A simultaneous equations model with interaction effects in y, x and ε, and a spatial reformulation of Okun's Law

Paper prepared for the III World Conference of Spatial Econometrics, Barcelona

J. Paul Elhorst

Faculty of Economics and Business, University of Groningen, P.O. Box 800,

9700 AV Groningen, The Netherlands, Phone: +31 50 363 3893, Fax: +31 50 363 7337,

j.p.elhorst@rug.nl

Annette Illy

University of Groningen, annette.illy@gmx.de

Abstract

We set forth a simultaneous equations model with one equation for every spatial unit in the sample. Each equation in this model is further extended to include endogenous interaction effects, exogenous interaction effects and interaction effects among the error terms. In contrast to the single equation model found in Manski (1993), we show that the interaction parameters of this model are identified under reasonable conditions.

Using this framework, we analyze Okun's law for 112 Western European regions over the period 1986-2001. This is the first time that Okun's law is estimated while including endogenous and exogenous interaction effects. Studies that have estimated Okun's law, including correlated error terms, already exist. The inclusion of endogenous and exogenous interaction effects turns out to have two main effects on Okun's coefficient. It decreases and breaks up almost equally into a local effect and a neighbor effect.

Keywords: Interaction effects, Identification, Spatial heterogeneity, Okun's law

***JEL* Codes**: C21, C31, E32, R11

1. Introduction

In his seminal book, Anselin (1988) points out that two problems may arise when data incorporate a locational component. The first problem originates in the fact that parameters are not homogeneous over space, but instead vary over different geographical locations. Anselin (1988, Ch. 9), but also Pesaran and Smith (1995) and Fotheringham et al. (2002), advocate abandoning the fundamental assumption of homogeneous parameters underlying pooled models. The main problem with these models is that they only capture representative behavior and do not show the differences in behavior found among individual spatial units. One reason for introducing relationships that can exhibit spatial variation is that the pooled model is a gross misspecification of reality, because one or more relevant variables are either erroneously omitted from the model or are represented by an incorrect functional form.

There are two methods available to account for spatial heterogeneity. Both assume a system of equations with one equation for every spatial unit in the sample and different response parameters from one spatial unit to the other. The first method requires space-time data because it selects a sub-sample of observations on each spatial unit over time in order to estimate the parameters of the explanatory variables in the regression equation of that spatial unit. The two best-known models are the fixed coefficients and the random coefficients models (Elhorst, 2003a). The second method only requires cross-sectional data because it selects a sub-sample of spatial units around each spatial unit in the sample in order to estimate the parameters of the explanatory variables in the regression equation of that spatial unit. This second method is known as geographically weighted regression (GWR) (Fotheringham et al., 2002).

The second problem that may arise when data incorporate a locational component is that there may be spatial dependence among the observations at each point in time. Manski (1993) indicates that three different types of interaction effects may explain why an observation associated with a specific location may be dependent on observations at other locations. These effects are as follows: (i) endogenous interaction effects, where the decision of a spatial unit (or its subjects) to behave in some way depends on the decision taken by other spatial units; (ii) exogenous interaction effects, where the decision of a spatial unit to behave in some way depends on independent explanatory variables of the decision taken by other spatial units – if the number of independent explanatory variables in a linear regression model is K, then the number of exogenous interaction effects is also K, provided that the intercept is considered as a separate variable; (iii) correlated effects, where similar unobserved environmental characteristics result in similar behavior. Manski (1993) concurrently points out that the interaction parameters are not identified if the dependent variable describing the behavior of a spatial unit is regressed on the independent explanatory variables and on these $2+K$ interaction effects. For identification, at least one of these $2+K$ interaction effects must be excluded from the single equation model.

To model spatial dependence between the observations, Anselin (1988) introduces two single equation models that have been the main focus of the spatial econometrics literature for a long time: the spatial error model and the spatial lag model. The first model incorporates a spatial autoregressive process in the error term, while the second model contains a spatially lagged dependent variable. In other words, both models contain only one interaction effect. Anselin et al. (1996) also develop a procedure to test which of these two models best describes the data.

In his keynote speech at the first World Conference of the Spatial Econometrics Association in 2007, Harry Kelejian advocated single equation models that include both a spatially lagged dependent variable and a spatially autocorrelated error term (based on Kelejian and Prucha, 1998 and related work), while James LeSage, in his presidential address at the 54th North American Meeting of the Regional Science Association International in 2007, advocated single equation models that include both a spatially lagged dependent variable and spatially lagged explanatory variables (based on LeSage and Pace, 2009). In other words, they both advocated models that contain more than just one interaction effect as in the spatial lag model or in the spatial error model. The reason they did not advocate models that contain all possible interaction effects is probably due to the identification problem advanced by Manski (1993). In this paper we go one step further and develop an econometric model that includes all possible interaction effects. This model is not a single equation model as in Manski (1993), but, following Anselin (1988, Ch. 9), Fotheringham et al. (2002) and Elhorst (2003a), is a simultaneous equations model with one equation for every spatial unit in the sample. Each equation within this model is extended to include endogenous interaction effects, exogenous interaction effects and an error term that is correlated with the error terms of the other spatial units. We also derive the conditions needed so that the interaction parameters of this simultaneous equations model can be identified. Importantly, these conditions will not require that the spatial weights matrix $\boldsymbol{W}$, the matrix used to describe the spatial arrangement of the spatial units in the sample, be different for different types of interaction effects. Anselin and Bera (1998) show that the use of the same spatial weights matrix may also pose identification problems when a single equation model with both a spatially lagged dependent variable and a spatially autocorrelated error term is estimated by maximum likelihood.

This paper is organized as follows. In Section 2, we present our simultaneous equations model, determine the minimum number of observations required for estimation and discuss the conditions required for identification when using space-time data. Next, we briefly discuss the possibilities as well as obstacles to the application of this model when using cross-sectional data. The method used to estimate the model will be 3SLS. As an application of this model, we analyze Okun's law using data of 112 Western European regions over the period 1986-2001 in Section 3. This will be the first time that Okun's law is estimated while including endogenous and exogenous interaction effects. Studies that have estimated Okun's law, with correlated error terms included, already exist. The inclusion of interaction effects appears to have two main effects on Okun's coefficient. It decreases and breaks up almost equally into a local effect and a neighbor effect. Section 4 summarizes our main findings and concludes the paper.

2. A simultaneous equations model with spatial interactions in y, x and ε

2.1 Specification

To explain the identification problem, we will first consider a simultaneous equations model without imposing any prior information about the nature of interactions over space. This model consists of a system of equations with one equation for every spatial unit in the sample and different response parameters from one spatial unit to the other. The i^{th} equation of this model takes the form

$$y_{it} = \sum_{j=1, j\neq i}^{N} \delta_{ij} y_{jt} + \alpha_i + \sum_{k=1}^{K} \beta_{ik} x_{it}^{k} + \sum_{k=1}^{K} \sum_{j=1, j\neq i}^{N} \gamma_{ijk} x_{jt}^{k} + \varepsilon_{it}\,, \tag{1}$$

where i is an index for the spatial units in the cross-sectional dimension, with i=1, ...,N, and t is an index for the time dimension, with t=1, ...,T. y_{it} is an observation on the dependent variable of spatial unit i at t and y_{jt} of spatial unit j at t. x_{it}^{k} is an observation on the k^{th} independent variable of spatial unit i at t and x_{jt}^{k} of spatial unit j at t. The total number of independent variables (besides the intercept) is K. In this model the dependent variable of spatial unit i at t, y_{it}, is taken to depend on an intercept, α_i, the independent variables of spatial unit i at t, x_{it}^{k} (k=1,…,K), and on the dependent and independent variables of all the other spatial units at t, y_{jt} (j=1,..,N, except for j=i) and x_{jt}^{k} (k=1,…,K; j=1,…,N, except for j=i). The intercept α_i and the slope parameters β_{ik} (k=1,…,K) and γ_{ijk} (j=1,…,N; k=1,…,K) are fixed but unknown parameters for i=1,…,N. In addition, it is assumed that $E(\varepsilon_{it})=0$ and that the error terms in different equations are correlated with each other at t, a phenomenon that is known as contemporaneous error correlation

$$E(\varepsilon_{it}\varepsilon_{jt}) = E(\varepsilon_{jt}\varepsilon_{it}) = \sigma_{ij} = \sigma_{ji}\,, \quad i,j=1,..,N;\ t=1,\ldots,T. \tag{2}$$

Since $\sigma_{ii} \neq \sigma_{jj}$ if $i \neq j$, the error terms may also be said to be heteroskedastic across the equations.

The parameters δ in this simultaneous equations model measure the endogenous interaction effects between the spatial units, the parameters γ the exogenous interaction effects between the spatial units, and the parameters σ the interaction effects between the variables omitted from the model through the error terms. So far these parameters can take any value, since restrictions upon or between them have not been imposed.

2.2 Required number of observations and identification conditions

We hasten to point out that the parameters of this unrestricted model cannot be estimated because the number of observations may well be insufficient and the unknown parameters not be identified.

The total number of unknown parameters to be estimated in the full model is comprised of: $N(N$-1) different parameters δ, N different intercepts α, NK different parameters β, $N(N$-1)K different parameters γ, and ½$N(N$+1) different parameters σ of the symmetric covariance matrix of the error terms. This means that the number of unknown parameters in each equation amounts to N(1+K)+½(N+1). Since the number of observations available for the estimation of these unknown parameters is T, this model is only of use when T is large and N is small. Unfortunately, most space-time data sets do not meet this requirement, even if N is small. Note that sampling more observations in the cross-sectional domain is no solution for insufficient observations in the time domain, since the number of unknown parameters increases as N increases, a situation known as the incidental parameters problem.

In addition, if T were to be sufficiently large, meaningful estimates of the structural parameters could still not be obtained. This is because every equation eventually contains the same set of variables either at the left-hand or at the right-hand side of the equals signs. This problem is known as the identification problem and is not a question of method of estimation, nor of sample size, but of whether the equations can be distinguished from each other. The explanation of this problem below is based on Hsiao (1983) and Johnston and Dinardo (1997, pp. 305-314).

To derive identification conditions, the simultaneous equations model at each point in time (t=1,…,T) is rewritten as

$$\begin{bmatrix} 1 & -\delta_{12} & . & -\delta_{1N} \\ -\delta_{21} & 1 & . & -\delta_{2N} \\ . & . & . & . \\ -\delta_{N1} & -\delta_{N2} & . & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ , \\ y_{Nt} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_{11} & . & \beta_{1K} & \gamma_{121} & . & \gamma_{12K} & . & \gamma_{1N1} & . & \gamma_{1NK} \\ \alpha_2 & \gamma_{211} & . & \gamma_{21K} & \beta_{21} & . & \beta_{2K} & . & \gamma_{2N1} & . & \gamma_{2NK} \\ . & . & . & . & . & . & . & . & . & . & . \\ \alpha_N & \gamma_{N11} & . & \gamma_{N1K} & \gamma_{N21} & . & \gamma_{N2K} & . & \beta_{N1} & . & \beta_{NK} \end{bmatrix} \begin{bmatrix} 1 \\ x_{1t}^1 \\ . \\ x_{1t}^K \\ . \\ x_{Nt}^1 \\ . \\ x_{Nt}^K \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ . \\ \varepsilon_{Nt} \end{bmatrix}, \quad (3)$$

or, equivalently,

$$\boldsymbol{BY}_t = \boldsymbol{CX}_t + \varepsilon_t . \quad (4)$$

Let θ_i denote the N(1+K) unknown parameters α, β, γ and δ of model equation i, and K* the total number of independent explanatory variables in that equation, including the intercept, K*=1+NK. Also let

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{C} \\ \boldsymbol{I}_{K*} \end{bmatrix}, \quad (5)$$

and Φ be a (N+K*)×R matrix with R the number of restrictions imposed on θ_i. Then identification of θ_i requires that the rank of $[\boldsymbol{P} \quad \Phi]$ be N+K*-1. Unfortunately, implementation of this rank condition is not usually feasible for systems other than for those for which N is small. However, one necessary condition for identification, known as the order condition, is easy to derive and to apply. The rank condition cannot hold if $[\boldsymbol{P} \quad \Phi]$ does not have at least N+K*-1 columns, that is, if $K^*+R \geq N+K^*-1$. This implies that $R \geq N-1$. To state it in words, the number of a priori restrictions on the coefficients in each equation should be at least as great as N-1.

2.3 Parameter restrictions to meet the identification conditions

Since we have not imposed any restrictions on the coefficients α, β, γ and δ in the model equations yet, the interaction coefficients in the simultaneous equations model cannot be estimated in a meaningful way, even if T were to be sufficiently large. The standard solution to this problem in spatial econometrics is to impose prior information about the nature of interactions over space and to assume one instead of N-1 parameters for each type of interaction effect (Anselin, 2006).

Prior information about the nature of interactions over space can be modeled using a spatial weights matrix ***W***, a pre-specified nonnegative matrix of order N describing the arrangement of the spatial units in the sample. The diagonal elements of ***W*** are set to zero by convention, since no spatial unit can be viewed as its own neighbor. For ease of interpretation, it is also common practice to normalize ***W*** such that the elements of each row sum to unity. Since ***W*** is nonnegative, this ensures that all the elements are between 0 and 1. This has the effect that the interaction effects of spatial unit i with the dependent variable, the independent variables and the error terms — $\sum_j w_{ij} y_{jt}$, $\sum_j w_{ij} x^k_{jt}$ (k=1,…,K) and $\sum_j w_{ij} \varepsilon_{jt}$, where w_{ij} is the (i,j)th element of ***W*** — can be interpreted as averaging neighboring values.

Given ***W***, the following restrictions can be imposed on the parameters so as to reduce the number of parameters for each type of interaction effect from N-1 to 1

$$\delta_{ij} = \delta_i w_{ij} \text{ for } i \neq j \text{ and } i,j=1,\ldots,N, \tag{6a}$$

$$\gamma_{ijk} = \gamma_{ik} w_{ij} \text{ for } i \neq j,\ i,j=1,\ldots,N, \text{ and } k=1,\ldots,K, \tag{6b}$$

$$\sigma_{ij} = \sigma_i w_{ij} \text{ for } i \neq j \text{ and } i,j=1,\ldots,N. \tag{6c}$$

Consequently, the i^{th} equation of the simultaneous equations model reduces to

$$y_{it} = \sum_{j=1, j\neq i}^{N} \delta_i w_{ij} y_{jt} + \alpha_i + \sum_{k=1}^{K} \beta_i x_{it}^k + \sum_{k=1}^{K} \sum_{j=1}^{N} \gamma_{ik} w_{ij} x_{jt}^k + \varepsilon_{it}, \tag{7}$$

while the covariance between the error terms reduces to

$$E(\varepsilon_{it}\varepsilon_{it}) = \sigma_{ii} \text{ for } i=j,\ i=1,\ldots,N \text{ and } t=1,\ldots,T. \tag{8a}$$

$$E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_i w_{ij} \text{ for } i\neq j,\ i,j=1,\ldots,N \text{ and } t=1,\ldots,T. \tag{8b}$$

The total number of unknown parameters to be estimated in this restricted model is comprised of: N different parameters δ, N different intercepts α, NK different parameters β, NK different parameters γ, and $2N$ different parameters σ of the – diagonal and non-diagonal elements of the – covariance matrix of the error terms. This means that the number of unknown parameters in each equation is reduced to 4+2K. Consequently, the minimum number of observations on each spatial unit required for estimation is also reduced to 4+2K, a number that is no longer dependent on N.

The number of restrictions on the coefficients α, β, γ and δ imposed in (6) amounts to $N(K+1)(N-2)$. Note that $\delta_{ij} = \delta_i w_{ij}$ and $\gamma_{ijk} = \gamma_{ik} w_{ij}$ for every pair of i and j counts as one restriction, except for the first one. Consequently, the number of restrictions per equation amounts to $(K+1)(N-2)$, which is more than the N-1 restrictions that are required to satisfy the order condition, provided that the number of spatial units in the cross-sectional domain is at least 3 and that the number of independent explanatory variables (besides the intercept) is at least 1: $N\geq 3$ and $K\geq 1$.

Instead of the parameter restrictions $\delta_{ij} = \delta_i w_{ij}$ and $\gamma_{ijk} = \gamma_{ik} w_{ij}$, one might also constrain the number of neighbors that each particular spatial unit has. If spatial unit j is assumed not to neighbor spatial unit i, then w_{ij} is equal to zero, which in turn implies that $\delta_{ij}=0$ and $\gamma_{ijk}=0$, and that the corresponding explanatory variables y_{jt} and x_{jt}^k are excluded from the i^{th} equation. These types of restrictions are known as exclusion restrictions because they limit the number of explanatory variables in the equation. If the parameter restrictions $\delta_{ij} = \delta_i w_{ij}$ and $\gamma_{ijk} = \gamma_{ik} w_{ij}$ for those spatial units that are neighbors of each other (w_{ij}>0) are maintained, the number of exclusion restrictions is exactly the same as the number of restrictions implied by (6a) and (6b) that are lost. This implies that the parameters are again identified if $N \geq 3$ and $K \geq 1$ and can again be estimated if $T \geq 4+2K$.

In sum, if $N \geq 3$, $K \geq 1$ and $T \geq 4+2K$ the parameters of the restricted model are identified and can be estimated.

2.4 Alternative restrictions and cross-sectional data

The parameters δ_{ij} and γ_{ijk} in the i^{th} equation can also be freely estimated for those spatial units that are neighbors of spatial unit i (w_{ij}>0) without imposing the parameter restrictions $\delta_{ij} = \delta_i w_{ij}$ and $\gamma_{ijk} = \gamma_{ik} w_{ij}$. Suppose that spatial unit i has q neighbors, with $q<N$, and thus N-q-1 spatial units that are not neighbors. For every spatial unit j that is not a neighbor of i, we then have $\delta_{ij} = 0$ and $\gamma_{ijk} = 0$ (k=1,…,K). The total number of parameter restrictions within the i^{th} equation is therefore (N-q-1)(1+K), which should be greater than or equal to N-1. This implies that $K \geq 1$ and that $N \geq q+1+q/K$. Both conditions, however, do not guarantee that the parameters are identified. If two spatial units i and j are neighbors of each other and the remaining q-1 neighboring spatial units of i and j fully overlap, then the estimation equations of these two spatial units again contain the same set of variables either at the left-hand or at the right-hand side of the equals signs, as a result of which the response parameters of these equations are not identified. Only if each equation has at least one unique independent explanatory variable, is the entire model identified (Greene, 2003, p. 393). This can be easily verified by constructing a binary matrix $\boldsymbol{W}^*$ of order N whose elements are 1 if i=j or if w_{ij}>0 for i,j=1,…,N and $i \neq j$, and zero otherwise. If this matrix $\boldsymbol{W}^*$ has full rank, rank($\boldsymbol{W}^*$)=N, then the condition that each equation should have at least one unique independent explanatory variable is satisfied.

To determine the minimum number of observations on each spatial unit required for estimation under these circumstances, we also need to take into account the restrictions on the covariance matrix, $\sigma_{ij} = 0$ when $w_{ij} = 0$. Since it is possible that *j* may be considered as being a neighbor of *i* for *i*,*j*=1,…,*N*, but not vice versa, we have ½(*N*-1)-*q* restrictions, which implies that $T \geq 2+K+qK+2q$.

When we have cross-sectional instead of space-time data and each equation is estimated by GWR, there are three complications. First, the number of observations *T*, available for estimation, may vary from one equation to the other, but will never be greater than *N*. Second, whereas in the space-time model every observation is sub-sampled only once, in GWR observations are sub-sampled several times. If each sub-sample around a particular spatial unit consists of *s* spatial units, with $s \leq N$, and *N* sub-samples are formed, then each spatial unit is sub-sampled $Ns/N=s$ times, on average. This implies that the condition rank($\boldsymbol{W}^*$)=*N* must also be satisfied if we impose the parameter restrictions $\delta_{ij} = \delta_i w_{ij}$ and $\gamma_{ijk} = \gamma_{ik} w_{ij}$. This is necessary in order to exclude the possibility that the regressions of two spatial units *i* and *j* are estimated on the same set of observations — this happens if two spatial units are neighbors of each other and the remaining neighboring spatial units of *i* and of *j* fully overlap — as a result of which the response parameters of the two regressions are not identified. Third, the covariance matrix of the error terms needs adjustment. Consider the error terms of two spatial units *i* and *j*, ε_i and ε_j. Then we not only have $E(\varepsilon_{it}\varepsilon_{jt}) = E(\varepsilon_{jt}\varepsilon_{it}) = \sigma_{ij} = \sigma_{ji}$, if *i* and *j* are part of different sub-samples, but also if *i* and *j* are part of the same sub-sample. This adjustment of the covariance matrix is the main obstacle to estimating the model using GWR.

2.5 Estimation

We will use 3SLS to estimate the simultaneous equations models. 3SLS has the advantage that it does not require information about the exact distribution of the error terms.[1] The intercept and the independent explanatory variables x_{it}^{k} for i=1,…,N and k=1,…,K are used as instruments for the endogenous explanatory variable $\sum_j w_{ij} y_{jt}$ in each equation, that is, 1+KN instruments in total. Spatially lagged independent variables are not used as instruments, since they form linear combinations of the variables x_{it}^{k} (i=1,…,N; k=1,…,K) and therefore would lead to perfect multicollinearity in the matrix of instruments.[2] Lee (2004) has pointed out that the method of instrumental variables will break down when there are no regressors in the model. However, since the parameters of the model will only be identified when K≥1, this problem will not occur.

A practical problem is that the value of N, which determines the number of equations in the simultaneous equations model, is restricted in some commercial econometrics software. For instance, the upper bound of N in LIMDEP (version 8.0) is 20 for the simultaneous equations model. Another practical problem is that every single restriction has to be specified. Since we have $N[(N-2)(K+1)+½(N-3)]$ restrictions in the system, this is quite a burden if N is large. Furthermore, it is difficult if not impossible to impose restrictions on the parameters of the covariance matrix in commercial software packages. For these reasons and for reasons of efficiency — the restrictions proposed in this paper can be easily implemented using a few

---

[1]Kelejian and Prucha (2004) have developed a generalized moments (GM) estimator of a simultaneous equations model with a spatially lagged dependent variable and a spatially autocorrelated error term in each equation, but in order to avoid confusion we would like to point out that their model is different from ours. The data set they take is a single cross-section, while we start with space-time data (or different sets of cross-sectional data according to GWR). They consider one equation for every dependent variable in the model, while we have only one dependent variable and consider one equation for every spatial unit in the sample. Consequently, every equation in their model is estimated on the same cross-section of observations, while the equations in our model are based on different sets of observations. Finally, they assume that the error terms in each equation follow a first-order spatial autoregressive process, while we assume that the error terms are contemporaneously correlated.

[2] This also (see previous footnote) distinguishes our 3SLS estimator from the GM estimator in Kelejian and Prucha (2004).

loops in i and j — we have programmed the 3SLS estimator in Matlab. This routine can be downloaded for free from the first author's Web site.

3. An application of Okun's Law in Western European Regions

3.1 Introduction

Okun's law, named after economist Arthur Okun (1962), is one of the basic rules of thumb of macroeconomics. It describes a relationship between the change in the rate of unemployment and the difference between actual and potential real GDP. In the United States, Okun's law can be stated as saying that for every one percentage point by which the actual unemployment rate exceeds the so-called "natural" rate of unemployment, real GDP is reduced by 2% to 3%. According to Perman and Tavera (2007), Okun's law can be regarded as a benchmark for policy-makers when measuring the cost of higher unemployment. There are three reasons why GDP may increase or decrease more rapidly than unemployment decreases or increases. As unemployment increases, (i) unemployed persons may drop out of the labor force, after which they are no longer counted in unemployment statistics, (ii) employed persons may work shorter hours, and (iii) labor productivity may decrease, perhaps because employers retain more workers than they need.

Since Okun's original publication, the existence of a trade-off between unemployment and output has been studied extensively. Studies that have appeared in the last decade can be divided into three groups:

a. Studies that use data from one country. Prachowny (1993), Weber (1995), Weber and West (1996), Moosa (1999), Cuaresma (2003), Silvapulle et al. (2004), and Huang and Lin (2006, 2008) used U.S. data, Attfield and Silverstone (1997, 1998) U.K. data, and Sögner (2001) Austrian data;

b. Studies that use data from more than one country, among which Moosa (1997), Virén (1999), Lee (2000), Freeman (2001), Sögner and Stiassny (2002), Izyumov and Vahaly (2002) and Perman and Tavera (2005, 2007);
c. Studies that use regional instead of country data, among which Freeman (2000), Apergis and Rezitis (2003), Chrisopoulos (2004), Adanu (2005), Kosfeld and Dreger (2006) and Villaverde and Maza (2007).

The main reason, according to the latter two groups of studies, for researching data from more than one country or of using regional instead of country data is to test for spatial differences in the responsiveness that output has to reductions in unemployment. Differences among countries or regions point to institutional differences that determine the rigidity or flexibility of national or regional labor markets (Moosa, 1997).

The commonly used empirical specification takes the form

$$(y_{it} - y_{it}^{*}) = \beta(u_{it} - u_{it}^{*}) + \varepsilon_{it}, \quad i=1,\dots,N;\ t=1,\dots,T, \tag{9}$$

where y is the logarithm of actual output and y^* of potential or equilibrium output (both measured as real gross domestic product), and u is the actual rate of unemployment and u^* the natural or the equilibrium rate of unemployment. The difference between the actual and the potential value of a variable is called a gap. The index i refers to a country or region and the index t to a time period. Okun's coefficient is represented by -β, the increase in the output gap (in percentage points) for every 1% decrease in the unemployment gap.[3] It is common practice to estimate Okun's law for each country or region in the sample separately from the others,

[3]Using a production function approach, Prachowny (1993) has argued, that the model in (9) should be extended to include variables measuring the difference between the actual and the potential utilization rate of capital (c-c^*), the difference between the actual and the potential supply of workers (l-l^*) and the difference between the actual and the potential number of working hours (h-h^*). However, this study has not only been criticized (Attfield and Silverstone, 1997), it has also failed to gain a firm foothold in later studies. It has been criticized because the levels and especially the potential values of c, l and h are extremely difficult to measure (see below), and because the labor supply gap and the capacity utilization gap are highly correlated. The overall conclusion from most studies is that equation (9) is still correctly specified when it is assumed that all other variables are either on their equilibrium paths or change pari passu with unemployed labor (see Freeman, 2001; Christopoulos, 2004).

which can be justified by the assumption that the error terms ε_{it} are independently and identically distributed for all i and t with zero mean and variance σ^2.[4]

Originally, Okun (1962) regressed the unemployment gap on the output gap and then used the reciprocal of the slope of this regression to predict the effect of unemployment on output. However, Barreto and Howland (1993) point out that this leads to an overestimation of Okun's coefficient. This is because the reciprocal of the slope of this regression not only measures the impact of unemployment on output, but also the expected value of the error term given the unemployment rate, which are not independent of each other, $E(\varepsilon|u\text{-}u^*)\neq 0$.[5] Barreto and Howland (1993) find that, had Okun adopted model (9), the slope would fall to 2, instead of the 3.2 reported by Okun (1962). Nevertheless, the literature can be divided into two camps, those who take the output gap and those who take the unemployment gap as the dependent variable. In this paper, we take equation (9) as our point of departure.[6]

Okun's law is difficult to use in practice because y^* and u^* can only be estimated, not observed. Modern empirical work offers a number of alternatives to the separation of trends and cycles in economic time series and thus for the derivation of these two variables, such as linear or quadratic trends, first differencing, or more complex methods such as the Beveridge-Nelson method, the Harvey structural time series approach, the Baxter-King bandpass filter and the Hodrick-Prescott filter. In this paper, we have adopted the Hodrick and Prescott (1997) filter (hereafter called the HP filter).[7]

[4] Note that not one single study discussed above treated the unemployment gap as an endogenous variable.
[5] For this reason, the coefficient of regressing y on x, where y and x are two variables, is generally not identical to the reciprocal of the coefficient of regressing x on y.
[6] The question of whether the unemployment gap affects the output gap or vice versa might be investigated using Granger causality tests, but this is not common practice within this literature.
[7]The HP filter has become a standard method in the business cycle literature for removing trend movements. Although the HP filter may be subject to criticism (see Freeman, 2001; Sögner, 2001; Silvapulle et al., 2004), it remains one of the standard methods for detrending. Freeman (2000), Cuaresma (2003), Adanu (2005), Villaverde and Maza (2007), and Apergis and Rezitis (2003) applied different detrending or filtering techniques to investigate to what extent they affected the estimate of Okun's coefficient. Freeman (2000) compared the performance of the Baxter and King bandpass filter with that of a quadratic trend, Cuaresma (2003) the Hodrick-Prescott filter with that of the Harvey structural time series approach, Adanu (2002) and Villaverde and Maza (2007) the Hodrick-Prescott filter with that of a quadratic trend, and Apergis and Rezitis (2004) the Hodrick-Prescott filter with that

The HP filter is a smoothing method for obtaining an estimate of the long-term trend component of a time series. It decomposes an economic time series (s_t) into an unobserved long-run growth component (s_t^*) and an irregular, or cyclical, component (s_t^c). The smoothed series s_t^* is computed by minimizing the variance of s_t around s_t^*, subject to a penalty for the variation in the second difference of the growth component s_t^*

$$\min_{s_t^*} \frac{1}{T}\sum_{t=1}^{T}(s_t - s_t^*)^2 + \frac{\lambda}{T}\sum_{t=2}^{T-1}((s_{t+1}^* - s_t^*) - (s_t^* - s_{t-1}^*))^2 . \tag{10}$$

The parameter λ is a positive number which penalizes variability in the growth component series. The larger the λ the smoother the series; as λ approaches infinity, the growth component corresponds to a linear time trend. For annual time series, Hodrick and Prescott (1997) suggest λ=100. The variable $s_t - s_t^*$ applied to the output and unemployment series produces the output and unemployment gap variables.

There are two additional issues that have received a lot of attention in the empirical literature. The first issue is whether the variables, the output gap and the unemployment gap, are stationary. Stationary variables are a prerequisite for regression analysis (Greene, 2003). In the next section we will apply two unit root tests to find out whether the output and unemployment gap variables are indeed stationary.[8]

The second issue is whether or not the error terms are serially autocorrelated. Studies that reject the hypothesis of no serial autocorrelation often consider a dynamic version of Okun's law, which takes the form

---

of the Baxter and King bandpass filter. All these studies conclude that the results are robust as to the choice of technique, so long as the same technique is used for both the output and the unemployment variables.

[8] Engle and Granger (1987) have argued that, if both variables appear to be stationary only after differencing, then these variables might still be cointegrated if the series ε_t is integrated to the order of zero. If cointegration exists, then equation (9) still represents a structural relationship that is not spurious. However, an error-correction framework and more sophisticated estimation techniques are required under these circumstances (see, for details, Attfield and Silverstone, 1997, 1998; Lee, 2000; Christopoulos, 2004).

$$(y_{it} - y_{it}^*) = \sum_{p=1}^{P} \rho_l (y_{it-l} - y_{it-l}^*) + \beta(u_{it-l} - u_{it-l}^*) + \varepsilon_{it} \,, \tag{11}$$

where β denotes the short-run effect and $\beta/(1-\rho_1-...-\rho_P)$ the long-run effect of unemployment on output (see Weber, 1995). One regional study that tested for and found evidence in favor of serial autocorrelation was that of Adanu (2005). He started with a maximum of five lags and then used a sequential testing approach to find out whether the number of lags could be reduced. Country-based studies that have tested for and found evidence in favor of serial autocorrelation are Moosa (1999), Cuaresma (2003), Perman and Tavera (2006) and Huang and Lin (2008). Most of these studies used quarterly data and found evidence in favor of two lags in the dependent variable. Studies that have tested for and rejected serial autocorrelation are those of Freeman (2000) and Apergis and Rezitis (2003), who used regional data, and Moosa (1997) and Freeman (2001), who used country data. Most of these studies used annual data. In the next section we will also test for serial autocorrelation and extend the model if necessary.

3.2 Data and empirical results for the time series approach

In this section we estimate Okun's law using data from 112 regions across eight countries in Western Europe (number of regions of each country given in parenthesis): Belgium (11), West Germany (30), Denmark (1), Spain (16), France (21), Italy (20), Luxembourg (1), and the Netherlands (12). The basic data have been taken from Eurostat's regional data file for the period 1986-2001, using the regional division of Eurostat at the NUTS 2 level. The unemployment rate is constructed by dividing the number of unemployed people by the number of people in the labor force, either employed, self-employed or unemployed. Unemployment data are often troubled by variations between countries and over time in terms of the ways unemployment rates are defined and/or measured. We have partly avoided such problems by using Eurostat's harmonized unemployment rates. Output is represented by real GDP at market prices. The deflators used to obtain real data are different for different countries, but the same for all regions within a country.

The output gap and the unemployment gap variables have been estimated using the Hodrick-Prescott filter with $\lambda=100$, as explained in the previous section. To test whether these gap variables are stationary, we performed the individual cross-sectionally augmented Dickey-Fuller (CADF) test developed by Pesaran (2007).[9] The null hypothesis of a unit root has been rejected for 100 of the 112 time series of the output gap variable and for all time series of the unemployment gap variable at five-percent significance and for 105 of the 112 time series of the output gap variable at ten-percent significance. However, given the short time-span of our time series — each time series consists of only 16 observations — the results may suffer as a result of low power. Therefore, we also performed the CADF panel data unit root test of Pesaran (2007). This test statistic is based on the average of the individual CADF tests. We found -14.15 for the output gap variable and -16.25 for the unemployment gap variable, which represents a rejection of a unit root in both variables at one-percent significance (the critical value according to Pesaran's Table II(a) is approximately -1.66). The conclusion must be that the variables are stationary and that Okun's law may be estimated for each region separately according to equation (9) and according to its spatial extension in equation (7), unless we find strong empirical evidence in favor of serial autocorrelation.

Table 1 reports the estimation results when ignoring spatial interaction effects, which has been the commonly used method for estimating Okun's law up until now. Just as in most previous studies, we started by estimating one equation for every spatial unit in the sample using OLS. Note that we did not include a constant, since the sum of the observations of any of the time series obtained by using the HP filter equals zero. Consequently, the constant is zero

[9] The test statistic is the t-value of the lagged dependent variable in a standard augmented Dickey-Fuller regression augmented with the cross-section averages of lagged levels and first-differences of the individual series. These additional variables are important since we will extend the model with spatial (which are cross-sectional) interaction effects in the next section.

by construction. The value of Okun's coefficient for the 112 regions varies between 0.17 and 4.17. Table 1 reports the unweighted average of these coefficients and its t-value, as proposed by Pesaran and Smith (1995). This mean is obtained by $\overline{\beta} = \frac{1}{N}\sum_i \beta_i$ (ignoring the minus sign) and its t-value by $\overline{\beta}/\sqrt{\mathrm{var}(\overline{\beta})}$. Note that $\mathrm{var}(\overline{\beta}) = \mathrm{var}(\frac{1}{N}\sum_i \beta_i) = \frac{1}{N^2}\sum_i \mathrm{var}(\beta_i)$, so long as it is assumed that interaction effects among the error terms of different spatial units do not exist. When testing these individual regressions for heteroskedasticity, we found that the hypothesis of homoskedastic error terms within the equations cannot be rejected for 109 regions at five-percent significance and cannot be rejected for any region at one-percent significance. By contrast, when testing for heteroskedasticity across the model equations, we found that the hypothesis $H_0 : \sigma_{11}^2 = \ldots = \sigma_{NN}^2 = \sigma^2$ had to be strongly rejected. In other words, heteroskedasticity across the equations is a serious problem that should be taken into account, whereas heteroskedasticity within the equations is not.

<< Table 1 around here >>

When testing these individual regressions for serial autocorrelation, we found that the hypothesis of no serial autocorrelation had to be rejected for 19 regions. We therefore extended the equations of these regions with lags in the dependent variable. We used Akaike's Information Criterion to determine the optimal lag length. The optimal lag length appeared to be 1 for seven regions, 2 for six regions, 3 for one region and 4 for five regions (4 is the maximum being considered). Okun's coefficient for these regions has been reestimated using (11). The results of this so-called dynamic OLS estimator, combined with the OLS estimates of the remaining 93 regions, are reported in Table 1. The conclusion must be that the mean, its t-value, and the range of Okun's coefficient of the OLS estimator and this dynamic OLS estimator are virtually the same. The average value of Okun's coefficient is 1.95 when the model is estimated by OLS and 1.90 when estimated by dynamic OLS.

One important question is how this Okun's coefficient of 1.90 relates to that of other studies. The average value of Okun's coefficient in studies based on U.S. data that took the output gap variable as the dependent variable is 1.74 (Prachowny, 1993; Attfield and Silverstone, 1997, 1998; Freeman, 2001; Lee, 2000) and the average value of Okun's coefficient in studies that took the unemployment gap variable as the dependent variable is 2.85 (Moosa, 1997, 1999; Sögner and Stiassny, 2001; Cuaresma, 2003; Silvapulle et al., 2004; Huang and Lin, 2008). This difference has been explained by Barreto and Howland (1993), as discussed above. The reciprocal of the slope of the regression of (u-u^*) on (y-y^*) overestimates Okun's coefficient, because it also includes the expected value of the error term given the unemployment rate. Using U.S. data, Freeman (2001) found that the value of Okun's coefficient for the overall U.S. economy is 1.99 and for the eight regions he considered varies within the range 1.84 to 3.57, with an average of 2.22. The explanation of why the average response at the regional level is larger than the national response is that regional labor markets are more flexible. The average value of Okun's coefficient in studies based on European data that took the output gap variable as the dependent variable is 1.43 (Freeman, 2000; Lee, 2001). The explanation of why the trade-off between output and unemployment in European countries is weaker than that in the U.S. is that European labor markets are more rigid and the average unemployment level is higher (Izyumov and Vahaly, 2002). Our value of Okun's coefficient of 1.90 fits within this picture. Since we studied the regional response and since regional labor markets tend to be more flexible, it is higher than the national response: 1.90 versus 1.43. At the same time, it is lower that the regional response from U.S. regions, 1.90 versus 2.22, since U.S. labor markets are generally thought to be more flexible than European labor markets.

3.3 Space-time approach with interaction effects

We will now turn to the model where spatial interaction effects in the output gap variable, the unemployment gap variable and the error terms are taken into account. The literature has produced many theoretical explanations for the existence of interaction effects among these variables. Endogenous interaction effects occur when output in one region is affected by output in other regions. If output in one region increases, output in other regions might also increase due to the import of goods and services. Trade can also encourage output growth in neighboring economies through diffusion of knowledge, since it opens up the possibility of cross-border learning-by-doing, as well as investment in research and development (Helpman, 2004, Ch. 5). The hypothesis that the relative location of an economy, that is, the effect of being located closer or further away from other specific economies, is a determinant of economic growth due to diffusion of knowledge has recently been underpinned by economic-theoretical models (López-Bazo et al., 2004; Ertur and Koch, 2007). Moreover, a vast empirical literature exists in support of the hypothesis that growth in one region is a determinant of economic growth in other regions (for an overview see Fingleton, 2003; Abreu et al., 2005; Rey and Janikas, 2005). Finally, investments in infrastructure such as road or high-speed rail infrastructure can make a whole area more attractive and benefit a region that would not be connected in and of itself (Elhorst and Oosterhaven, 2008).

Exogenous interaction effects occur when households change their consumption and labor supply decisions depending on the market conditions in the home region compared to other regions. If unemployment falls and regional income increases, consumers may increase their consumption expenditures not only in their own region but also in neighboring regions. People also do not necessarily live and work in the same region. Any person, whether employed or unemployed, may supply his labor outside his home region when the wage rate in a nearby region is higher due to lower unemployment, and this higher wage rate compensates for the greater time and commuting costs. Thus, consumption spending and commuting are two potential reasons for the existence of an exogenous interaction effect related to the unemployment rate.

An interaction effect in the error term occurs when factors that are omitted from the model equations do not respect regional boundaries. For example, shocks to the exchange rate (Euro) will not only influence output in one region, but in almost all regions in our sample.

Freeman (2001) argued that the error terms for different regions will be correlated in the Euro Zone especially because the economic ties between countries in the Euro Zone are very tight as a result of the common monetary policy pursued by the European Central Bank.

Despite the growing literature pointing to interaction effects, not one single study on Okun's law discussed above has included endogenous and/or exogenous interaction effects, while only four country-based studies assumed that the error terms were contemporaneously correlated, namely those of Moosa (1997), Freeman (2001) and Perman and Tavera (2005, 2007). Their motivation was to account for potential cross-border error correlation, for example, due to common shocks.[10]

To estimate the interaction effects, we took into account the first-order neighbors of each region. If two regions i and j share a common border, we set $w_{ij}=1$, and if they do not we set $w_{ij}=0$. If two regions are located in different countries and share a common border we also assumed that $w_{ij}=1$. The reason for this is that Overman and Puga (2002) have pointed out that both a cross-regional and a transnational dimension should be added to national anti-unemployment policies. Up until now, we used to think about differences in unemployment rates as differences across countries. However, regional unemployment has a strong geographical component that goes across national boundaries (see also Elhorst, 2003b). In addition, we assumed that the Spanish and Italian islands are connected to the mainland, so that each region has at least one neighbor (the average number of neighbors for each region is 4.56). The spatial weights matrix that was obtained as a result is known as a binary contiguity matrix. Finally, this matrix was row-normalized. Alternative specifications of the spatial weights matrix have also been considered. We will come back to this later.

[10] Kosfeld and Dreger (2006) is the only study that corrects for spatial autocorrelation among the error terms using a first-order spatial autoregressive process $\varepsilon_{it}=\Sigma_j \tau_t w_{ij} \varepsilon_{jt}+\mu_{it}$, where the spatial autocorrelation coefficient τ_t has been estimated for every single time period t (t=1,…,T). This is the specification that is commonly used to specify interaction effects for the error term in a single equation model. To cope with spatial error autocorrelation in a simultaneous equations model, we assume that the error terms of the different equations, one for every spatial unit in the sample, are contemporaneously autocorrelated. This approach goes back to White and Hewings (1982).

The row-normalized binary contiguity matrix imposes two types of restrictions on the parameters of the model. First, for those regions that, according to the spatial weights matrix, do not neighbor each other (w_{ij}=0, i,j=1,…,N; $i \neq j$), we have δ_{ij}=0, γ_{ij}=0 (we may drop the superscript k from this parameter, since we only have one explanatory variable (K=1), which is the unemployment gap variable), and σ_{ij}=0. Second, for those regions that, according to the spatial weights matrix, do neighbor each other (w_{ij}>0, i,j=1,…,N; $i \neq j$), we have $\delta_{ij} = \delta_i w_{ij}$, $\gamma_{ij} = \gamma_i w_{ij}$ and $\sigma_{ij} = \sigma_i w_{ij}$. Since we also have N=112, T=16 and K=1, the conditions $N \geq 3$, $K \geq 1$ and $T \geq 4+2K$ are satisfied, as a result of which the parameters are identified and can be estimated. We also found that the matrix $\boldsymbol{W}^*$ has full rank, rank($\boldsymbol{W}^*$)=112.[11] Although it is not necessary for the latter condition to be satisfied when using space-time data, this shows that each equation has at least one unique independent explanatory variable.

Since the constant is not included[12], each equation contains five parameters to be estimated: the coefficient δ which reflects the endogenous interaction effect of the output gap variable, the coefficient β which reflects the effect of the unemployment gap variable in the region itself, the coefficient γ which reflects the exogenous interaction effect of the unemployment gap variable in neighboring regions, the coefficient σ which reflects the interaction effect among the error terms (the non-diagonal elements of the variance-covariance matrix) and, finally, the variance of the error term of the regions themselves (the diagonal elements of the variance-covariance matrix).

Table 2 reports the estimation results when spatial interaction effects in the output gap variable, the unemployment gap variable and the error terms are taken into account.[13] Just as in Table 1, we computed the mean and the variance of the mean for each coefficient. This calculation is slightly different from that in Table 1, because the error terms of the different equations are now assumed to be correlated with each other. The formulas for the mean have

---

[11] If $\boldsymbol{W}_b$ denotes the binary contiguity matrix before it is row-normalized, then $\boldsymbol{W}^* = \boldsymbol{I}_N + \boldsymbol{W}_b$, where $\boldsymbol{I}_N$ is the identity matrix for the N=112 regions in the sample.
[12] Formally, the condition $T \geq 4+2K$ changes into $T \geq 3+2K$ as a result.
[13] Note that notation of variables is changed into vector form.

remained the same, but the variance of the mean coefficient of the unemployment gap variable is now calculated as $\text{var}(\overline{\beta}) = \text{var}(\frac{1}{N}\sum_i \beta_i) = \frac{1}{N^2}[\sum_i \text{var}(\beta_i) + \sum_i \sum_j \text{cov}(\beta_i, \beta_j)]$. Similar formulas apply to the coefficients $\overline{\delta}$ and $\overline{\gamma}$. The results in Column (1) of Table 2 show that there is a positive and significant endogenous interaction effect in the output gap variable (0.015; t-value 5.79), but that its impact is nonetheless very small. This not only follows from its mean, but also from the range of this interaction effect over different regions. With a range of -0.03 to 0.06, the endogenous interaction effect is almost negligible. By contrast, the positive and significant exogenous interaction effect in the unemployment gap variable (0.702; t-value 3.10) is almost as large as the impact of the unemployment gap variable of the region itself (0.732; t-value 3.96). Finally, the interaction effect among the error terms appears to be positive. Unfortunately, we are not able to ascertain whether this effect is significant, since the 3SLS estimator only produces the variance-covariance matrix of the response parameters of the model, and not of the parameters determining the variance-covariance matrix of the error terms.

<< Table 2 around here >>

To test for serial dependence among the observations over time, we considered two extensions of the model. The first extension augments the matrix of explanatory variables of each equation with an additional column containing the serial lagged residuals. If no fit is found, the hypothesis of serial dependence in the error terms is rejected. This test is known as Breusch and Godfrey's Lagrange Multiplier test. The second extension augments the matrix of explanatory variables for each equation with an additional column containing the serial lagged dependent variable. Mizon (1995) has argued that serial autocorrelation correction cannot be considered as a serious effort to find the "correct" equation and that one may better start with a linear dynamic regression model in which the dependent variable is regressed on its serial lagged value. The results of these two auxiliary regressions are reported in Columns (2) and (3) of Table 2, and show that the coefficients of the serial lagged residuals and of the serial lagged dependent variable are not significant. In other words, the hypothesis of serial dependence in the residuals or in the dependent variable may be rejected. For this reason, we prefer the results reported in Column (1) of Table 2.

3.4 Direct and indirect effects

Since our model contains both endogenous and exogenous interaction effects, we will follow LeSage and Pace (2009) in computing Okun's coefficient. They show that a change in a single explanatory variable in region i has a "direct impact" on region i as well an "indirect impact" on other regions $j \neq i$. Since our model of Okun's law does not contain an intercept and since we only have one explanatory variable (K=1), the reduced form can be written as (see also Eq. 7)

$$\begin{bmatrix} (y-y^*)_1 \\ (y-y^*)_2 \\ . \\ (y-y^*)_N \end{bmatrix} = \begin{bmatrix} 1 & -\delta_1 w_{12} & . & -\delta_1 w_{1N} \\ -\delta_2 w_{21} & 1 & . & -\delta_2 w_{2N} \\ . & . & . & . \\ -\delta_N w_{N1} & -\delta_N w_{N1} & . & 1 \end{bmatrix}^{-1} \begin{bmatrix} \beta_1 & \gamma_{11} w_{12} & . & \gamma_{11} w_{1N} \\ \gamma_{21} w_{21} & \beta_2 & . & \gamma_{21} w_{2N} \\ . & . & . & . \\ \gamma_{N1} w_{N1} & \gamma_{N1} w_{N2} & . & \beta_N \end{bmatrix} \begin{bmatrix} (u-u^*)_1 \\ (u-u^*)_2 \\ . \\ (u-u^*)_N \end{bmatrix} \quad (12)$$

where the subscript t has been left aside for simplicity. In matrix notation this model takes the form ($\boldsymbol{Y}$-$\boldsymbol{Y}^*$) =$\boldsymbol{B}^{-1}\boldsymbol{C}$($\boldsymbol{U}$-$\boldsymbol{U}^*$). Consequently, the direct or local effect of a change in (u-u^*) in region i on (y-y^*) in region i, and the indirect or neighbor effect of a change in (u-u^*) in region j on (y-y^*) in region i are, respectively,

$$\frac{\partial (y-y^*)_i}{\partial (u-u^*)_i} = (\boldsymbol{B}^{-1}\boldsymbol{C})_{ii} \quad \text{and} \quad \frac{\partial (y-y^*)_i}{\partial (u-u^*)_j} = (\boldsymbol{B}^{-1}\boldsymbol{C})_{ij} \,. \quad (13)$$

Since both effects are different for different regions, we are again computing averages. The average direct/local effect is simply the average of the first formula in (13) calculated over N regions. However, when computing the indirect/neighbor effect, we have two possibilities. We can compute the effect on output in a particular region from reducing unemployment in all other regions, that is, the average of the second formula in (13) calculated over index j given i.

Alternatively, we can compute the effect of reducing unemployment in a particular region on the output of all the other regions, that is, the average of the second formula in (13) calculated over index i given j. LeSage and Pace (2009) show that the numerical magnitudes of both calculations are the same, as a result of which it does not matter which one is used. In sum, we have

$$\text{Direct/local effect} = \frac{1}{N}\sum_{i=1}^{N} (\boldsymbol{B}^{-1}\boldsymbol{C})_{ii}, \tag{14a}$$

$$\text{Indirect/neighbor effect} = \frac{1}{N}\sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} (\boldsymbol{B}^{-1}\boldsymbol{C})_{ij} = \frac{1}{N}\sum_{j=1}^{N} \sum_{i=1, i\neq j}^{N} (\boldsymbol{B}^{-1}\boldsymbol{C})_{ij}, \tag{14b}$$

Okun’s coefficient = Direct/local effect + Indirect/neighbor effect. (14c)

Using these formulas, Okun’s coefficient appears to be 1.45, on average, and to consist of a direct/local effect of 0.732 and an indirect/neighbor effect of 0.717. In other words, the inclusion of interaction effects has two main effects on Okun’s coefficient. First, there is a drop from 1.90 to 1.45. The estimate of 1.90 was the result of our analysis when interaction effects were not included (Table 1), which has been the commonly used method of estimating Okun’s law up until now. The explanation for the drop in Okun’s coefficient is that regions may not be treated as independent units. The two basic assumptions underlying previous studies are that observations on the explanatory variables do not convey information about the expected value of the disturbances and that the disturbances convey no information about each other. However, since we found significant interaction effects in the dependent variable as well as in the error terms, both assumptions have been violated. Second, Okun’s coefficient breaks up almost equally into the direct/local effect and the indirect/neighbor effect.

Since the endogenous interaction effects are relatively small, so are the differences between the direct/local effect and the coefficient of (***U***-***U****), and between the indirect/neighbor effect and the coefficient ***W***(***U***-***U****). This finding indicates that Okun’s law constitutes a local rather than a global relationship. Only if the endogenous interaction effects were to be larger, would Okun's law turn into a global relationship, where a reduction in unemployment in one particular region would benefit output in all regions of the European continent. This, however, turns out not to be the case. By reducing unemployment in one particular region, the output of that region as well as the output of its neighbors will increase, but not in any regions located farther away. Nevertheless, due to this local/neighbor effect, the benefits of any anti-unemployment policy still partly accrue to neighboring regions. According to Overman and Puga (2002), simply adding a regional component to this policy is therefore not enough. Since politicians gain no votes or tax revenues from these spillovers, they are likely to underestimate the true benefit of this policy when the cost is completely borne by the region itself, and thus they tend to not implement it vigorously enough.

Figure 1 maps Okun's coefficient for the 112 regions in the sample. Moran's I shows that there is a match between value similarity and locational similarity. Using univariate LISA (Anselin, 1995), two clusters of regions could be identified, namely eight adjacent regions in Spain with Okun's coefficient in the highest quartile and seven adjacent regions in France with Okun's coefficient in the lowest quartile. However, it is conceivable that there are also clusters of regions which do not neighbor each other, for example, peripheral regions and centrally located regions. The analysis of whether additional clusters exist and whether differences in Okun's coefficient among the regions may be explained by both national and regional institutions that determine the rigidity or flexibility of national and regional labor markets is an interesting topic for further research.[14]

<< Figure 1 around here >>

[14] We also found a match between value similarity and locational similarity for the endogenous interaction effect (Moran's I=0.153, p-value 0.01). Low values for the endogenous interaction effect in the output gap variable were found in France, the north of the Netherlands and the south of West Germany, whereas high values were found in Belgium, Spain (with some regions excepted), the south and the middle of the Netherlands, the north and the middle of West Germany, Italy (except for the south-east) and the north-east of France. For the local coefficient estimates of the other variables, (***U-U****) and ***W***(***U-U****), and the interaction effect in the error term, ***W***ε, Moran's I appeared to be insignificant (-0.022 (p-value 0.42), -0.088 (p-value 0.09), and -0.010 (p-value 0.53), respectively).

3.5 Okun's coefficient and robustness checks

We have also tested the robustness of our results using another spatial weights matrix and by renouncing the use of a spatial weights matrix. If the model is estimated using an inverse distance instead of a binary contiguity matrix, Okun's coefficient has the tendency to increase. Although the impact from regions located farther away decreases with distance, one objection to this matrix is that every region has the same neighborhood set. A cut-off point beyond which regions are no longer considered to be neighboring partly solves this problem but, just as in considering first-order neighbors alone, any postulated choice of this cut-off point is arbitrary. In principle, the same would apply to the postulated functional form of the distance decay function.

If the model is estimated without using a spatial weights matrix, the parameters δ_{ij}, γ_{ij} and σ_{ij} are freely estimated if j is considered to be a neighbor of i, and set to zero if j is considered not to be a neighbor of i. Since K=1, the model equations do not contain a constant and the covariance matrix may be asymmetric, and thus we only have sufficient observations in relation to the number of parameters to be estimated under these circumstances if $T \geq 1+K+qK+2q=2+3q$. This implies that the maximum number of neighbors of each region should not be greater than 4. In addition, we should have rank($\boldsymbol{W}$*)=112, otherwise the parameters will not be identified. We have experimented with q=1, 2, 3 and 4.

If q=1, the results of our experiments were virtually the same as those for the binary contiguity matrix. Although it is clear that one might do better to consider the average of all neighboring regions, this does indicate that, in considering only one neighbor, the model has already been able to catch the interaction effects. By contrast, if q>1, Okun's coefficient has the tendency to decrease. However, one objection to this setup was that we often found opposite signs for the various interaction effects of the q different neighbors and that these q interaction effects partly cancelled each other out when taken together. The explanation for this is that the number of observations in relation to the number of parameters is relatively small and that the correlations between the time series of different neighbors, both for the output gap variables and the unemployment gap variables, are relatively high. It is obvious that these problems are avoided by considering the average of all neighboring regions.

The results of our robustness checks point out that more research is needed on the geographical scale of Okun's law and that the best way to do this is by using a spatial weights matrix whose elements are parameterized such that the distance decay effect can be empirically estimated instead of imposed a priori. At the same time, it should be stressed that this would require nonlinear 3SLS estimation techniques. Since we have one equation for every unit in the sample and thus large numbers of parameters to be estimated, this would complicate the analysis considerably, both from a programming as well as from a numerical point of view.

3.6 The impact of ignoring one or more interaction effects

In the introduction to this paper, we have pointed out that the interaction parameters in a single equation model are not identified if a researcher wants to include endogenous interaction effects, exogenous interaction effects, as well as interaction effects among the error terms. The researcher must perforce abandon one or more interaction effects. Table 3 shows the bias in Okun's coefficient as a result of one or more of these restrictions. This bias is computed by comparing Okun's coefficient of the restricted model with that of the unrestricted model, that is, if all interaction effects are included.

<< Table 3 around here >>

The greatest bias in Okun's coefficient occurs when the endogenous interaction effects $\boldsymbol{W}(\boldsymbol{Y}\text{-}\boldsymbol{Y}^*)$ are ignored. The bias ranges from 33.7% to 48.1%. Even though the coefficient of the endogenous interaction effects appeared to be small, its impact on Okun's coefficient is apparently tremendous. When endogenous interaction effects are included, the bias in Okun's coefficient falls to 5.0% up to 8.2% (in absolute values), dependent on whether exogenous interaction effects, interaction effects among the error terms, or both are ignored. However, the estimates of the direct/local effect and the indirect/neighbor effect can still be severely biased. The smallest bias of 5.0% is obtained only when interaction effects among the error terms are ignored. The explanation for this is simple. If the interaction effects among the error terms are ignored, every equation can be separately estimated from the others by 2SLS. Although this estimator has the disadvantage that it will lose its property of being efficient, it will not lose its property of being unbiased. Consequently, the difference between a) Okun's coefficient of the model with endogenous and exogenous interaction effects only and b) Okun's coefficient of the model with endogenous and exogenous interaction effects and interaction effects among the error terms will be small. By contrast, if endogenous and/or exogenous interaction effects are ignored, we have an omitted variable bias, as a result of which the difference becomes larger.

In the introduction to this paper, we also stated that Harry Kelejian advocated single equation models that include both a spatially lagged dependent variable and a spatially autocorrelated error term, while James LeSage advocated single equation models that include both a spatially lagged dependent variable and spatially lagged explanatory variables. One argument in favor of LeSage's call is that ignoring spatial error autocorrelation in a single equation model only leads to a loss of efficiency, while ignoring exogenous interaction effects leads to biases in the parameter estimates. One argument in favor of Kelejian's call is that a tolerably biased estimator with a small variance should be preferable to an unbiased estimator with a larger variance. Given the results of our empirical study, LeSage's call is to be more warmly recommended than Kelejian's call, but this is no guarantee that another empirical study will find comparable results.

4. Conclusions

We developed a simultaneous equations model having endogenous interaction effects, exogenous interaction effects and interaction effects among the error terms. In contrast to the single equation model in Manski (1993), the interaction parameters of this model are identified and can be estimated by space-time data, provided that:

1. The number of spatial units in the sample is greater than or equal to 3, $N \geq 3$;
2. The number of independent explanatory variables is greater than or equal to 1, $K \geq 1$;
3. The number of observations is sufficiently large, $T \geq 4+2K$; and
4. Information is used about the nature of interactions over space in the form of a spatial weights matrix $\boldsymbol{W}$ in order to reduce the number of each type of interaction effect to one.

Alternative conditions when using space-time data as well as those conditions required when using cross-sectional data have also been derived.

Using this model, we estimated Okun's law for 112 Western European regions over the period 1986-2001. Okun's coefficient appeared to be 1.45, on average, and to consist of a direct/local effect of 0.732 and an indirect/neighbor effect of 0.717. We compared this figure of 1.45 with one where interaction effects are not included, which has been the commonly used method for estimating Okun's law up until now. Without the interaction effects, Okun's coefficient appeared to be 1.90, on average, which represents an overestimation of 36.2%. We also computed the bias in Okun's coefficient if interaction effects were included, although not all of them, in order to simulate the disadvantage of the single equation model where the researcher must perforce abandon one or more interaction effects for reasons of identification. If endogenous interaction effects are left aside, the bias appears to be 35% at a minimum, whereas the bias ranges from 5% to 9% (in absolute values) if endogenous interaction effects are included but exogenous interaction effects and/or the interaction effect among the error terms are left aside. Since the coefficient estimate of the endogenous interaction effects appeared to be small (between -0.03 and 0.06, with an average of 0.015), Okun's law should be interpreted as a local rather than a global relationship. By reducing unemployment in a particular region, the output of that region as well as the output of its neighbors will increase, but will not do so in any regions located farther away.

In addition, we have identified three interesting topics for further research. The first topic is whether there is a relationship between Okun's coefficient and national and regional institutions that determine the rigidity or flexibility of national and regional labor markets. The second topic is the determination of the geographical scale of Okun's law by using a spatial weights matrix whose elements are parameterized such that the distance decay effect can be empirically estimated instead of imposed a priori. The third topic is the adjustment of the covariance matrix of the disturbances such that the model can also be estimated by Geographically Weighted Regression when using cross-sectional data.

References

Abreu, M., H.L.F. de Groot and R.J.G.M. Florax (2005) Space and Growth: A Survey of Empirical Evidence and Methods. *Région et Développement* 21: 13-44.

Adanu, K. (2005) A Cross-Province Comparison of Okun's Coefficient for Canada. Applied Economics 37: 561-570.

Anselin, L. (1988) Spatial Econometrics: Methods and Models. Kluwer: Dordrecht.

Anselin, L. (1995) Local Indicators of Spatial Association—LISA. *Geographical Analysis* 27: 93-115.

Anselin, L. and A. Bera (1998) Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In: Ullah, A. and D. Giles (eds) Handbook of Applied Economics Statistics, pp. 237-289. Marcel Dekker: New York.

Anselin, L., A. Bera, R. Florax and M. Yoon (1996) Simple Diagnostic Tests for Spatial Dependence. *Regional Science and Urban Economics* 26: 77-104.

Apergis, N. and A. Rezitis (2003) An Examination on Okun's Law: Evidence from Regional Areas in Greece. *Applied Economics* 35: 1147-1151.

Attfield, C.L.F. and B. Silverstone (1997) Okun's Coefficient: A Comment. *Review of Economics and Statistics* 79: 326-329.

Attfield, C.L.F. and B. Silverstone (1998) Okun's Law, Cointegration and Gap Variables. *Journal of Macroeconomics* 20: 625-637.

Barreto, H. and F. Howland (1993) There are Two Okun's Law Relationships between Output and Unemployment. Mimeo, Wabash college, Crawfordsville (USA).

Christopoulos, D.K. (2004) The Relationship between Output and Unemployment : Evidence from Greek Regions. *Papers in Regional Science* 83: 611-620.

Cuaresma, J.C. (2003) Okun's Law Revisited. *Oxford Bulletin of Economics and Statistics* 65: 439-451.

Elhorst, J.P. (2003a) Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review* 26: 244-268.

Elhorst, J.P. (2003b) The Mystery of Regional Unemployment Differentials: Theoretical and Empirical Explanations. *Journal of Economic Surveys* 17: 709-748.

Elhorst, J.P. and J. Oosterhaven (2008) Integral Cost-Benefit Analysis of Maglev Projects under Market Imperfections. *Journal of Transport and Land Use*, 1: 65-87.

Engle, R. and C. Granger (1987) Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica* 35: 251-276.

Ertur, C. and W. Koch (2007) Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence. *Journal of Applied Econometrics* 22: 1033-1062.

Fingleton, B. (2003) European Regional Growth. Springer: Berlin.

Fotheringham, A.S., C. Brunsdon and M. Charlton (2002) Geographically Weighted Regression. Wiley: Chichester.

Freeman, D.G. (2000) Regional Tests on Okun’s Law. *International Advances in Economic Research* 6: 557-570.

Freeman, D.G. (2001) Panel Tests of Okun’s Law for Ten Industrial Countries. *Economic Inquiry* 39: 511-523.

Greene, W.H. (2003). Econometric analysis, 5th edition. Prentice Hall, New Jersey.

Helpman, E. (2004) The Mystery of Economic Growth. Harvard University Press: Cambridge (USA).

Hodrick, R. and E. Prescott (1997) Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29: 1-16.

Hsiao, C. (1980) Identification. In: Griliches, Z. and M.D. Intriligator (eds.) Handbook of Econometrics, Volume 1, pp. 223-283. North-Holland: Amsterdam.

Huang, H.-C. and S.-C. Lin (2006) A Flexible Nonlinear Inference to Okun’s Relationship. *Applied Economics Letters* 13: 325-331.

Huang, H.-C. and S.-C. Lin (2008) Smooth-Time Varying Okun’s Coefficients. *Economic Modelling* 25: 363-375.

Izyumov, A. and J. Vahaly (2002) The Unemployment-Output Tradeoff in Transition Economies; Does Okun’s Law Apply? *Economics of Planning* 35: 317-331.

Johnston, J. and J. DiNardo (1997) Econometric Methods, 4th edition. McGraw-Hill: New York.

Kelejian, H.H. and I.R. Prucha (2004) Estimation of Simultaneous Equations Systems of Spatially Interrelated Cross Section Equations. *Journal of Econometrics* 118: 27-50.

Kosfeld, R. and C. Dreger (2006) Thresholds for Employment and Unemployment: A Spatial Analysis of German Regional Labour Markets, 1992-2000. *Papers in Regional Science* 85: 523-542.

Lee, J. (2000) The Robustness of Okun’s Law: Evidence from OECD Countries. *Journal of Macroeconomics* 22: 331-356.

Lee, L. (2004) Asymptotic distribution of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* 72: 1899-1925.

LeSage, J.P. and R.K. Pace (2009) Introduction to Spatial Econometrics. CRC Press/Taylor & Francis Group.

López-Bazo. E., E. Vayá and M. Artís (2004) Regional Externalities and Growth: Evidence from European Regions. *Journal of Regional Science* 44: 43-73.

Manski, C.F. (1993) Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies* 60: 531-542.

Mizon, G.E. (1995) A Simple Message for Autocorrelation Correctors: Don’t. *Journal of Econometrics* 69: 267-288.

Moosa, I.A. (1997) A Cross-Country Comparison of Okun’s Coefficient. *Journal of Comparative Economics* 24: 335-356.

Moosa, I.A. (1999) Cyclical Output, Cyclical Unemployment, and Okun’s Coefficient; A Structural Time Series Approach. *International Review of Economics and Finance* 8: 293-304.

Okun, A.M. (1962) Potential GNP: Its Measurement and Significance. American Statistical Association. *Proceedings of the Business and Economic Statistics Section:* 98-104.

Overman, H.G. and D. Puga (2002) Unemployment Clusters across Europe's Regions and Countries. *Economic Policy*, April, 116-147.

Perman, R. and C. Tavera (2005) A Cross-Country Analysis of the Okun’s Law Coefficient Convergence in Europe. *Applied Economics* 37: 2501-2513.

Perman, R. and C. Tavera (2007) Testing for Convergence of the Okun's Coefficient in Europe. *Empirica* 34: 45-61.

Pesaran, M.H. (2007) A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence. *Journal of Applied Econometrics* 22: 265-312.

Pesaran, M.H. and R. Smith (1995) Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics* 68: 79-113.

Prachowny, M.F.J. (1993) Okun's Law: Theoretical Foundations and Revised Estimates. *Review of Economics and Statistics* 75: 331-336.

Rey, S.J. and M.J. Janikas (2005) Regional Convergence, Inequality, and Space. *Journal of Economic Geography* 5: 155-176.

Silvapulle, P., I.A. Moosa and M.J. Silvapulle (2004) Asymmetry in Okun's Law. *Canadian Journal of Economics* 37: 353-374.

Sögner, L. (2001) Okun's Law, Does the Austrian Unemployment-GDP Relationship Exhibit Structural Breaks. *Empirical Economics* 26: 553-564.

Sögner, L. and A. Stiassny (2002) An Analysis on the Structural Stability of Okun's Law —A Cross-Country Study. *Applied Economics* 14: 1775-1787.

Villaverde, J. and A. Maza (2007) Okun's Law in the Spanish Regions. *Economics Bulletin* 18: 1-11.

Weber, C.E. (1995) Cyclical Output, Cyclical Unemployment, and Okun's Coefficient: A New Approach. *Journal of Applied Econometrics* 10: 433-445.

Weber, C.E. and J.E. West (1996) Functional Form in Regression Models of Okun's Law. *Applied Economics Letters* 3: 607-609.

White, E.N. and G.J.D. Hewings (1982) Space-Time Employment Modeling: Some Results Using Seemingly Unrelated Regression Estimators. *Journal of Regional Science* 22: 283-302.

Table 1 Coefficient estimates of Okun's law without interaction effects

Dependent variable y-y*		OLS			Dynamic OLS		
Explanatory variable	Coef.	Coef.	t-value	Range	Coef.	t-value	Range
u-u *	$-\overline{\beta}$	1.953	9.37	0.17/4.17	1.902	8.14	0.19/4.17

Table 2 Coefficient estimates of Okun's law with interaction effects

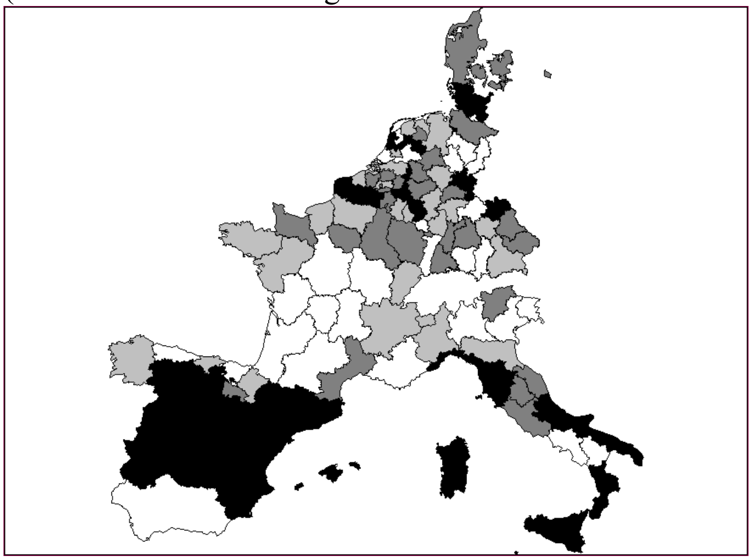
Dependent variable ***Y-Y****		(1) 3SLS			(2) 3SLS		(3) 3SLS	
Explanatory variables	Coef.	Coef.	t-value	Range	Coef.	t-value	Coef.	t-value
U-U*	$-\overline{\beta}$	0.731	3.96	-1.74/3.95	0.788	3.61	1.902	8.14
***W*(*U-U**)**	$-\overline{\gamma}$	0.702	3.10	-3.18/3.64	0.612	2.45	0.840	3.61
***W*(*Y-Y**)**	$\overline{\delta}$	0.015	5.79	-0.03/0.06	0.019	6.84	0.635	2.45
***W*ε**	$\overline{\sigma}$	0.172		-0.01/0.37	0.179		0.180	
ε(t-1)	$\overline{\lambda}$				-0.033	-0.78		
Y(t-1)-***Y****(t-1)	$\overline{\rho}$						-0.026	-0.67

Table 3 Coefficient estimates and bias in Okun's coefficient when ignoring one or more interaction effects

Interaction effects ignored	Coef. ***W*(*Y-Y**)** $\overline{\delta}$	Coef. ***W*ε** $\overline{\sigma}$	Coef. ***W*(*U-U**)** $-\overline{\gamma}$	Coef. (***U-U****) $-\overline{\beta}$	Direct/ local effect	Indirect/ neighbor effect	Okun's coef.	% bias Okun's coef.
None	0.015	0.172	0.702	0.731	0.732	0.717	1.449	0
***W*(*Y-Y**)**, ***W*(*U-U**)**, ***W*ε**				1.953	1.953	-	1.953	34.8
***W*(*Y-Y**)**, ***W*(*U-U**)**		0.165		1.937	1.937	-	1.937	33.7
***W*(*Y-Y**)**, ***W*ε**			1.516	0.613	0.613	1.516	2.130	47.0
***W*(*U-U**)**, ***W*ε**	0.019			1.340	1.340	0.021	1.361	-6.1
***W*(*Y-Y**)**		0.176	1.386	0.760	0.766	1.386	2.146	48.1
***W*(*U-U**)**	0.019	0.169		1.307	1.307	0.023	1.330	-8.2
***W*ε**			0.855	0.651	0.652	0.869	1.522	5.0

The direct/local effect, the indirect/neighbor effect and Okun's coefficient are obtained by Eqs. (14).

Figure 1 Okun's coefficient subdivided into four quartiles (darker colors indicate higher values for Okun's coefficient)



Moran's I=0.132 (p-value 0.02)