WP 39_08

Johan H.L. Oud
Radboud University Nijmegen, The Netherlands

Henk Folmer
University of Groningen, The Netherlands
Northwest A&F University, China

Roberto Patuelli
University of Lugano, Switzerland
The Rimini Centre for Economic Analysis (RCEA), Italy

Peter Nijkamp
VU University Amsterdam, The Netherlands
The Rimini Centre for Economic Analysis (RCEA), Italy

CONTINUOUS-TIME MODELLING WITH SPATIAL DEPENDENCE

Continuous-Time Modelling with Spatial Dependence

Johan H.L. Oud
Radboud University Nijmegen, The Netherlands
j.oud@pwo.ru.nl

Henk Folmer[1]
University of Groningen, The Netherlands
Northwest A&F University, China
Henk.Folmer@wur.nl

Roberto Patuelli
University of Lugano, Switzerland
The Rimini Centre for Economic Analysis, Italy
roberto.patuelli@usi.ch

Peter Nijkamp
VU University Amsterdam, The Netherlands
pnijkamp@feweb.vu.nl

Abstract

(Spatial) panel data are routinely modelled in discrete time (DT). However, there are compelling arguments for continuous time (CT) modelling of (spatial) panel data. Particularly, most social processes evolve in CT, so that statistical analysis in DT is an oversimplification, gives an incomplete representation of reality and may lead to misinterpretation of estimation results. The most compelling reason for a CT approach is that, in contrast to DT modelling, it allows adequate modelling of dynamic adjustment processes. The paper introduces spatial dependence in a CT modelling framework. We propose a nonlinear Structural Equation Model (SEM) with latent variables for estimation of the Exact Discrete Model (EDM), which links the CT model parameters to the DT observations. The use of a SEM with latent variables makes it possible to take measurement errors in the variables into account, leading to a reduction of attenuation bias (i.e., disattenuation). The SEM-CT model with spatial dependence developed here is the first dynamic structural equation model with spatial dependence. The spatial econometric SEM-CT framework is illustrated on the basis of a simple regional labour market model for Germany made up of the endogenous state variables unemployment change and population change and of the exogenous input variables change in regional average wage and change in the structure of the manufacturing sector.

Keywords: continuous-time modelling, structural equation modelling, latent variables, spatial dependence, panel data, disattenuation, measurement errors, unemployment change, population change, Germany. **JEL codes**: C33, E24, O18, R11.

This version: 5/10/2010

---

[1] Corresponding author.

1. Introduction

Socio-economic processes such as the development of prices, unemployment rates, investments, migration and political preferences are the outcomes of various decisions taken by different actors at different points in time. This basic feature gives rise to continuously evolving socio-economic dynamics, rather than to processes that change at specific discrete points in time only. The analyst, however, only observes the processes at discrete points in time (for example, yearly observations of regional unemployment). The typical approach in conventional (that is, discrete) time series modelling and panel data analysis is to ignore the continuous nature of the processes underlying discrete time observations. Consequently, discrete time series and discrete panel data analysis are simplifications of reality and may lead to bias in the mapping of dynamic adjustment processes of socio-economic phenomena and to misinterpretation of estimation results. Discrete-time (DT) analysis is at best a simplified approximation of real-world processes which take place in continuous time (CT) (2008).

   CT econometrics models the continuous nature of social processes by means of systems of differential equations. It departs from the assumption that different agents take different actions at different points in time. This assumption implies that there is no obvious time interval that can serve as a natural unit. This is in contrast to DT models (which are made up of systems of difference equations), which are typically formulated in relation to the data available (for example yearly or monthly data).

   A DT model estimated on the basis of, for example, monthly data will be different from a model estimated on the basis of annual data. CT models, however, are independent of the observation interval, and thus provide a common basis for accurate comparison of processes (Oud and Jansen 2000). These features enable the analyst to obtain predictions and simulations for any time interval, rather than for the time interval inherent to the data, as in the case of DT modelling.

   CT modelling is particularly useful for the analysis of dynamic adjustment processes (Gandolfo, 1993). Whereas in DT models it may not be possible to obtain an estimate of the adjustment speed when the time lags are short compared to the observation period, in CT models it is in general possible to obtain an asymptotically unbiased estimate of it. Specifically, CT modelling makes it possible to determine at what pace an effect builds up over continuously increasing intervals, at or between which observation points the maximum impact of an effect occurs, and at what pace it dies out. A CT model therefore allows a more satisfactory treatment of distributed-lag processes.

   CT modelling has a long history in the social sciences. In econometrics it was pioneered by, amongst others, Bartlett (1946), Koopmans (1950) and Phillips (1959); in sociology by Simon (1952) and Coleman (1968). For an overview, see amongst others, Bergstrom 1988). To our best knowledge, little attention has been paid to CT modelling in spatial econometrics.[2] The reverse also holds: In CT modelling no attention has been

---

[2] For interesting applications, see Donaghy and Plotnikova (2003), and Piras et al (2007). Note that in spatial analyses the units of observation are usually discrete. There is, however, a growing interest in theoretical work on continuous-space modelling (see, for example, Puu 1997). A great challenge in spatial modelling and spatial econometrics is to explore the relationships between continuous time and continuous space modelling and to integrate both (see, amongst others, Cressie 1993; Wackernagel 1995; Donaghy 2001).

paid to spatial dependence nor, more generally, to dependence among units of observation (testees).

In this paper we introduce spatial dependence in a CT modelling framework. Because of its relative unfamiliarity in geography and regional science, we also present a brief overview of the main characteristics of CT modelling compared to DT modelling and of estimation – by means of a Structural Equation Model (SEM) – of the Exact Discrete Model (EDM), which links the CT model to observations at discrete time points. We illustrate the CT modelling framework by a simple regional unemployment model for Germany. The model is made up of the endogenous state variables unemployment and population change and of the exogenous input variables change in regional average wage and change in the structure of the manufacturing sector.

2. Main Characteristics of CT Modelling

As observed above, in CT modelling the parameters estimated are independent of the observation intervals which implies that CT modelling makes it possible to 'fill out' the 'gaps' between the discrete observation time points by model-based estimates of complete autoregressive and cross-lagged effect curves as well as mean trajectories. Figure 1, which is taken from Delsing and Oud (2008), illustrates this feature. It presents the mean curve $E[\mathbf{x}(t)]$, that is, the estimated mean trajectory in the population; an individual subject-specific mean curve $E[\mathbf{x}(t)|\kappa]$ (κ the random parameter with the subject specific value) and the trajectory $E[\mathbf{x}(t)|\mathbf{y}]$ for the same individual subject. The latter, the conditional mean (conditional on the individual's data vector $\mathbf{y}$) or state space smoother (Durbin & Koopman, 2001), represents the best estimate of the individual's sample trajectory through CT on the basis of the total data vector $\mathbf{y}$.[3] $E[\mathbf{x}(t)|\kappa]$ and $E[\mathbf{x}(t)|\mathbf{y}]$, though both relating to the same individual subject, differ because they exploit different kinds of information. $E[\mathbf{x}(t)|\mathbf{y}]$ is based on all information available for the subject, while the subject-specific mean curve $E[\mathbf{x}(t)|\kappa]$ is the model expectation, conditioned on the single subject-specific parameter value (Oud & Singer, 2008).

Figure 1 shows that for $E[\mathbf{x}(t)|\mathbf{y}]$ the gaps between the measurements in 1986, 1988, 1990 and 1992 are filled out by interpolation. Similarly for the prediction interval 1992–98. Furthermore, uncertainty of the interpolations and predictions is given by confidence intervals (dotted lines).[4] The confidence intervals for $E[\mathbf{x}(t)|\mathbf{y}]$ typically go to zero at the observation points 1986, 1988, 1990 and 1992, but increase considerably in the prediction period after 1992. The confidence intervals show that the subject's sample trajectory $E[\mathbf{x}(t)|\mathbf{y}]$ is almost everywhere far above the mean curve $E[\mathbf{x}(t)]$, even in the prediction period, and for time points close to the measurement points below (for the first year) or above the individual's subject-specific mean curve $E[\mathbf{x}(t)|\kappa]$ as well. However, in the prediction period $E[\mathbf{x}(t)|\mathbf{y}]$ converges to $E[\mathbf{x}(t)|\kappa]$.

---

[3] Observe that the conditional mean includes error components.

[4] Note that the confidence intervals of $E[\mathbf{x}(t)|\mathbf{y}]$ have been determined by the Kalman-smoother (Commandeur et al. 2010) which accounts for prediction errors and, if present, measurement errors in the data, but ignores sampling fluctuations in parameter estimates. Since $E[\mathbf{x}(t)|\kappa]$ and $E[\mathbf{x}(t)]$ only depend on parameter values and are not conditioned on the data, no Kalman-smoother confidence intervals are calculated.

Both in CT and DT modelling the model parameters are estimated on the basis of discrete observations. However, in DT modelling the model parameters are used to estimate $E[\mathbf{x}(t)]$, $E[\mathbf{x}(t)|\kappa]$ and $E[\mathbf{x}(t)|\mathbf{y}]$ for the four observation time points (1986, 1988, 1990, 1992) only, and possibly used for extrapolation for the three prediction time points (1994, 1996, 1998). In contrast, CT modelling uses the model estimates to estimate the complete curves $E[\mathbf{x}(t)]$, $E[\mathbf{x}(t)|\kappa]$ and $E[\mathbf{x}(t)|\mathbf{y}]$ for the entire interval 1986–98 by interpolation, and for the prediction interval 1992–98 by prediction. Clearly, the interpolations and predictions in CT are subject to uncertainty. However, the uncertainty is quantified in the model by means of confidence intervals.[5]
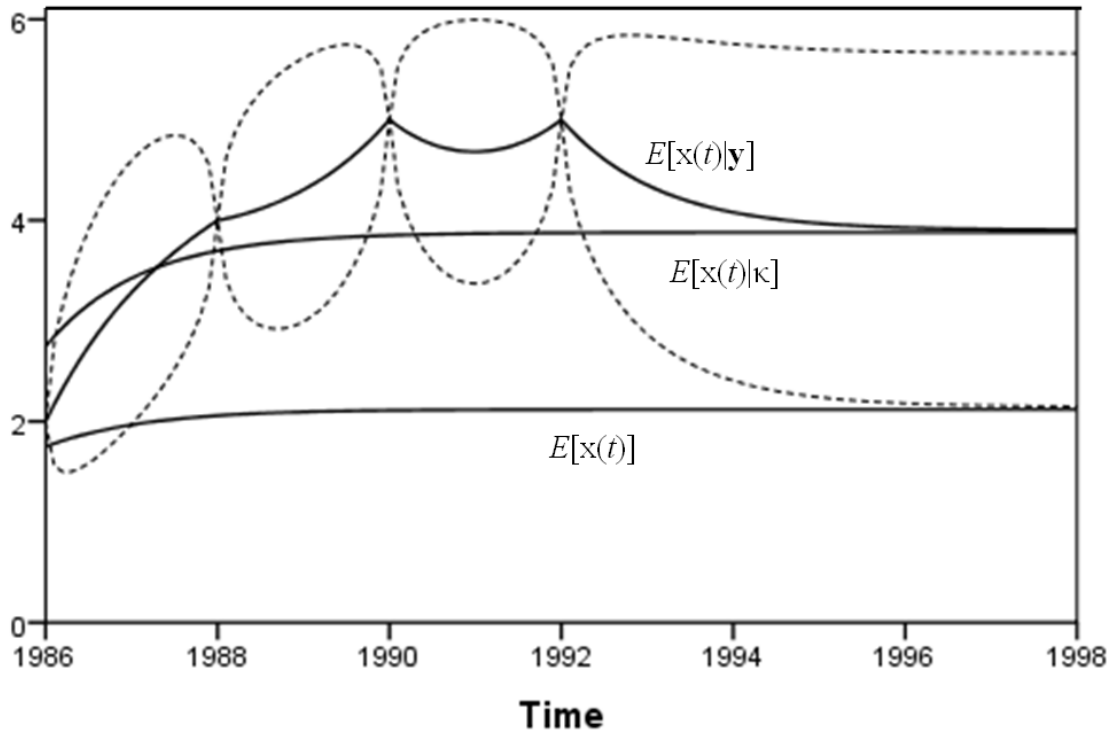

Figure 1. Estimated mean curve $E[\mathbf{x}(t)]$, subject-specific mean curve $E[\mathbf{x}(t)|\kappa]$ and sample
trajectory curve $E[\mathbf{x}(t)|\mathbf{y}]$ (for the same subject as in $E[\mathbf{x}(t)|\kappa]$) in CT
(confidence intervals for the sample trajectory curve represented by dotted
lines)

DT modelling may be misleading in the case of unequal observation intervals within the same study, or when comparisons between different studies are made. This is shown in Figure 2, where (hypothetical) autoregression functions in CT of two studies, A and B, with two different observation intervals (0.50 in study A and 1.00 in study B) are depicted. Because CT autoregression function B exceeds autoregression function A everywhere, the obvious conclusion is that the autoregression in study A is lower than in study B. Nevertheless, in DT, for example at observation interval 0.50 (half-yearly observation interval), study A finds the autoregression value of 0.61 which is

---

[5] For statistical details, see Section 4.

considerably larger than the value of 0.50 found at the interval 1.00 (yearly observation interval) in study B, which might lead to the erroneous conclusion that the autoregression in study A is larger than in study B. To obtain the correct answer, CT modelling is required, which makes it possible to compare the complete autoregression functions. An important corollary is that CT modelling allows combining the data of several studies with different observation intervals into one data set, testing whether the underlying CT parameters are equal and, if so, presenting one and the same autoregression function (Oud 2001).
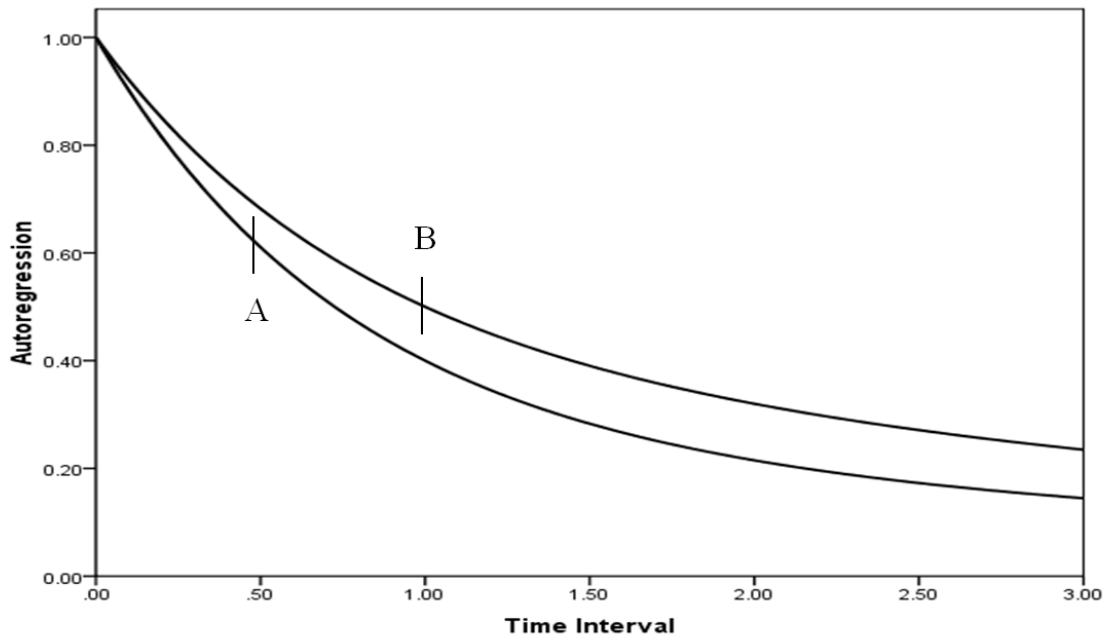

Figure 2.Two different autoregression functions in two different studies A and B.

The misleading results of DT modelling in the case of unequal observation intervals may perhaps seem irrelevant when there are conventions to conduct surveys at similar intervals. However, using equal observation intervals is often no solution to the problems inherent to DT modelling. To see this, consider the two (hypothetical) CT reciprocal cross-lagged effect functions for x_1 and x_2 in Figure 3.[6] Unlike autoregression functions, which start at value 1, cross-lagged effect functions have starting value 0 (different variables cannot have any influence on each other over a time interval of length zero), increase until the maximum is reached (in Figure 3 the maxima 0.250 and 0.240 are reached at the time points 1.02 and 1.64, respectively), and next taper off to 0 (in a stable model). Figure 3 shows that the two cross-lagged effect functions cross at interval 1.44. Both have here the same value 0.239, but are different over all other intervals. Particularly, for observation intervals < 1.44, the effect of x_1 on x_2 is larger than the reverse effect, while the opposite holds for intervals > 1.44. Using one and the same interval in DT does not resolve the problem. To illustrate this, assume that in DT only the

---

[6] A cross-lagged effect function gives the effect of one variable on another as it develops over time.

observation interval 1.44 was used. This would give that the effect of x_1 on x_2 is equal to the effect of x_2 on x_1. However, the conclusion that both effects are equal everywhere would be false. Furthermore, using in DT equal intervals > 1.44 would lead to the finding that the effect of x_2 on x_1 exceeds the effect x_1 on x_2 . Again, the conclusion that the effect of x_2 on x_1 exceeds the effect of x_1 on x_2 everywhere would be false. CT analysis, particularly estimating the continuous-time effects and displaying the full autoregression and cross-lagged effect functions over the entire time axis, would be the way to obtain the correct answer. Note that the problems of using equal observation intervals in DT modelling, which are explained here for the case of crossing cross-lagged effect functions only, would also apply in the case of crossing autoregression functions.
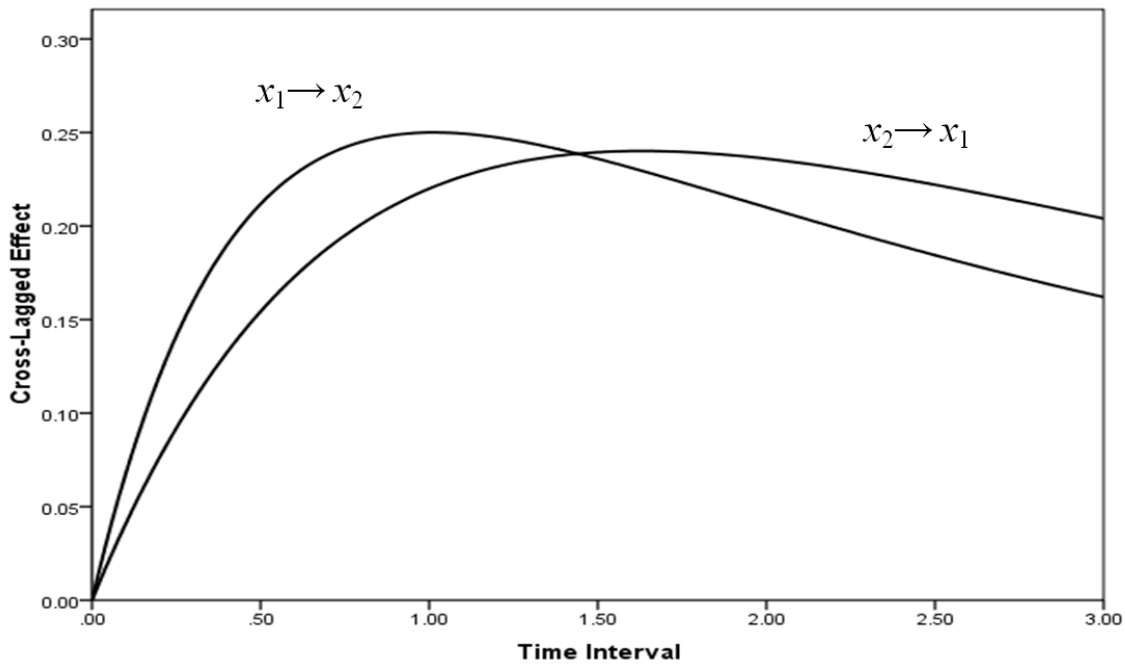

Figure 3. Cross-lagged effect functions for the reciprocal effects between x_1 and x_2

3. Specification and Estimation of the Spatial-Dependence Continuous-Time Model

Let $\mathbf{x}(t)$ be the n-dimensional (endogenous) vector of state variables,[7] $\mathbf{u}(t)$ the r-dimensional vector of (exogenous) fixed input variables, $\boldsymbol{\kappa}$ the n-dimensional vector of subject (region-) specific deviations from the fixed-mean intercept (that is, a vector of

---

[7] In addition to observed state variables, we consider the possibility that the state variables are latent variables, that is, variables that cannot be directly observed due to unobservability or measurement error. Latent variables are measured by one or more indicators. For instance, the latent variable socioeconomic status is usually measured by means of the indicators income, education, and profession. In the present paper, we only deal with latent variables that are measured by one indicator, though with error. The presence of latent variables requires the use of a measurement model relating the latent variables to their indicators, and a structural model which presents the relationships between the latent variables. For details, see Oud and Folmer (2008).

5

'random subject effects' or 'unobserved unit heterogeneity') and $\mathbf{W}(t)$ the standard multivariate Wiener process. We consider the following spatial error model for regions $i = 1, 2, ..., N$:

$$\frac{d\tilde{\mathbf{x}}(t)}{dt} = \tilde{\mathbf{A}}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{B}}\tilde{\mathbf{u}}(t) + \tilde{\mathbf{\kappa}} + \frac{d\tilde{\mathbf{z}}(t)}{dt}, \tag{1}$$

$$\frac{d\tilde{\mathbf{z}}(t)}{dt} = \tilde{\mathbf{R}}\tilde{\mathbb{C}}\frac{d\tilde{\mathbf{z}}(t)}{dt} + \tilde{\mathbf{G}}\frac{d\tilde{\mathbf{W}}(t)}{dt}, \tag{2}$$

where $\tilde{\mathbf{x}}_{(Nn \times 1)} = \text{rowvec } \mathbf{X}_{(N \times n)}$ and $\tilde{\mathbf{u}}_{(Nr \times 1)} = \text{rowvec } \mathbf{U}_{(N \times r)}$ row-vectorize the data matrices $\mathbf{X}$ and $\mathbf{U}$ (for each region i there is a row with n values for the n state variables and a row with r values for the fixed input variables). Similarly for the random subject effects $\tilde{\mathbf{\kappa}}$. Furthermore, $\tilde{\mathbf{W}}$ has the same dimension $(Nn \times 1)$ as $\tilde{\mathbf{x}}$, $\tilde{\mathbf{A}} = \mathbf{I}_N \otimes \mathbf{A}$, $\tilde{\mathbf{B}} = \mathbf{I}_N \otimes \mathbf{B}$, $\tilde{\mathbf{G}} = \mathbf{I}_N \otimes \mathbf{G}$, where the drift matrix $\mathbf{A}$ contains the coefficients of the causal relationships among the state variables, $\mathbf{B}$ the coefficients of the effects of the fixed input variables on the state vector, and the lower triangular matrix $\mathbf{G}$ transforms the uncorrelated standard multivariate Wiener process $\mathbf{W}(t)$ with variance t at time t into a process with variances possibly $\neq t$ at time t and correlations between its elements possibly $\neq 0$.[8]

Matrix $\mathbb{C}$ is the spatial $(N \times N)$ connectivity matrix. For the multivariate case we specify the $(Nn \times Nn)$ matrix $\tilde{\mathbb{C}} = \mathbb{C} \otimes \mathbf{I}_n$. Associated with $\tilde{\mathbb{C}}$ are the spatial dependence parameters. In the general case of a different spatial dependence parameter for each state variable, we have the $(n \times n)$ spatial parameter matrix $\mathbf{R}$, which for the N subjects becomes $\tilde{\mathbf{R}} = \mathbf{I}_N \otimes \mathbf{R}$. In this paper, we assume one and the same spatial dependence parameter for the n state variables, that is, $\mathbf{R} = \rho \mathbf{I}_n$. This simplification safeguards the commutative property and has two advantages: (a) conventional standard procedures can be used to solve the stochastic differential equation implied by Equations (1) and (2); and (b) standard spatial econometric methods can be applied. Observe that Model (1)–(2) includes three parameter matrices to be estimated ($\mathbf{A}$, $\mathbf{B}$, and $\mathbf{G}$) in addition to the spatial parameter ρ.

From Equations (1) and (2), we derive:

$$\frac{d\tilde{\mathbf{x}}(t)}{dt} = \tilde{\mathbf{A}}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{B}}\tilde{\mathbf{u}}(t) + \tilde{\mathbf{\kappa}} + (\mathbf{I} - \tilde{\mathbf{R}}\tilde{\mathbb{C}})^{-1}\tilde{\mathbf{G}}\frac{d\tilde{\mathbf{W}}(t)}{dt}. \tag{3}$$

Equation (3) is solved over intervals $[t - \Delta t, t)$ of length Δt by:

---

[8] Observe that $d\mathbf{W}(t)/dt$ is not a stochastic process in the ordinary sense (see for example Arnold 1974) nor can $\int_{t_0}^{t} \mathbf{F}(s)d\mathbf{W}(s)$ with possibly time-varying function $\mathbf{F}(t)$ be defined as an ordinary Riemann-Stieltjes integral (Kuo 2006, pp. 9–10). The solution of Equations (1) and (2) nevertheless contains this nonstandard type of integral. However, as described in the references in this note, the problems of defining and handling the stochastic integral have been completely solved.

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{A}}_{\Delta t}\tilde{\mathbf{x}}(t-\Delta t) + \int_{t-\Delta t}^{t}\tilde{\mathbf{A}}_{t-s}\mathrm{d}s\,\tilde{\mathbf{B}}\tilde{\mathbf{u}}(t-\Delta t) + \int_{t-\Delta t}^{t}\tilde{\mathbf{A}}_{t-s}\mathrm{d}s\,\tilde{\mathbf{\kappa}}$$
$$+ (\mathbf{I}-\tilde{\mathbf{R}}\tilde{\mathbb{C}})^{-1}\int_{t-\Delta t}^{t}\tilde{\mathbf{A}}_{t-s}\tilde{\mathbf{G}}\mathrm{d}\tilde{\mathbf{W}}(s), \tag{4}$$

where $\tilde{\mathbf{A}}_{\Delta t} = \mathbf{I}_N \otimes \mathbf{A}_{\Delta t}$, $\mathbf{A}_{\Delta t} = \mathrm{e}^{\mathbf{A}\Delta t}$, $\tilde{\mathbf{A}}_{t-s} = \mathbf{I}_N \otimes \mathbf{A}_{t-s}$, $\mathbf{A}_{t-s} = \mathrm{e}^{\mathbf{A}(t-s)}$, and where for convenience sake, it is assumed that the input $\tilde{\mathbf{u}}(t)$ can be approximated by constants over the relevant intervals $[t-\Delta t, t)$ (for time-varying inputs, see Oud and Jansen 2000). Note that Δt and $(t-s)$ are subscripts in $\mathbf{A}_{\Delta t}$ and $\mathbf{A}_{t-s}$ but multiplied by $\mathbf{A}$ in the exponents of $\mathrm{e}^{\mathbf{A}\Delta t}$ and $\mathrm{e}^{\mathbf{A}(t-s)}$. Observe furthermore the important role of the matrix exponential $\mathbf{A}_{\Delta t} = \mathrm{e}^{\mathbf{A}[t-(t-\Delta t)]} = \mathrm{e}^{\mathbf{A}\Delta t}$, as well as the matrix exponential $\mathbf{A}_{t-s} = \mathrm{e}^{\mathbf{A}(t-s)}$, which appears three times inside the integrals. Particularly, $\mathrm{e}^{\mathbf{A}\Delta t}$ gives the effect of $\mathbf{x}(t-\Delta t)$ on $\mathbf{x}(t)$ over the interval Δt. Over this interval input, subject, and noise effects enter continuously at time points $s \in [t-\Delta t, t)$ with effects $\mathrm{e}^{\mathbf{A}(t-s)}$ on $\mathbf{x}(t)$. These effects (from successive time points s to t) must be 'summed' (via the integrals) to obtain the total effect. For an explicit expression of the integral $\int_{t-\Delta t}^{t}\mathbf{A}_{t-s}\mathrm{d}s$ we refer to Oud and Jansen (2000).

Formula (4) is based on the assumption that the state $\mathbf{x}(t-\Delta t)$ summarizes the entire history of the system up to $t-\Delta t$ at the single time point $(t-\Delta t)$ (Polderman and Willems 1998, pp. 115-149) such that the single matrix exponential $\mathrm{e}^{\mathbf{A}\Delta t}$ adequately captures the effects from the past over the interval Δt. The input, subject and noise effects after time point $(t-\Delta t)$, however, are exogenous and thus not captured by $\mathbf{x}(t-\Delta t)$. Therefore, their effects need to be summed over the interval Δt and separately added to the system.

We write Equation (4) in compact form as follows:

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{A}}_{\Delta t}\tilde{\mathbf{x}}(t-\Delta t) + \tilde{\mathbf{B}}_{\Delta t}\tilde{\mathbf{u}}(t-\Delta t) + \tilde{\mathbf{H}}_{\Delta t}\tilde{\mathbf{\kappa}} + (\mathbf{I}-\tilde{\mathbf{R}}\tilde{\mathbb{C}})^{-1}\tilde{\mathbf{w}}(t-\Delta t),$$
$$\text{where } \tilde{\mathbf{B}}_{\Delta t} = \int_{t-\Delta t}^{t}\tilde{\mathbf{A}}_{t-s}\mathrm{d}s\,\tilde{\mathbf{B}},$$
$$\tilde{\mathbf{H}}_{\Delta t} = \int_{t-\Delta t}^{t}\tilde{\mathbf{A}}_{t-s}\mathrm{d}s, \tag{5}$$
$$\text{and } \tilde{\mathbf{w}}(t-\Delta t) = \int_{t-\Delta t}^{t}\tilde{\mathbf{A}}_{t-s}\tilde{\mathbf{G}}\mathrm{d}\tilde{\mathbf{W}}(s).$$

For an explicit expression of the covariance matrix of $\tilde{\mathbf{w}}(t-\Delta t)$ we refer to Oud and Jansen (2000).

We now turn to estimation of the CT parameters on the basis of DT observation points $t_i \in \{t_1, ..., t_T\}$. For this purpose, a 'link' or 'tool' is needed between the CT model (5) and the observations at the DT observation points $t_i \in \{t_1, ..., t_T\}$. Several links are available in the literature including the Approximate Discrete Model (ADM) and the Exact Discrete Model (EDM). In this paper we apply the EDM which is preferable to alternatives like

7

the ADM and several other estimation procedures in the literature that merely approximate the CT parameter matrices (Singer 1990). The EDM imposes nonlinear restrictions on the DT parameters (which are obtained by solving the differential equation) to ensure that the parameters estimated are exactly equal to the parameters of the underlying differential equation model.

The EDM corresponding to Equation (5) reads as follows (Oud and Jansen 2000):

$$\tilde{\mathbf{x}}_{t_i} = \tilde{\mathbf{A}}_{\Delta t_i}\tilde{\mathbf{x}}_{t_i - \Delta t_i} + \tilde{\mathbf{B}}_{\Delta t_i}\tilde{\mathbf{u}}_{t_i - \Delta t_i} + \tilde{\mathbf{H}}_{\Delta t_i}\tilde{\boldsymbol{\kappa}} + (\mathbf{I} - \tilde{\mathbf{R}}\tilde{\mathbb{C}})^{-1}\tilde{\mathbf{w}}_{t_i - \Delta t_i}. \tag{6}$$

where the parameter matrices are as in Equation (5).

Equations (5) and (6) look very similar. However, whereas the former is a CT model defined for all t, the latter is a DT model defined for the discrete observation points $t_i \in \{t_1, ..., t_T\}$. The notation in Equation (6) is chosen to make clear that $t_i \in \{t_1, ..., t_T\}$ is a discrete observation time point in the continuous time interval. Observe that the CT parameter matrices in Equation (5) impose nonlinear restrictions on the DT parameter matrices in Equation (6) so that the latter exactly satisfy the CT model structure (1)–(2). The nonlinearity is apparent from the exponentials $\mathbf{A}_{\Delta t} = e^{\mathbf{A}\Delta t}$ and $\mathbf{A}_{t-s} = e^{\mathbf{A}(t-s)}$ in (5) and (6). Traditional linearly-oriented programs for estimating a Structural Equation Model (SEM) such as LISREL (Jöreskog and Sörbom 1996) do not provide the exponential function and so cannot be employed. However, EDM (6) and thus the CT parameters in (1)–(2) can be estimated by means of a nonlinear SEM (using, for example, the Mx software by Neale et al. 2003). We formulate the SEM model by first defining the state, input, and error vectors $\vec{\mathbf{x}}$, $\vec{\mathbf{u}}$, and $\vec{\mathbf{w}}$, for successive observation time points t_i.[9]

$$\vec{\mathbf{x}} = \left[\tilde{\mathbf{x}}_{t_0}', ..., \tilde{\mathbf{x}}_{t_{T-1}}'\right]',$$
$$\vec{\mathbf{u}} = \left[\tilde{\mathbf{u}}_{t_0}', ..., \tilde{\mathbf{u}}_{t_{T-1}}'\right]', \tag{7}$$
$$\vec{\mathbf{w}} = \left[\tilde{\mathbf{x}}_{t_0}' - E(\tilde{\mathbf{x}}_{t_0}'), \tilde{\mathbf{w}}_{t_0}', ..., \tilde{\mathbf{w}}_{t_{T-2}}'\right]'.$$

Next, we write Equation (6) over all observation time points in comprehensive form as:

$$\vec{\mathbf{x}} = \vec{B}\vec{\mathbf{x}} + \vec{\Gamma}_{\mathrm{u}}\vec{\mathbf{u}} + \vec{\Gamma}_{\kappa}\tilde{\boldsymbol{\kappa}} + (\mathbf{I} - \vec{\mathbf{R}}\mathbb{C})^{-1}\vec{\mathbf{w}}, \tag{8}$$

where we put all the $T-1$ ($Nn \times Nn$) matrices $\tilde{\mathbf{A}}_{\Delta t_i}$ at the appropriate places in the ($TNn \times TNn$) matrix $\vec{B}$, the $T-1$ ($Nn \times Nr$) matrices $\tilde{\mathbf{B}}_{\Delta t_i}$ in the ($TNn \times TNr$) matrix $\vec{\Gamma}_{\mathrm{u}}$, and the $T-1$ ($Nn \times Nn$) matrices $\tilde{\mathbf{H}}_{\Delta t_i}$ in the ($TNn \times Nn$) matrix $\vec{\Gamma}_{\kappa}$. The block-diagonal ($TNn \times TNn$) matrix $\bar{\mathbb{C}} = \mathbf{I}_T \otimes \tilde{\mathbb{C}}$ has (possibly asymmetric) blocks $\tilde{\mathbb{C}}$ on its

---

[9] To our best knowledge there are no dynamic structural equation models with spatial dependence. For a cross-sectional spatial structural equation model we refer to e.g. Gebremariam et al. (2010a, b) who apply Kelejian and Prucha's (2004) Feasible Generalized Three-Stage Least Squares Estimator.

diagonal. Because of the assumption $\mathbf{R} = \rho \mathbf{I}_n$, we can write $(\mathbf{I} - \vec{\mathbf{R}} \bar{\mathbb{C}})^{-1}$ as $(\mathbf{I} - \rho \bar{\mathbb{C}})^{-1}$ with T blocks $(\mathbf{I} - \rho \tilde{\mathbb{C}})^{-1}$ on its diagonal.[10]

Putting Equation (8) in terms of the spatially lagged variables $\vec{\mathbf{x}}_{\mathbb{C}} = \bar{\mathbb{C}} \vec{\mathbf{x}}$ and $\vec{\mathbf{u}}_{\mathbb{C}} = \bar{\mathbb{C}}_u \vec{\mathbf{u}}$ gives:

$$\vec{\mathbf{x}} = \bar{B} \vec{\mathbf{x}} + \rho(\mathbf{I} - \bar{B}) \vec{\mathbf{x}}_{\mathbb{C}} + \bar{\Gamma}_u \vec{\mathbf{u}} - \rho \bar{\Gamma}_u \vec{\mathbf{u}}_{\mathbb{C}} + \bar{\Gamma}_\kappa \tilde{\boldsymbol{\kappa}}_{\mathbb{C}} + \vec{\mathbf{w}}, \tag{9}$$

Observe that in the derivation of Equation (9) from Equation (8) we have made use of the equality $\bar{\mathbb{C}} \bar{\Gamma}_u \vec{\mathbf{u}} = \bar{\Gamma}_u \bar{\mathbb{C}}_u \vec{\mathbf{u}}$ for $\bar{\mathbb{C}}_u = \mathbf{I}_T \otimes \mathbb{C} \otimes \mathbf{I}_r$. Moreover, the transformed unobserved heterogeneity $\tilde{\boldsymbol{\kappa}}_{\mathbb{C}}$ is related to the original $\tilde{\boldsymbol{\kappa}}$ in Equation (6) as follows: $\tilde{\boldsymbol{\kappa}}_{\mathbb{C}} = (\mathbf{I} - \tilde{\mathbf{R}} \tilde{\mathbb{C}}) \tilde{\boldsymbol{\kappa}}$. Observe also that in the derivation of Equation (9) we have made use of the commutative property several times.

Equation (9) can be specified as a latent variables SEM as follows:

$$\vec{\boldsymbol{\eta}} = \bar{B} \vec{\boldsymbol{\eta}} + \bar{\Gamma} \vec{\boldsymbol{\xi}} + \vec{\boldsymbol{\zeta}}$$
$$\text{for} \tag{10}$$
$$\vec{\boldsymbol{\eta}} = \vec{\mathbf{x}}, \vec{\boldsymbol{\xi}} = \left[\vec{\mathbf{x}}_{\mathbb{C}}' \ \vec{\mathbf{u}}' \ \vec{\mathbf{u}}_{\mathbb{C}}' \ \tilde{\boldsymbol{\kappa}}_{\mathbb{C}}' \right]', \vec{\boldsymbol{\zeta}} = \vec{\mathbf{w}}, \bar{\Gamma} = \left[\rho(\mathbf{I} - \bar{B}) \ \ \bar{\Gamma}_u \ \ -\rho \bar{\Gamma}_u \ \ \bar{\Gamma}_\kappa \right],$$

which is conventionally written in variable form (rather than in terms of units of observation) as follows:

$$\boldsymbol{\eta} = B \boldsymbol{\eta} + \Gamma \boldsymbol{\xi} + \boldsymbol{\zeta}. \tag{11}$$

Further to Footnote 7 a SEM can contain both latent and observed variables.[11] If a SEM contains latent variables in the structural model, measurement equations are required which specify how the latent variables are measured, that is, how the observed variables $\mathbf{y}$ are related to the latent variables $[\boldsymbol{\eta}' \ \boldsymbol{\xi}']'$. If the variables in Equation (11) are latent, the measurement equation reads:

$$\mathbf{y} = \Lambda \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\xi} \end{bmatrix} + \boldsymbol{\varepsilon}. \tag{12}$$

Matrix Λ in Equation (12) contains the loadings, whereas the measurement errors are given by $\boldsymbol{\varepsilon}$ (with covariance matrix Θ). The measurement model parameter matrices Λ and Θ are estimated simultaneously with the other parameter matrices of Equation (11). For reasons of interpretation and identification, it is customary to specify unifactorial

---

[10] The notation in Equation (8) is a compromise that combines the standard notations in state-space modelling and structural-equation modelling. Although it would be possible to introduce a new notation, we prefer to apply the combined notation so as to facilitate access to the constituting literatures.

[11] A major advantage of SEM is that measurement errors can be straighforwardly taken into account to reduce attenuation bias (Oud and Folmer 2008).

observed variables only, which means that each observed variable in **y** has a loading on only one single latent variable in $[\mathbf{\eta}' \quad \mathbf{\xi}']'$.

The vector $\vec{\mathbf{y}}$ and its spatially lagged counterpart $\vec{\mathbf{y}}_{\mathbb{C}}$ are defined analogously to $\vec{\mathbf{x}}$ and $\vec{\mathbf{x}}_{\mathbb{C}}$. Therefore, we assume that their loading matrices $\vec{\mathbf{L}}$ (which are submatrices of the matrix $\mathbf{\Lambda}$ in Equation (12)) and measurement intercept vectors $\vec{\mathbf{d}}$ are equal for $\vec{\mathbf{y}}$ and $\vec{\mathbf{y}}_{\mathbb{C}}$. This gives the following measurement model for $\vec{\mathbf{x}}$ and $\vec{\mathbf{x}}_{\mathbb{C}}$.

$$
\begin{aligned}
\vec{\mathbf{y}} &= \vec{\mathbf{L}}\vec{\mathbf{x}} + \vec{\mathbf{d}} + \vec{\mathbf{v}}, \\
\vec{\mathbf{y}}_{\mathbb{C}} &= \vec{\mathbf{L}}\vec{\mathbf{x}}_{\mathbb{C}} + \vec{\mathbf{d}} + \vec{\mathbf{v}}_{\mathbb{C}}.
\end{aligned}
\tag{13}
$$

We impose no equality constraints between the measurement error variances of $\vec{\mathbf{v}}$ and $\vec{\mathbf{v}}_{\mathbb{C}}$, since the measurement errors in $\vec{\mathbf{v}}_{\mathbb{C}}$ are linear combinations of the measurements errors in $\vec{\mathbf{v}}$ and therefore typically have lower variance. In addition to $\vec{\mathbf{y}}$ and $\vec{\mathbf{y}}_{\mathbb{C}}$ the vector of observables includes the input vectors $\vec{\mathbf{u}}'$ and $\vec{\mathbf{u}}'_{\mathbb{C}}$. As they are fixed, their measurement model reduces to identity relationships, i.e. they get loadings equal to 1's in $\mathbf{\Lambda}$ and 0's in $\mathbf{\epsilon}$ in (12).

If there is (as in the present case study) only one observed variable in $\vec{\mathbf{y}}$ for a given latent variable in $\vec{\mathbf{x}}$, as well as one observed variable in $\vec{\mathbf{y}}_{\mathbb{C}}$ for a given latent variable in $\vec{\mathbf{x}}_{\mathbb{C}}$, then $\vec{\mathbf{y}}$ and $\vec{\mathbf{x}}$ as well $\vec{\mathbf{y}}_{\mathbb{C}}$ and $\vec{\mathbf{x}}_{\mathbb{C}}$ are equal, except for the measurement errors in $\vec{\mathbf{v}}$ and $\vec{\mathbf{v}}_{\mathbb{C}}$.[12] In such a case, $\vec{\mathbf{L}} = \mathbf{I}$ and $\vec{\mathbf{d}} = \mathbf{0}$, so that no loadings or measurement intercepts are estimated. This model can be rendered identified by specifying the measurement error variances to be equal for the repeated measurements of the same variable in $\vec{\mathbf{y}}$ as well as $\vec{\mathbf{y}}_{\mathbb{C}}$.[13]

Estimation of SEM models basically comes down to minimizing, in some metric, the distance between the theoretical covariance or moment matrix of the observed variables (as determined by the model specifications) and the corresponding sample matrix. Various estimators for SEMs exist, including maximum likelihood (ML). Oud and Folmer (2008) show that in the case of ML estimation the standard SEM likelihood function for a spatial dependency model needs to be augmented by the Jacobian correction term $\ln|\mathbf{I} - \rho\vec{\mathbb{C}}|$, where ln denotes the natural logarithm. The size of the Jacobian correction depends on the number of dependent variables. In a conventional spatial error model with only a single dependent variable and weights matrix $\mathbb{C}$, the correction is:

$$
\ln|\mathbf{I} - \rho\mathbb{C}|. \tag{14}
$$

---

[12] In the full model (13) (including both matrices $\vec{\mathbf{L}}$ and measurement intercept vectors $\vec{\mathbf{d}}$) more than one (unifactorial) indicator per latent variable may be specified.

[13] Of course, a measurement model would not be needed if there were only observed variables in the model, in which case we would have $\vec{\mathbf{y}} = \vec{\mathbf{x}}$ and $\vec{\mathbf{y}}_{\mathbb{C}} = \vec{\mathbf{x}}_{\mathbb{C}}$.

In a model with n dependent variables with equal spatial dependence parameter for the n variables with $(Nn \times Nn)$ matrix $\tilde{\mathbb{C}}$, the Jacobian correction is:

$$\ln | \mathbf{I} - \rho\tilde{\mathbb{C}} | = n \ln | \mathbf{I} - \rho\mathbb{C} |. \qquad (15)$$

In a longitudinal analysis with n dependent variables and T observations and $\bar{\mathbb{C}}$ of order $(TNn \times TNn)$, the correction is:

$$\ln | \mathbf{I} - \rho\bar{\mathbb{C}} | = Tn \ln | \mathbf{I} - \rho\mathbb{C} |. \qquad (16)$$

Finally, if each of the n latent dependent variables is measured by m indicators, a $(TNnm \times TNnm)$ matrix $\bar{\tilde{\mathbb{C}}}$ applies and the correction becomes (Oud and Folmer 2008):

$$\ln | \mathbf{I} - \rho\bar{\tilde{\mathbb{C}}} | = Tnm \ln | \mathbf{I} - \rho\mathbb{C} |. \qquad (17)$$

We observe that nonlinear SEM programs like Mx (Neale et al. 2003) can be applied to estimate the Equations (11) and (12), including all linear and nonlinear restrictions implied by both the CT and the spatial dependence specifications.

4. Illustration: Regional Unemployment and Population Change in Germany

The purpose of this section is to illustrate the CT modelling approach by estimating and interpreting a simple CT simultaneous equations model relating to the state variables regional unemployment change and regional population change, controlling for the effects of the input variables change of manufacturing work force and wage change on both state variables.[14] In the first equation, unemployment change is explained by population change, average wage change of fulltime workers, and change of manufacturing workforce. The expected effect of population change on unemployment change is ambiguous, since it has an impact on both labour supply (giving a positive impact on unemployment development) and labour demand (giving a negative impact on unemployment development). For the input variables, we expect a positive impact of wage change, since higher wages represent greater costs to firms (as, for example, in Hall (1972) and Layard et al. (1991)). For the change in manufacturing workforce (regional specialization), we expect a negative impact because of improved employment opportunities.

Population change is explained by the changes in unemployment, wages and manufacturing workforce. We expect higher unemployment to have a negative effect on population change, particularly because of an induced increase of outmigration. Since higher wages will tend to attract individuals towards a region (inmigration), we expect a positive wage change coefficient. For the change of manufacturing specialization, we

---

[14] It is possible to endogenize the latter two variables as well. See Oud et al. (2010).

expect a positive effect, though weak, on population change because of increased inmigration (see, for example, Budd et al. 1987).

To sum up, the model is made up of:

- The state variables unemployment change and population change. For each state variable, we expect an autoregressive effect. Moreover, we hypothesize a positive cross effect of population change on unemployment change (because we expect the labour supply effect to be stronger than the demand effect), and a negative cross effect from unemployment change to population change.
- The fixed input variables wage change and change of the manufacturing workforce. We assume a positive wage change effect and a negative manufacturing workforce change effect on unemployment change. For population change we assume positive effects of both exogenous variables.
- A first-order spatial lag for each of the state variables. We assume the spatial dependence parameter for the state variables to be equal such that $\tilde{\mathbf{R}}$ in Equation (2) contains only a single spatial parameter ρ.

The CT model is presented in Equation (18), where $ud(t)$ is unemployment change, $pd(t)$ population change, $wd(t)$ wage change, and $md(t)$ change of the manufacturing workforce:

$$
\begin{aligned}
\frac{\mathrm{d}ud(t)}{\mathrm{d}t} &= a_{11}ud(t) + a_{12}pd(t) + b_{11}wd(t) + b_{12}md(t) + b_1 + \kappa_1 + \frac{\mathrm{d}z_1(t)}{\mathrm{d}t}, \\
\frac{\mathrm{d}pd(t)}{\mathrm{d}t} &= a_{21}ud(t) + a_{22}pd(t) + b_{21}wd(t) + b_{22}md(t) + b_2 + \kappa_2 + \frac{\mathrm{d}z_2(t)}{\mathrm{d}t}.
\end{aligned}
\tag{18}
$$

Coefficients a_{11} and a_{22} represent the CT autoregressive effects of ud and pd, a_{12} and a_{21} the cross-effects of pd on ud and of ud on pd, respectively, whereas the effects of the input variables wd and md are given by b_{11} and b_{12} (on ud) and b_{21} and b_{22} (on pd). Finally, b_1 and b_2 are the intercepts, and κ_1 and κ_2 are the region-specific random intercepts. Because the random subject effects κ_1 and κ_2 represent deviations from the fixed intercepts b_1 and b_2, $\mathrm{E}(\kappa_1) = \mathrm{E}(\kappa_2) = 0$. The squared deviations show up in the model as the variances φ_{κ_1} and φ_{κ_2}; their covariance is $\varphi_{\kappa_1 \kappa_2}$.

It should be observed that the CT effects basically are the limits of the corresponding effects in DT for the observation interval going to zero. Due to the nonlinear relationship between CT and DT effects, the parameter values may considerably differ between CT and DT, however. Although the infinitesimal 'cross-effects' a_{12} and a_{21} in CT can be interpreted similarly to the corresponding cross-lagged effects $a_{12\Delta t}$ and $a_{21\Delta t}$ in DT, the values of a_{12} and a_{21} may differ considerably from $a_{12\Delta t}$ and $a_{21\Delta t}$; even the signs may change when going from the DT coefficients to the CT coefficients and vice versa. The reason is that the CT analysis also accounts for the autoregressive effects during the observation interval. In fact, the estimated cross-lagged DT effects are mixtures of the CT cross- and autoregressive effects. A variable with a high autoregressive CT effect, meaning that there is a strong tendency to sustain its value over time, also tends to retain the influence of other variables over a longer time interval than a variable with a low

autoregressive effect. So, even a relatively small CT cross-effect can result in a relatively high cross-lagged effect in DT, if the variable influenced has a high autoregressive effect. But the converse can also be true: a relatively strong CT cross-effect may have only a small impact over a DT interval because of a rather low autoregressive effect in the dependent variable.

The CT autoregressive coefficients (direct feedback-effects) a_{11} and a_{22} are to be interpreted differently from the corresponding DT autoregressive coefficients. Suppose the cross-effects are zero, then a CT autoregressive effect of 0 in $\mathbf{A}$ (no change) corresponds to a DT autoregression of 1 in $\mathbf{A}_{\Delta t}$, and a CT autoregressive effect of $-\infty$ in $\mathbf{A}$ (maximum negative feedback) to a DT autoregressive effect of 0 in $\mathbf{A}_{\Delta t}$. So, for zero cross-effects CT autoregressive effects in the range $(-\infty, 0)$ are transformed to DT autoregressive effects in the range $(0,1)$.

For the error components in Equation (18), Equation (2) applies, with spatial dependence parameter ρ and the parameters g_{11}, g_{22} and g_{21} in matrix $\mathbf{G} = \begin{pmatrix} g_{11} & 0 \\ g_{21} & g_{22} \end{pmatrix}$.

This matrix transforms the two independent standard Wiener processes

$$\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t} = \begin{pmatrix} \dfrac{\mathrm{d}W_1(t)}{\mathrm{d}t} \\ \dfrac{\mathrm{d}W_2(t)}{\mathrm{d}t} \end{pmatrix}$$

into correlated general Wiener processes for *ud* and *pd*.

Finally, we discuss the parameters relating to the initial time point when the process starts. First of all, there are the initial state means $\mu_{x_{1t_0}}$ and $\mu_{x_{2t_0}}$, their variances $\varphi_{x_{1t_0}}$, $\varphi_{x_{2t_0}}$, and covariance $\varphi_{x_{1t_0} x_{2t_0}}$. Because the initial means may differ in regions with different levels of *wd* and *md* at the initial point, one may regress $ud(t_0)$ and $pd(t_0)$ on $wd(t_0)$ and $md(t_0)$ (which gives the regression coefficients b_{11t_0}, b_{12t_0}, b_{21t_0} and b_{22t_0}), and subtract the regression means $[b_{11t_0} \overline{wd(t_0)} + b_{12t_0} \overline{md(t_0)}]$ and $[b_{21t_0} \overline{wd(t_0)} + b_{22t_0} \overline{md(t_0)}]$ to obtain the 'pure' initial means. Moreover, conditional on the initial inputs, the initial conditional variances and covariance of the state variables can be calculated. Finally, since the region specific random effects κ_1 and κ_2 are assumed to influence the state variables before as well as after initial time point t_0, the four covariances between the initial state variables and the random region specific effects $\varphi_{\kappa_1 x_{1t_0}}$, $\varphi_{\kappa_1 x_{2t_0}}$, $\varphi_{\kappa_2 x_{1t_0}}$ and $\varphi_{\kappa_2 x_{2t_0}}$ are estimated, since they cannot, in general, be taken to be zero.

4.1. *Empirical Analysis*

The empirical analysis relates to 439 German districts (NUTS-3 level of geographical aggregation, shown in Figure 4)[15] over the period 2000–03. All observed variables in the model (two endogenous (state) variables and two exogenous (input) variables) are defined as changes, that is, differences between successive years.[16] For convenience we divide the changes by 1,000. For instance, unemployment change in region *r* in 2000 is measured as (*1/1000*) * (unemployed in region *r* in 2000 – unemployed in region *r* in 1999).[17]

The nonlinear restrictions imposed by Equation (5) on the DT parameters in Equation (6) are implemented as follows. The DT parameter matrix $\mathbf{A}_{\Delta t_i}$ for successive intervals $\Delta t_i = 1$ is related to the matrix exponential as follows:

$$\mathbf{A}_{\Delta t_i} = \begin{bmatrix} a_{11\Delta t_i} & a_{12\Delta t_i} \\ a_{21\Delta t_i} & a_{22\Delta t_i} \end{bmatrix} = e^{\mathbf{A}\Delta t_i} = e^{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\Delta t_i} = \mathbf{M}e^{\mathbf{V}\Delta t_i}\mathbf{M}^{-1} = \mathbf{M}e^{\begin{bmatrix} v_1 & 0 \\ 0 & v_2 \end{bmatrix}\Delta t_i}\mathbf{M}^{-1}, \quad (19)$$

where $\mathbf{M}$ is eigenvector matrix and $\mathbf{V}$ the diagonal eigenvalue matrix of $\mathbf{A}$. The maximum likelihood procedure proceeds by inserting values for the CT parameters $a_{11}, a_{12}, a_{21}, a_{22}$ in $\mathbf{A}$ and for all other CT parameters which via (19) impose restrictions on the DT parameter matrix $\mathbf{A}_{\Delta t_i}$ and on other DT parameter matrices. The resulting theoretical (restricted) moment or covariance matrix is estimated by minimizing the distance between the latter and the observed moment or covariance which gives the maximum of the likelihood function. The Mx code used to estimate the present labour market model can be downloaded from http://www.socsci.kun.nl/~hano/.

To illustrate the advantage of SEM in reducing attenuation bias we estimate two model types: one without (I), and one with (II) measurement errors for the state variables and their spatially lagged counterparts. Compared to Model I, there are four additional parameters in Model II: the measurement error variances θ_{v1} and θ_{v12} for the observed state variables, and measurement error variances $\theta_{v1_{\mathbb{C}}}$ and $\theta_{v2_{\mathbb{C}}}$ for their spatially lagged counterparts (see Equation (13)).

Both Model I and Model II are estimated by ML. Since $T = 4$, $n = 2$ and $m = 1$ (a single indicator per latent variable), the Jacobian correction term added to the likelihood function in both models is $\ln|\mathbf{I} - \rho\tilde{\tilde{\mathbb{C}}}| = Tnm\ln|\mathbf{I} - \rho\mathbb{C}| = 8\ln|\mathbf{I} - \rho\mathbb{C}|$.

---

[15] For NUTS-3 units, spatial dependence is likely, due to, amongst others, commuting. The data used in this application were provided by the German Institute for Employment Research (*Institut für Arbeitsmarkt- und Berufsforschung*, IAB), and include the entire Germany, consisting of 326 districts in the former West Germany, and 113 in the former East Germany. They have previously been analysed by, amongst others, Patuelli et al. (2010).

[16] The variables in our model are point observations of changes for which the CT framework presented above applies. It applies also for mixed stock and flow variables (Bergstrom 1984; Singer 1992).

[17] All variables are of orders of magnitude of thousands. By dividing the changes by 1000 the magnitudes of the variances and covariances are reduced which facilitates the calculations without affecting the results.

Figure 4. Map of the 439 German NUTS-3 regions

The estimation results are given in Table 1. First of all, we refer to the spatial dependence parameter ρ, which is 0.375 in Model I and 0.378 in Model II. In both models, ρ is highly significant. Since it is rather restrictive, we relaxed the assumption of equal spatial dependence for *ud* and *pd* in the SEM model. This relaxation, however, did not lead to any significant improvement in model fit, as measured by the χ^2-difference test. We conclude that the assumption of the same spatial dependence parameter ρ for both *ud* and *pd* is not contradicted by the data.

Table 1. ML parameter estimates, associated z-values and (selected) goodness of fit statistics for model I (without measurement errors) and model II (with measurement errors)

Par.	Model I		Model II	
	Parameter estimate	z	Parameter estimate	z
Spatial parameter				
ρ	0.375	16.30*	0.378	14.00*
Measurement error variances				
θ_{v1}			0.422	16.88*
θ_{v2}			0.081	1.09
$\theta_{v1_{\mathbb{C}}}$			0.015	0.82
$\theta_{v2_{\mathbb{C}}}$			0.044	4.04*
State effects				
a_{11}	−1.664	−11.64*	−0.594	−6.83*
a_{12}	−0.009	−0.17	0.013	0.54
a_{21}	−0.252	−2.00*	−0.498	−4.64*
a_{22}	−1.363	−6.75*	−1.168	−5.01*
Input effects				
b_{11}	0.016	0.39	0.005	0.19
b_{12}	0.047	0.98	0.021	0.78
b_{21}	0.001	0.01	0.017	0.49
b_{22}	0.033	0.89	0.014	0.34
Fixed intercepts				
b_1	0.633	10.21*	0.463	11.87*
b_2	−0.136	−1.97*	−0.101	−1.65
Random intercept variance				
φ_{κ_2}	0.679	2.23*	0.439	1.64
Error parameters				
g_{11}	1.316	23.93*	0.156	1.61
g_{22}	1.196	19.93*	1.000	16.40*
g_{21}	−0.010	−0.18	−0.059	−0.31
Initial (reduced) state means				
$\mu_{x_{1_{t_0}}}$	−0.656	−13.39*	−0.647	−12.68*
$\mu_{x_{2_{t_0}}}$	−0.146	−2.24*	−0.148	−2.24*
Initial (conditional) state variances and covariance				
$\varphi_{x_{1_{t_0}}}$	0.419	14.96*	0.052	1.85
$\varphi_{x_{2_{t_0}}}$	1.270	14.76*	1.202	11.13*
$\varphi_{x_{1_{t_0}} x_{2_{t_0}}}$	−0.014	0.40	0.015	0.42

Par.	Model I		Model II	
	Parameter estimate	z	Parameter estimate	z
Initial time point regression coefficients				
b_{11t_0}	0.030	0.86	0.010	0.30
b_{12t_0}	−0.057	−1.63	−0.049	−1.44
b_{21t_0}	0.021	−0.51	−0.018	−0.42
b_{22t_0}	0.029	0.69	0.027	0.63
Covariances between random intercept and initial states				
$\varphi_{\kappa_2 x_{1t_0}}$	−0.085	−2.24[*]	−0.091	−2.52[*]
$\varphi_{\kappa_2 x_{2t_0}}$	0.940	5.25[*]	0.792	4.10[*]
Fit				
χ^2	770.2		650.7	
Df	206		202	
RMSEA	0.081		0.071	

[*] p-value < 0.05

From the significance of the measurement error variance of *ud* (0.422) and of the spatially lagged *pd* (0.044) it follows that Model II is more adequate than Model I. This conclusion is supported by the significant improvement of model fit when the assumption of no measurement errors is dropped, as shown by the difference test ($\chi^2_{\text{dif}} = 770.2 - 650.7 = 119.5$ for *df* = 206 – 202 = 4). Finally, the RMSEA fit measure of Model II (0.071) is smaller than for Model I (0.08), and meets the criterion of a 'reasonable' fit (Jöreskog and Sörbom 1996, p. 124). We conclude that Model II is preferable to Model I. Nevertheless, we initially consider both models below to illustrate the attenuation bias due to ignoring measurement error.

Table 2 presents the CT state effect matrices **A.** To facilitate interpretation we also present the effect matrices $\mathbf{A}_{\Delta t}$ for observation interval $\Delta t = 1$ derived from the CT state effect matrices.[18] Moreover, in Table 2 we present both the standardized (by the initial standard deviations of *ud* and *pd*) and unstandardized effects in **A** and $\mathbf{A}_{\Delta t}$ (in Table 1, only unstandardized effects are presented). Observe that unstandardized and standardized autoregression effects are equal, because the standard deviations in numerator and denominator are equal.[19]

Comparison of Models I and II in Table 2 illustrates the disattenuation effect due to explicitly accounting for measurement errors. With the exception of the standardized CT effect of unemployment change on population change (*ud* → *pd*) the coefficients in Model II are larger than the corresponding coefficients in Model I, indicating that the

---

[18] Observe that for intervals other than $\Delta t = 1$, the results would be totally different, as shown in Figure 5a,b.

[19] Standardization prevents the effects from being dependent on the measurement scale unit of the variables involved.

latter are subject to attenuation bias due to measurement errors. Therefore, for the remainder we restrict ourselves to Model II.

Table 2. CT state effect matrices **A** and derived DT state effect matrices $\mathbf{A}_{\Delta t} = e^{\mathbf{A}\Delta t}$ for Model I (without measurement errors) and Model II (with measurement errors) for a one-year interval ($\Delta t = 1$), in unstandardized and standardized form

Par.	Model I		Model II	
	A	$\mathbf{A}_{\Delta t=1}$	**A**	$\mathbf{A}_{\Delta t=1}$
Unstandardized				
ud	$a_{11} = -1.664^*$	$a_{11\Delta t=1} = 0.190$	$a_{11} = -0.594^*$	$a_{11\Delta t=1} = 0.550$
pd → ud	$a_{12} = -0.009$	$a_{12\Delta t=1} = -0.002$	$a_{12} = 0.013$	$a_{12\Delta t=1} = 0.006$
ud → pd	$a_{21} = -0.252^*$	$a_{21\Delta t=1} = -0.056$	$a_{21} = -0.498^*$	$a_{21\Delta t=1} = -0.209$
pd	$a_{22} = -1.363^*$	$a_{22\Delta t=1} = 0.256$	$a_{22} = -1.168^*$	$a_{22\Delta t=1} = 0.310$
Standardized				
ud	$a_{11} = -1.664^*$	$a_{11\Delta t=1} = 0.190$	$a_{11} = -0.594^*$	$a_{11\Delta t=1} = 0.550$
pd → ud	$a_{12} = -0.016$	$a_{12\Delta t=1} = -0.003$	$a_{12} = 0.063$	$a_{12\Delta t=1} = 0.027$
ud → pd	$a_{21} = -0.145^*$	$a_{21\Delta t=1} = -0.032$	$a_{21} = -0.104^*$	$a_{21\Delta t=1} = -0.044$
pd	$a_{22} = -1.363^*$	$a_{22\Delta t=1} = 0.256$	$a_{22} = -1.168^*$	$a_{22\Delta t=1} = 0.310$

* p-value < 0.05

From Table 1 and Table 2, it follows that, in accordance with our expectations (see Section 3), both *ud* and *pd* show substantial autoregressive effects (–0.594 and –1.168 in **A**; 0.550 and 0.310 in $\mathbf{A}_{\Delta t}$). In both models the coefficients a_{11} and a_{22} are negative and significant, implying that the models are stable. Moreover, as hypothesized, *ud* has a negative effect on *pd* (standardized value of –0.104), which is highly significant (see Table 1). The cross-effect of *pd* on *ud,* however, is not significant. This result is in line with the hypothesized dual population effect on unemployment in that supply factors (such as a potentially larger workforce) may be counterbalanced by demand factors (such as an increased demand for goods).

Table 1 shows that neither of the two input variables (wage change and change of the structure of the manufacturing sector) has a significant effect on either of the two state variables. The insignificant effect of wage change on unemployment change is counter-intuitive. A possible explanation is the rigidity of wage setting in Germany (wages are often set nationally rather than regionally or at firm level), such that regional wages insufficiently reflect the regional unemployment structure. Its insignificant effect on population change may be due to constraints on labour mobility (such as, for example, inefficiencies in the housing market). Longer time lags (and hence longer time series) might show significant effects.

The fixed intercept of *ud* is positive, substantial (0.463) and highly significant, whereas the fixed intercept of *pd* is negative (–0.101) and not significant at the 5 per cent level. The random intercept variance and covariances for κ_1 ($\varphi_{\kappa_1}, \varphi_{\kappa_1 x_{1_{t_0}}}, \varphi_{\kappa_1 x_{2_{t_0}}}$) were not significant and have been left out of the final analysis, since they considerably affect the

estimates of all other model parameters. The covariances of κ_2 with the initial state variables, $\varphi_{\kappa_2 x_{1t_0}}$ and $\varphi_{\kappa_2 x_{2t_0}}$, are significant which supports a random intercept for *pd* (variance $\varphi_{\kappa_2} = 0.439$). From these results it follows that regions resemble each other much more with regard to unemployment change than population change. Particularly, it implies that each region is following its own expected population change curve. For unemployment change on the other hand, the data support *one* single expected curve towards which each region in Germany regresses.

With regard to the error variances of the structural equations, we find that g_{11} and g_{21} are insignificant. However, we do not impose restrictions of the type $g_{11} = 0$ and $g_{21} = 0$, since it is unrealistic to assume that the model explains all the variance in *ud* and all the covariance between *ud* and *pd*.

The initial state variance of *ud* (0.052) is not significant. This outcome, together with the insignificance of its random intercept variance, means that regression for *ud* is not only towards the same expected curve for all regions, but also that the variance of the regions around this common expected curve is quite small. The initial state variance of *pd* (1.202), however, is significant meaning that from the start in 2000 the regions show clear differences in population change and, because of the random intercept variance, keep differences over time.

The initial means of the state variables over the 439 German local labour markets show that both unemployment and population decrease at the beginning of 2000 (–0.647[20] and –0.148, respectively). The initial mean unemployment change is much larger (in absolute value) and has a higher z-value than the initial mean population change ($z = -12.68$ versus $z = -2.24$). It should be noted that these means have been reduced by subtracting the (insignificant) regression means $[b_{11t_0} \overline{wd(t_0)} + b_{12t_0} \overline{md(t_0)}]$ and $[b_{21t_0} \overline{wd(t_0)} + b_{22t_0} \overline{md(t_0)}]$ from the initial means. The uncorrected initial means are even larger in absolute value (–0.975 for *ud* and –0.288 for *pd*).

4.2. Autoregression Functions, Cross-Lagged Effect Functions, and Means Trajectories

The estimates of Model II have been used to depict the autoregression functions of *ud* and *pd* (Figure 5a), the standardized cross-lagged effect functions of *ud* → *pd* and *pd* → *ud* (Figure 5b), and the means trajectories of *ud* and *pd* (Figure 5c) in CT. Because the model is asymptotically stable, both the autoregression functions and the cross-lagged effect functions go to zero.

Figure 5a shows the decay of the autoregressive effects of *ud* and *pd*. The decay is smaller for *pd* than for *ud*. For *pd*, the decay is approximately 70 per cent and for *ud* approximately 90 per cent after two years.

---

[20] Meaning a mean decrease of 647 unemployed persons per local labour market over the previous year.
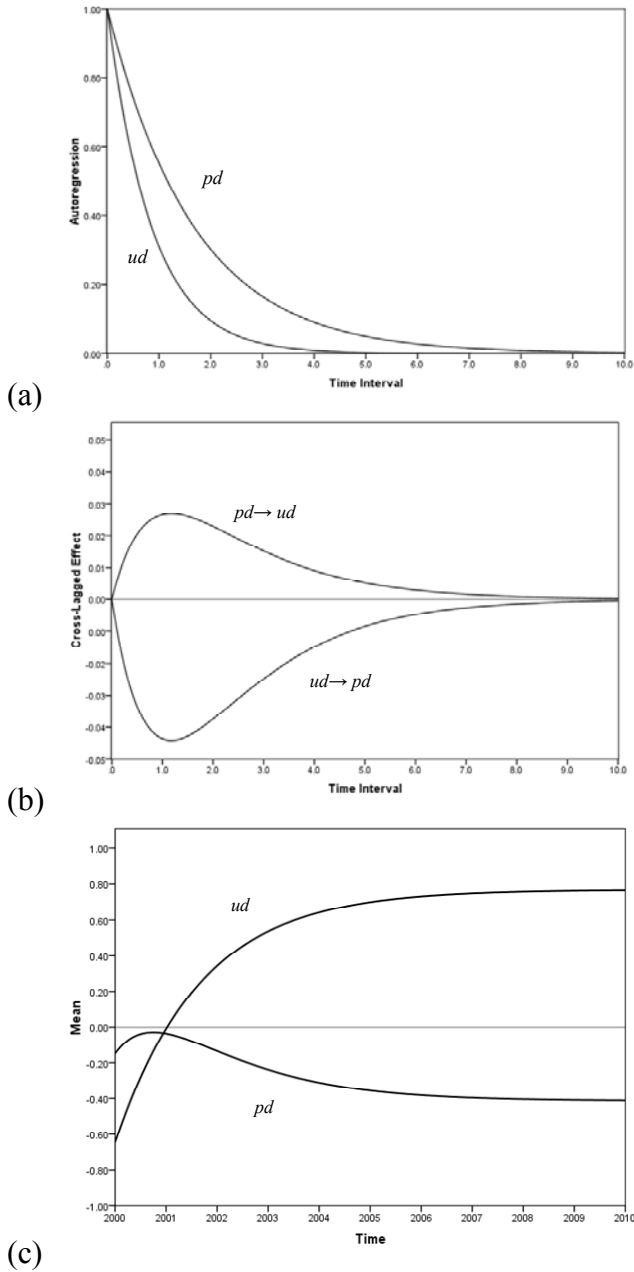
(a)

(b)

(c)

Figure 5. Autoregression functions (a); standardized cross-lagged effect functions (b);
Means trajectories (c)[21]

The standardized cross-lagged effect functions in Figure 5b show the effects in terms of standard deviation units[22] of the dependent variable from a standard deviation unit increase in the explanatory variable. The cross-lagged effect functions start from zero, then reach a maximum ($pd \rightarrow ud$), and a minimum ($ud \rightarrow pd$), respectively, and finally die out towards zero. In both directions ($ud \rightarrow pd$ and $pd \rightarrow ud$), the effects are rather

[21] All functions based on Model II parameters.
[22] Standard deviations are over regions.

small and in both cases the extreme values are reached after 1.15 years. Figure 5b shows that, at the extremes a standard deviation increase in unemployment change diminishes population change by 0.044 standard deviations while a standard deviation increase in population change increases employment change by 0.027 standard deviations. Observe that the effects die out rather slowly: in both cases, after four years, more than a quarter of the maximum impact is left.

Figure 5c depicts the autonomous developments of the means of *ud* and *pd* independent from input effects. They are given by (see Oud and Jansen 2000):

$$E[\mathbf{x}(t)] = e^{\mathbf{A}(t-t_0)} E[\mathbf{x}(t_0)] + \mathbf{A}^{-1}[e^{\mathbf{A}(t-t_0)} - \mathbf{I}]\mathbf{b}, \tag{20}$$

where **b** includes the fixed intercepts. Equation (20) shows that the mean development is driven by two components: the autoregression effect of the initial means $E[\mathbf{x}(t_0)]$ and the integrated effect of the intercepts **b** over the entire time period considered. Model II estimates are used for the calculation of **A**, **b**, and **x**(t_0) in Equation (20). Figure 5c shows that over the observation period 2000–03, shortly after 2001, the mean unemployment decrease turned into an unemployment increase which starts levelling off after 2003. The mean population decrease diminished until shortly before 2001, and then the downward trend increased again. Both trajectories tend to a stable equilibrium. This stable equilibrium implies for *ud* a mean unemployment increase of 770.1 per region and, for *pd*, a mean population decrease of 414.4.

5. Conclusions

In this paper we have incorporated spatial dependence in a continuous-time (CT) modelling framework, as introduced in econometrics in the 1950s by, amongst others, Koopmans (1950) and Phillips (1959), and in sociology by Simon (Simon 1952) and Coleman (1968). For this purpose we have combined the CT modelling approach developed by Oud and Jansen (2000) with the spatial dependence approach by Oud and Folmer (2008).

We have shown that CT modelling, in contrast to discrete-time (DT) modelling, takes explicitly into account the continuous nature of (socioeconomic) processes. Whereas DT modelling estimates a model made up of difference equations on the basis of observations at discrete points in time, CT modelling estimates a model made up of differential equations on the basis of the same observations; however, under restrictions derived from the underlying CT model. The model thus estimated allows to 'fill the gaps between observation points' by means of model-based interpolations and extrapolations. The main advantages compared to DT modelling are that CT models are independent of the observations interval and thus provide a common basis for accurate comparison of differently time-spaced models of the same process. Accordingly, predictions and simulations can be made for any time interval rather than for the observation points, as in DT modelling. This feature is especially relevant for the analysis and comparison of dynamic adjustment processes.

We have estimated the Exact Discrete Model (EDM) that links the CT model to the observations at discrete time points in a Structural Equation Model (SEM). A SEM

allows simultaneous handling of both latent (e.g. socioeconomic status or utility) and observable variables (e.g. income) in one model framework. Whereas the latter have direct empirical meaning, latent variables can only be observed via observables. The use of latent variables does not only allow a closer correspondence between theory (which usually contains latent variables) and empirics, but also makes it possible to reduce multicollinearity and attenuation bias (Folmer and Oud 2008; Folmer 2009).

We have extensively illustrated the CT-SEM model framework including spatial dependence by means of a data set of 439 German districts (NUTS-3 level of geographical aggregation) over the period 2000–03. The simultaneous equation model presents the reciprocal effects of regional unemployment change and regional population change, taking into account the effects of the input variables change of manufacturing workforce and wage change on both state variables. In a follow-up study the CT labour market model used here for illustrative purposes will be the core objective, and the model will be extended to include longer time periods. Moreover, the exogenous input variables wage change and change of economic structure will be made endogenous, in particular to test the responsiveness of wages to unemployment.

References

Arnold L. (1974) *Stochastic Differential Equations*. Wiley: New York

Bartlett M.S. (1946) On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series. *Journal of the Royal Statistical Society* 8 (1), 28-41

Bergstrom A.R. (1984) Continuous Time Stochastic Models and Issues of Aggregation over Time. In Z. Griliches and M.D. Intriligator (eds), *Handbook of Econometrics* (Vol. 2). North-Holland: Amsterdam, pp. 1145-212

Bergstrom A.R. (1988) The History of Continuous-Time Econometric Models. *Econometric Theory* 4 (3), 365-83

Budd P., R. Levine and P.N. Smith (1987) Long-term Unemployment and the Unemployment-Vacancies Relationship. *The Economic Journal* 98 (363), 1071-91

Coleman J.S. (1968) The Mathematical Study of Change. In H.M. Blalock and A. Blalock (eds), *Methodology in social research*. McGraw-Hill: New York, pp. 428-78

Commandeur J.J.F., S.J. Koopman and K. van Montfort (2010) State Space Methods for Latent Trajectory and Parameter Estimation by Maximum Likelihood. In K. van

Montfort, J.H.L. Oud and A. Satorra (eds), *Longitudinal Research with Latent Variables*. Springer-Verlag: Berlin Heidelberg, pp. 177-99

Cressie N.A.C. (1993) *Statistics for Spatial Data*. Wiley: New York

Delsing M.J.M.H. and J.H.L. Oud (2008) Analyzing Reciprocal Relationships by means of the Continuous-Time Autoregressive Latent Trajectory Model. *Statistica Neerlandica* 62, 58-82

Donaghy K.P. (2001) Solution and Econometric Estimation of Spatial Dynamic Models in Continuous Space and Continuous Time. *Journal of Geographical Systems* 3 (3), 257-70

Donaghy K.P. and M. Plotnikova (2003) Econometric Estimation of a Spatial Dynamic Model in Continuous Space and Continuous Time: An Empirical Demonstration. In A. Getis, J. Mur and H.G. Zoller (eds), *Spatial econometrics and spatial statistics* (Vol. 89-104). Palgrave Macmillan: Basingstoke

Folmer H. (2009) Why Sociology is Better Conditioned to Explain Economic Behaviour than Economics. *Kyklos* 62 (2), 258-74

Folmer H. and J. Oud (2008) How to Get Rid of W? A Latent Variables Approach to Modelling Spatially Lagged Variables. *Environment and Planning A* 40 (10), 2526-38

Hall R.E. (1972) Turnover in the Labour Force. *Brookings Papers on Economic Activity* 3, 709-64

Jöreskog K.G. and D. Sörbom (1996) *LISREL 8: User's Reference Guide*. Scientific Software International: Chicago

Kelejian H.H. and I.R. Prucha (2004) Estimation of Simultaneous Systems of Spatially Interrelated cross Sectional Equations. *Journal of Econometrics* 118 (1-2), 27-50

Koopmans T.C. (1950) Models Involving a Continuous Time Variable. In T.C. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*. John Wiley & Sons: New York, pp. 384–92

Kuo H.-H. (2006) *Introduction to Stochastic Integration*. Springer: New York

Layard P.R.G., S.J. Nickell and R. Jackman (1991) *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford University Press: Oxford

Neale M.C., G. Boker, G. Xie and H.M. Maes (2003) *Mx: Statistical Modeling* (6th ed.). Department of Psychiatry: Richmond

Oud J.H.L. (2001) Quasi-Longitudinal Designs in SEM State Space Modeling. *Statistica Neerlandica* 55, 200-20

Oud J.H.L. and R.A.R.G. Jansen (2000) Continuous Time State Space Modeling of Panel Data by Means of SEM. *Psychometrika* 65 (2), 199-215

Oud J.H.L. and H. Singer (2008) Continuous Time Modeling of Panel Data: SEM Versus Filter Techniques. *Statistica Neerlandica* 62 (1), 4-28

Oud J.H.L. and H. Folmer (2008) A Structural Equation Approach to Models with Spatial Dependence. *Geographical Analysis* 40 (2), 152-66

Oud J.H.L., H. Folmer, R. Patuelli and P. Nijkamp (2010) 'Endogenous and Exogenous Wages in German Regional Labour Markets: A Continuous-Time, Spatial-Dependence Panel Approach'. Paper presented at the 50th European Congress of the Regional Science Association International, Jönköping

Patuelli R., D.A. Griffith, M. Tiefelsdorf and P. Nijkamp (2010) Spatial Filtering and Eigenvector Stability: Space-Time Models for German Unemployment Data. *International Regional Science Review* (forthcoming)

Phillips R. (1959) Dissipative Operators and Hyperbolic Systems of Partial Differential Equations. *Transactions of the American Mathematical Society* 90 (2), 193-254

Piras G., K.P. Donaghy and G. Arbia (2007) Nonlinear Regional Economic Dynamics. *Journal of Geographical Systems* 9, 311-44

Polderman J.W. and J.C. Willems (1998) *Introduction to Mathematical Systems Theory: A Behavioral Approach*. Springer-Verlag: New York Berlin Heidelberg

Puu T. (1997) *Mathematical Location and Land Use Theory: An Introduction*. Springer: Berlin New York

Simon H.A. (1952) A Formal Theory of Interaction in Small Groups. *American Sociological Review* 17, 202-11

Singer H. (1990) *Parameter Schätzung in Zeitkontinuierlichen Dynamischen Systemen [Parameter Estimation in Continuous-Time Dynamic Systems]*. Hartung-Gorre: Konstanz

Singer H. (1992) *Zeitkontinuierliche Dynamische Systeme*. University of Frankfurt: Frankfurt

Wackernagel H. (1995) *Multivariate Geostatistics: An Introduction with Applications*. Springer: Berlin New York