534	0.0643
(U,L,U*,L*,A,V)	0.0787	0.1172	0.1545	0.0624	0.0560	0.1002	0.0765	0.0738	0.0618	0.0605
ALMP	0.0634	0.1412	0.1442	0.1054	0.1040	0.1147	0.0867	0.1727	0.0690	0.0667
vacancies	0.0637	0.1183	0.1283	0.1112	0.1036	0.1196	0.0826	0.1120	0.0785	0.0739
dax	0.0681	0.1197	0.1269	0.1162	0.1037	0.1182	0.0827	0.1170	0.0826	0.0773
ifo-lage	0.0583	0.1194	0.1274	0.1105	0.1031	0.1181	0.0879	0.1148	0.0751	0.0661
zew-lage	0.0674	0.1189	0.1250	0.1144	0.1032	0.1183	0.0833	0.1156	0.0819	0.0762
fmm	0.0679	0.1188	0.1262	0.1165	0.1032	0.1190	0.0833	0.1160	0.0824	0.0759
nmm	0.0675	0.1193	0.1264	0.1163	0.1027	0.1188	0.0835	0.1157	0.0822	0.0749
rss	0.0552	0.0958	0.1117	0.0948	0.0976	0.1077	0.0689	0.0864	0.0787	0.0668
tnn	0.0718	0.1174	0.1265	0.1069	0.1031	0.1173	0.0830	0.1136	0.0841	0.0677



Table 7: Root Mean Squared Forecast Error, Employment

	BB	BW	BY	H	Nord	NRW	NSB	RPS	S	SAT
Forecast horizon 3										
(U,L)	0.0046	0.0032	0.0038	0.0034	0.0032	0.0044	0.0038	0.0036	0.0065	0.0066
(U,L,U*,L*)	0.0046	0.0017	0.0030	0.0027	0.0028	0.0029	0.0037	0.0029	0.0061	0.0070
(U,L,A,V)	0.0045	0.0024	0.0037	0.0035	0.0036	0.0038	0.0037	0.0039	0.0063	0.0069
(U,L,U*,L*,A,V)	0.0042	0.0018	0.0031	0.0033	0.0025	0.0029	0.0038	0.0030	0.0050	0.0070
ALMP	0.0046	0.0023	0.0032	0.0035	0.0036	0.0037	0.0033	0.0034	0.0066	0.0070
vacancies	0.0042	0.0023	0.0032	0.0029	0.0028	0.0039	0.0033	0.0032	0.0063	0.0065
dax	0.0043	0.0030	0.0035	0.0031	0.0028	0.0041	0.0034	0.0035	0.0065	0.0066
ifo-lage	0.0041	0.0028	0.0034	0.0031	0.0029	0.0043	0.0035	0.0034	0.0066	0.0066
zew-lage	0.0043	0.0028	0.0034	0.0032	0.0029	0.0041	0.0034	0.0034	0.0065	0.0066
fmm	0.0043	0.0028	0.0034	0.0031	0.0028	0.0041	0.0033	0.0034	0.0066	0.0066
nmm	0.0043	0.0028	0.0034	0.0030	0.0027	0.0040	0.0035	0.0033	0.0065	0.0066
rss	0.0040	0.0018	0.0033	0.0030	0.0027	0.0034	0.0032	0.0032	0.0063	0.0067
tnn	0.0045	0.0028	0.0034	0.0032	0.0028	0.0041	0.0034	0.0036	0.0067	0.0068
Forecast horizon 6										
(U,L)	0.0081	0.0054	0.0061	0.0052	0.0052	0.0074	0.0059	0.0056	0.0109	0.0109
(U,L,U*,L*)	0.0080	0.0026	0.0046	0.0031	0.0047	0.0040	0.0064	0.0042	0.0095	0.0110
(U,L,A,V)	0.0078	0.0038	0.0061	0.0054	0.0061	0.0058	0.0057	0.0069	0.0109	0.0118
(U,L,U*,L*,A,V)	0.0071	0.0027	0.0049	0.0048	0.0038	0.0039	0.0066	0.0046	0.0086	0.0114
ALMP	0.0080	0.0041	0.0053	0.0056	0.0056	0.0061	0.0057	0.0059	0.0118	0.0114
vacancies	0.0075	0.0038	0.0051	0.0040	0.0042	0.0063	0.0056	0.0047	0.0100	0.0105
dax	0.0079	0.0049	0.0057	0.0047	0.0044	0.0069	0.0057	0.0053	0.0110	0.0109
ifo-lage	0.0073	0.0048	0.0056	0.0046	0.0047	0.0071	0.0059	0.0052	0.0109	0.0105
zew-lage	0.0079	0.0047	0.0055	0.0047	0.0045	0.0069	0.0056	0.0054	0.0109	0.0109
fmm	0.0078	0.0047	0.0055	0.0047	0.0044	0.0069	0.0055	0.0053	0.0111	0.0109
nmm	0.0079	0.0047	0.0055	0.0045	0.0043	0.0069	0.0057	0.0052	0.0110	0.0109
rss	0.0066	0.0024	0.0048	0.0045	0.0043	0.0058	0.0053	0.0048	0.0114	0.0112
tnn	0.0083	0.0047	0.0056	0.0047	0.0044	0.0070	0.0055	0.0054	0.0115	0.0113
Forecast horizon 12										
(U,L)	0.0147	0.0127	0.0122	0.0106	0.0115	0.0128	0.0096	0.0109	0.0178	0.0171
(U,L,U*,L*)	0.0149	0.0045	0.0073	0.0048	0.0079	0.0059	0.0116	0.0073	0.0164	0.0184
(U,L,A,V)	0.0143	0.0092	0.0116	0.0104	0.0132	0.0103	0.0094	0.0139	0.0185	0.0193
(U,L,U*,L*,A,V)	0.0134	0.0050	0.0095	0.0083	0.0061	0.0059	0.0122	0.0085	0.0167	0.0203
ALMP	0.0149	0.0095	0.0107	0.0116	0.0109	0.0112	0.0103	0.0153	0.0240	0.0192
vacancies	0.0144	0.0082	0.0092	0.0079	0.0094	0.0105	0.0095	0.0088	0.0161	0.0178
dax	0.0153	0.0102	0.0108	0.0094	0.0096	0.0119	0.0097	0.0106	0.0189	0.0186
ifo-lage	0.0136	0.0101	0.0107	0.0089	0.0103	0.0124	0.0106	0.0105	0.0182	0.0169
zew-lage	0.0152	0.0098	0.0105	0.0092	0.0099	0.0119	0.0094	0.0107	0.0188	0.0186
fmm	0.0151	0.0099	0.0105	0.0094	0.0097	0.0120	0.0093	0.0105	0.0189	0.0186
nmm	0.0152	0.0099	0.0105	0.0091	0.0094	0.0120	0.0096	0.0104	0.0190	0.0185
rss	0.0118	0.0044	0.0083	0.0093	0.0104	0.0114	0.0094	0.0091	0.0237	0.0206
tnn	0.0164	0.0099	0.0108	0.0086	0.0096	0.0123	0.0093	0.0103	0.0200	0.0188

Most interesting is the comparison of the forecast performance across the indicators. We find that, in general, the simple bivariate VAR is never the best model, it can always be improved by adding some information. On the other hand, the models containing one of the indicators are often outperformed by the comparison models  $(U, L, U^*, L^*)$ ,  $(U, L, A, V)$  and  $(U, L, U^*, L^*, A, V)$ . The GVAR model without indicator,  $(U, L, U^*, L^*)$ , is outperformed by an indicator model at all horizons only in the regions BB, BW, BY, and NSB; then the best performing model is typically one augmented by a climate indicator, mostly precipitation (rss). Economic business cycle indicators in contrast seem not be able to contribute to an improvement of accuracy in regional labour-market forecasting; there is just one case (employment in SAT, 12-months horizon) where one of the standard cycle indicators (ifo-lage) leads to a more accurate forecast and where there is no better predictor.

There may be two reasons for these results. First, our models could be overfitting, since the number of estimated coefficients is high in relation to the short observation period; then, too many degrees-of-freedom are used, the coefficient estimates are rather imprecise. Second, the lead of most indicators seems high. Often they show the highest correlation with the regional labour market series after two years or more (see Table 5), a time span that seems quite unreasonable. When a variable improves the forecast, its lead is typically within half a year. On the other hand, a short lead is not a guarantee for an accurate forecast since the correlation (the information content) might be poor.

## 6 Conclusion

The focus of this paper is on forecasting regional labour markets, and our aim is twofold. First, we establish a framework that considers spatio-temporal dynamics in a multivariate setting. The presented model has the advantage to allow for both strong (i.e. dynamic factor) and weak (spatial) cross-sectional dependence. A provided test distinguishes between the two form inside the modeling framework.

Second, we use this framework to analyse the prospective information regarding the economy due to spatial co-development. Furthermore, we examine with the same model the information content of a set of commonly used business-cycle and labour-market indicators, and compare it with the predictive information provided by climate variables. The tested business-cycle indicators seem to have no information regarding labour-market prediction, climate indicators little. In contrast, including information about simultaneous labour-market policies and vacancies, and in particular accounting for lagged and contemporaneous spatial dependence can improve the forecasts relative to a simple bivariate model.

## References

- Beaulieu, J. and Miron, J. (1993), ‘Seasonal Unit Roots in Aggregate U.S. Data’, *Journal of Econometrics* **55**, 305–328.
- Cahuc, P. and Zylberberg, A. (2004), *Labor Economics*, MIT Press, Cambridge (MA), London.
- Gaggermeier, C. (2006), Indikatoren-Modelle zur Kurzfristprognose der Beschäftigung in Deutschland, Forschungsbericht 6, IAB, Nürnberg.
- Hampel, K., Kunz, M., Schanne, N., Wapler, R. and Weyh, A. (2008), Regional Employment Forecasts with Spatial Interdependencies, *in* C. Knobel, B. Kriechel and A. Schmid, eds, ‘Regional Forecasting on Labour Markets’, Rainer Hampp Verlag, München, Mering, chapter 5, pp. 68–88.
- Holly, S., Pesaran, M. and Yamagata, T. (2010a), ‘A Spatio-temporal Model of House Prices in the US’, *Journal of Econometrics* **forthcoming**.
- Holly, S., Pesaran, M. and Yamagata, T. (2010b), Spatial and Temporal Diffusion of House Prices in the UK, Discussion paper 4694, IZA.
- Longhi, S. and Nijkamp, P. (2007), ‘Forecasting regional labour market developments under spatial autocorrelation’, *International Regional Science Review* **30**(2), 100–119.
- Lütkepohl, H. (2001), Vector autoregressions, *in* B. Baltagi, ed., ‘A Companion to Theoretical Econometrics’, *Companions to Contemporary Economics*, Blackwell, Malden (MA), chapter 32, pp. 678–699.
- Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer, Berlin, Heidelberg, New York.
- Patacchini, E. and Zenou, Y. (2007), ‘Spatial Dependence in Local Unemployment Rates’, *Journal of Economic Geography* **7**, 169–191.
- Pesaran, M., Schuermann, T. and Weiner, S. (2004), ‘Modeling Regional Interdependencies using a Global Error-Correcting Macroeconometric Model’, *Journal of Business and Economics Statistics* **22**, 129–162.
- Pesaran, M. and Tosetti, E. (2007), Large Panels with Common Factors and Spatial Correlations, Discussion Paper 3032, IZA.
- Schanne, N., Wapler, R. and Weyh, A. (2009), ‘Regional Unemployment Forecasts with Spatial Interdependencies’, *International Journal of Forecasting* **forthcoming**.
- Stock, J. (2001), Forecasting Economic Time Series, *in* B. Baltagi, ed., ‘A Companion to Theoretical Econometrics’, *Companions to Contemporary Economics*, Blackwell, Malden (MA), chapter 27, pp. 562–584.