university of groningen / faculty of economics and business

13006-EEF

State transfers at different moments in time: A spatial probit approach

Paul Elhorst
Pim Heijnen
Anna Samarina
Jan Jacobs

SOM is the research institute of the Faculty of Economics & Business at the University of Groningen. SOM has six programmes:

- Economics, Econometrics and Finance
- Global Economics & Management
- Human Resource Management & Organizational Behaviour
- Innovation & Organization
- Marketing
- Operations Management & Operations Research

Research Institute SOM
Faculty of Economics & Business
University of Groningen

Visiting address:
Nettelbosje 2
9747 AE Groningen
The Netherlands

Postal address:
P.O. Box 800
9700 AV Groningen
The Netherlands

T +31 50 363 7068/3815

www.rug.nl/feb/research

State transfers at different moments in time: A spatial probit approach

Paul Elhorst
University of Groningen
j.p.elhorst@rug.nl

Pim Heijnen
University of Groningen

Anna Samarina
University of Groningen

Jan Jacobs
University of Groningen
CAMA, Australian National University and CIRANO

State transfers at different moments in time: a spatial probit approach

J. Paul Elhorst[a,*]   Pim Heijnen[a]   Anna Samarina[a]
Jan P.A.M. Jacobs[a,b]

[a] University of Groningen
[b] CAMA, Australian National University and CIRANO

This version: June 2013

Abstract

This paper adopts a spatial probit approach to explain interaction effects among geographical units when the dependent variable takes the form of a binary response variable and state transfers occur at different moments in time. The model has two spatially lagged variables, one for units that are still in state 0 and one for units that already transferred to state 1. The parameters are estimated on observations for those units that are still in state 0 at the start of the different time periods, whereas observations on units after they transferred to state 1, just as in the literature on duration modeling, are discarded. Consequently, neighboring units that did not yet transfer may have a different impact than units that already transferred. We illustrate our approach with an empirical study of the adoption of inflation targeting for a sample of 58 countries over the period 1985–2008.

Keywords: spatial probit model, state transition, inflation targeting

JEL Classification: C23, C25, E52

*Corresponding author: Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. Phone: +31 50 363 3893; Fax: +31 50 363 7337; Email: `j.p.elhorst@rug.nl`.

1 Introduction

Spatial binary response models have seen increasing use in the spatial econometrics literature. This holds especially for the spatial probit model based on the normal distribution, on which we focus in this paper. Spatial probit may be used to explain interaction effects among geographical units when the dependent variable takes the form of a binary response variable. However, one shortcoming of this model is that it cannot be fruitfully used to explain the transition from one state to another when this transition for one geographical unit takes place at a different moment in time than for another unit.

Suppose there are two states, 0 and 1, that y_i denotes the state a particular unit i ($i = 1, \ldots, N$) is in, and that unit i transfers from state 0 to state 1 at time t_i ($t = 1, \ldots, T$). We are interested in the determinants of transfer from state 0 to state 1. This paper proposes a spatial probit model with two spatially lagged variables, one for units that are still in state 0 and one for units that already transferred to state 1. The parameters of this model will be estimated based on observations of those units that are still in state 0 at the start of the different time periods being considered; observations on units after they transferred to state 1 are removed. The dependent variable and the first spatial term are both specified in terms of unobserved choices, i.e., the propensity towards state 1, while the second spatial term is specified in terms of observed choices, i.e., the actual outcomes. Consequently, we allow neighboring units that did not yet transfer to have a different impact than units that already transferred.

Our setting differs from LeSage et al. (2011), who investigate the decision of firms in New Orleans to reopen their stores dependent on the decision made by other firms 0-3 months, 0-6 months, and 0-12 months in the aftermath of Hurricane Katrina. However, since they use their data in cross-section rather than splitting up the sample into different time periods, they implicitly assume that the transition to state 1 of all firms that reopened their stores took place at the same point in time. As a result, they cannot answer the question why some firms reopened their stores earlier than others and which role the interaction among firms at different points in time played in this transition process.

Mukherjee and Singer (2008) analyze the decision of 78 countries to adopt a monetary policy strategy known as inflation targeting dependent on the decision taken by other countries using time-series cross-section data over the period 1987–2003. Due to the fact that the coefficient of this interaction term is positive, the probability that a country will transfer to state 1 increases if other countries have preceded. However, by just pooling cross-sectional data over time, they implicitly assume that the period that has expired since a neighboring country has taken a positive decision, has no impact. In addition, they assume that neighboring countries that did not take a positive decision yet, have the same impact as countries that already adopted inflation targeting.

These and related issues have been widely discussed in the literature on duration modeling (see Cameron and Trivedi, 2005, Chapter 17 for an excellent overview). Generally, duration models are used to explain the time τ_i that has passed to the moment when unit i transfers from state 0 to state

1.[1] This literature has produced two results that are relevant to our study. Firstly, if the data are observed in discrete time intervals, one can use a discrete-time transition model, since in each time interval two outcomes are possible: the transition takes place or it does not (Cameron and Trivedi, 2005, p. 602). A probit model based on the normal distribution function which restricts the coefficients of the regressors to be constant over time, except for the intercept, is then a straightforward and legitimate choice. Secondly, observations on units after they transferred to state 1 are generally removed from the sample. This is because explanatory variables that change over time may exhibit feedback and hence may not be strictly exogenous; once a unit has transferred to state 1, the explanatory variables may change as a result of this transition.

The standard probit model as suggested in Cameron and Trivedi (2005) for duration data is not appropriate for our setting since individual units are treated as independent entities in duration models. Interaction effects result in additional complications. In duration models the process that is observed may have begun at different points in calendar time for different units in the sample. In other words, the models consider relative time rather than calendar time. In our setting not only the time that has passed before units transfer to state 1 is important, but also the time that has passed relative to the transfer of other units. Therefore, the transfer process can only be modeled adequately if the starting point of the observation period is the same for every unit in the sample.

The paper is structured as follows. Section 2 summarizes the literature

[1]A well-known example is the Cox proportional hazard model (Cox, 1972).

on the basic spatial probit model, its extensions and estimation issues. A detailed description of our model is provided in Section 3. Section 4 illustrates our approach with an empirical study of the adoption of inflation targeting for a sample of 58 countries over the period 1985–2008. Section 5 concludes.

2 Spatial probit models: a review

The basic spatial probit model

The basic spatial probit model is a linear regression model with spatially correlated error terms ε_i for a cross-section of N observations ($i = 1, \ldots, N$). In vector notation, the spatial error probit model reads as

$$\boldsymbol{Y}^* = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \ \boldsymbol{\varepsilon} = \lambda \boldsymbol{W\varepsilon} + \boldsymbol{v}, \tag{1}$$

where $\boldsymbol{Y}^*$ is an $N \times 1$ vector consisting of one observation on the latent dependent variable for every unit in the sample, and $\boldsymbol{X}$ is an $N \times K$ matrix of explanatory variables with parameters contained in a $K \times 1$ vector $\boldsymbol{\beta}$. $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)'$ and $\boldsymbol{v} = (v_1, \ldots, v_N)'$ represent the error terms of the model; $\boldsymbol{\varepsilon}$ reflects the spatially correlated error term with coefficient λ, while $\boldsymbol{v}$ follows a multivariate normal distribution with mean $\mathbf{0}$ and variance $\boldsymbol{I}$. We use $\boldsymbol{I}$ rather than $\boldsymbol{\sigma}^2\boldsymbol{I}$ here since $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ cannot be separately identified. For this reason, σ^2 is set to 1. $\boldsymbol{W}$ is an $N \times N$ pre-specified non-negative spatial weights matrix describing whether or not the spatial units in the sample are neighbors of each other. Its diagonal elements are zero, since no unit can be viewed as its own neighbor. Usually, the spatial weights matrix is normalized such that the elements of each row sum to one. The spatial error model is consistent with a situation where determinants of the dependent variable omitted from the model are spatially autocorrelated, and with a situation where unobserved shocks follow a spatial pattern. The spatial error probit model in (1) can be rewritten as

$$\boldsymbol{Y}^* = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} = \boldsymbol{X\beta} + (\boldsymbol{I} - \lambda \boldsymbol{W})^{-1}\boldsymbol{v}, \tag{2}$$

which implies that the covariance matrix of $\boldsymbol{\varepsilon}$ is $\boldsymbol{\Omega}_\lambda = [(\boldsymbol{I}-\lambda\boldsymbol{W})'(\boldsymbol{I}-\lambda\boldsymbol{W})]^{-1}$.

The basic problem that needs to be solved in estimating this model is that the likelihood function cannot be written as the product of N one-dimensional normal probabilities as is the case with the standard (non-spatial) probit model. This is because the individual error terms ε_i ($i = 1, \ldots, N$) are dependent on each other, as a result of which the likelihood function

$$L(\boldsymbol{\beta}, \lambda) = \int_{\boldsymbol{Y}^*} \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Omega}_\lambda|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}'\boldsymbol{\Omega}_\lambda^{-1}\boldsymbol{\varepsilon}\right\} d\boldsymbol{\varepsilon}, \tag{3}$$

is an N-dimensional integral.

Another problem might be the inversion of the matrix $(\boldsymbol{I} - \lambda\boldsymbol{W})$ for large values of N when using a numerical algorithm to find the optimum of λ, especially if this inversion needs to be repeated several times. This is because the number of steps most practical algorithms require to determine the inverse of an $N \times N$ matrix is proportional to N^3. Nevertheless, for small or moderate values of N (<1000), as in most empirical studies, this is not really a problem. The spatial error probit model has mainly been used to present solutions to these methodological problems, see McMillen (1992), Pinkse and Slade (1998), LeSage (2000), Beron and Vijverberg (2004), Fleming (2004), Klier and McMillen (2008), but it has rarely been used in empirical applications. One exception is Pinkse and Slade (1998), who adopt this model to explain the location choice of certain economic activities.[2]

The spatial lag probit model

Another popular spatial probit model is the spatial lag probit model: a linear regression model with endogenous interaction effects among the unobserved dependent variable

$$\boldsymbol{Y}^* = \rho \boldsymbol{W} \boldsymbol{Y}^* + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{v}, \tag{4}$$

where ρ represents the spatial autoregressive coefficient. Endogenous interaction effects are typically considered as the formal specification for the equilibrium outcome of a spatial or social interaction process, in which the value of the dependent variable for one agent is jointly determined with that of neighboring agents. By rewriting the spatial lag probit model as

$$\boldsymbol{Y}^* = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{v} = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{5}$$

it can be seen that the covariance matrix of $\boldsymbol{\varepsilon}$ in this model is similar to that of the spatial error probit model, $\boldsymbol{\Omega}_\rho = [(\boldsymbol{I} - \rho \boldsymbol{W})'(\boldsymbol{I} - \rho \boldsymbol{W})]^{-1}$, the difference

[2] Partly because of its mathematical convenience, Bolduc et al. (1997) replace probit by the logit specification in their empirical application; whereas the probability $P(y_1 = 1)$ has an analytical solution when adopting the logit specification, the probit specification has not.

being that the parameter λ is replaced by ρ. To estimate this model, not only the integration of N-dimensional integral needs to be accounted for, but also the endogeneity of the variable $\boldsymbol{WY}^*$. Many studies have considered this model from a methodological viewpoint: McMillen (1992), LeSage (2000), Beron and Vijverberg (2004), Fleming (2004), Klier and McMillen (2008), LeSage and Pace (2009, Chapter 10), Franzese Jr. and Hays (2010), Smirnov (2010), Pace and LeSage (2012). In contrast to the spatial error probit model, it has also been used in many empirical studies, among which, Beron et al. (2003), Mukherjee and Singer (2008), and LeSage et al. (2011).[3]

An important variant of the spatial lag probit model for the analysis to be conducted in this paper is

$$\boldsymbol{Y}^* = \rho \boldsymbol{WY} + \boldsymbol{X\beta} + \boldsymbol{v}, \tag{6}$$

where the latent dependent variable $\boldsymbol{Y}^*$ depends on observed choices represented by $\boldsymbol{WY}$ rather than unobserved ones. The only application of this model we could find in the spatial econometrics literature is Qu and Lee (2012). They derive LM tests for spatial correlation in a standard probit model not only if the alternative model is Equation (4) but also if the alternative model is Equation (6). In addition, they consider Tobit models. Soetevent and Kooreman (2007) apply Equation (6) to analyze social interactions in high school teen behavior. They assume that the unobserved choice of individual i depends on the observed choices of other individuals. One of the basic problems of this interaction model is that it does not have

[3] Klier and McMillen (2008) replace the probit by the logit specification.

a unique equilibrium, but different equilibria depending on the sign of the interaction parameter ρ and on the sample size N (Soetevent and Kooreman, 2007, Propositions 2 and 3). To estimate the model they assume that the probability that one particular equilibrium occurs is equal to one over the total number of equilibria.

Extensions of the spatial probit model

The spatial probit model has been extended further in several ways. LeSage and Pace (2009) present a spatial probit model where more than two alternatives are observed that can be ordered. Bhat et al. (2010) deal with ordered-response models in general, among which probit and logit. Wang and Kockelman (2009) construct a dynamic ordered spatial error probit model. Apart from spatial correlation, they add a (latent) dependent variable lagged in time to control for temporal dependence in the data. Flores-Lagunes and Schnier (2012) extend the spatial error probit model to a so-called Tobit type II model, i.e., they first transform the dependent variable into a binary variable and explain this variable by a spatial error probit model; then they explain the magnitude of the dependent variable by a regular spatial error model for only those units that are in state 1. Some studies also deal with heteroskedasticity (McMillen, 1992, LeSage, 2000, Fleming, 2004).[4] Although ignoring heteroskedasticity, when present, will result in inefficient parameter estimates, the estimators are still consistent. Therefore, the basic problem

[4]Generally, the diagonal elements of the variance-covariance matrix $\boldsymbol{\Omega}_\lambda = [(\boldsymbol{I} - \lambda\boldsymbol{W})'(\boldsymbol{I} - \lambda\boldsymbol{W})]^{-1}$ are not equal to each other. Some studies characterize this as heteroskedasticity too (Pinkse and Slade, 1998, Klier and McMillen, 2008). However, this type of heteroskedasticity is explicitly taken into account in the estimation of the spatial lag and the spatial error model.

that needs to be solved, the multidimensional integration problem, remains unaltered. It is perhaps for this reason that heteroskedasticity received less attention in later work.

Estimation

The expectation-maximization (EM) algorithm adapted by McMillen (1992) for the spatial probit model is one of the earliest attempts to deal with the multidimensional integration problem. The E-step takes the expectation of the log-likelihood function for the latent variable y_i^* conditional on its observed value y_i and the parameter vector. The initial parameter vector is obtained by estimating the spatial model as if the dependent variable is continuous, while subsequent values are obtained from the previous iteration. The M-step maximizes the likelihood function for the parameter vector conditional on the expected value of y_i obtained from the E-step, which boils down to estimating a regular spatial model for a continuous variable. These steps are then repeated until the parameter vector converges. This algorithm, however, has been severely criticized. Firstly, there is a substantial computational burden in the repetitions of the algorithm (Fleming, 2004). Both the EM algorithm and the maximization of the regular spatial model in each M-step requires an iterative two-stage procedure. Secondly, it does not produce an estimate of the variance-covariance matrix needed to determine the standard errors and t-values of the parameter estimates (LeSage, 2000, Fleming, 2004, Smirnov, 2010). It should be stressed that this is because of another important methodological shortcoming that has not been discussed in the literature before. Whereas the expectation of the latent variable y_i^* in the EM algorithm is determined conditional on the observed value y_i of the unit itself, it should be determined conditional on the observed values of *all* other units. Consequently, this algorithm produces inconsistent parameter estimates.

A similar type of problem applies to the Bayesian MCMC estimation procedure initially developed by LeSage (2000). This procedure is based on sequentially drawing model parameters from their conditional distributions. This process of sampling parameters continues until the distribution of draws converges to the targeted joint posterior distribution of the model parameters. Two different sampling schemes are used: the Gibbs sampler for model parameters that have standard conditional distributions ($\boldsymbol{\beta}$, $\boldsymbol{Y}^*$), and the Metropolis-Hastings sampler for the spatial parameter λ in the spatial error model or ρ in the spatial lag model, both of which have a non-standard distribution (LeSage and Pace, 2009, Chapter 7). The key problem is to sample $\boldsymbol{Y}^*$. In LeSage (2000), the individual elements of $\boldsymbol{Y}^*$ are obtained by sampling from a sequence of univariate truncated normal distributions. In later work, LeSage and Pace (2009, p. 285) point out that "this *cannot* be done for the case of a truncated multivariate distribution" (emphasis in original). Draws for individual elements y_i^* should be based on the distribution of y_i^* conditional on all other $N-1$ elements $[y_1^*, \cdots, y_{i-1}^*, y_{i+1}^*, \cdots, y_N^*]$. Probably because James LeSage has made a Matlab routine of the (improved) Bayesian MCMC estimator of the spatial lag probit model available at his Web site `www.spatial-econometrics.com`, it has been frequently used in empirical research (Bolduc et al., 1997, Mukherjee and Singer, 2008, Wang and Kockelman, 2009, LeSage et al., 2011). Another reason might be that Bayesian MCMC is faster than other estimation techniques (Franzese Jr. and Hays, 2010). One drawback of the Bayesian MCMC estimator is that it is difficult to verify whether convergence actually occurs. In some experiments we carried out it clearly did not, even though LeSage's Matlab routine simply reported final estimation results after the pre-specified number of draws had been passed through. In addition, the convergence to the joint posterior distribution is sometimes sensitive to the choice of the prior distributions (Franzese Jr. and Hays, 2010).

A third estimation method is the Generalized Method of Moments (GMM), initially proposed by Pinkse and Slade (1998) for the estimation of a spatial error probit model.[5] To deal with the endogeneity of the spatially lagged dependent variables in case of the spatial lag model, the variable $\boldsymbol{WY}^*$ is instrumented by $[\boldsymbol{X}\ \boldsymbol{WX} \ldots \boldsymbol{W}^g\boldsymbol{X}]$, where g is a pre-selected constant. Typically, one would take $g = 1$ or $g = 2$, dependent on the number of regressors and the type of model (see Kelejian et al., 2004). To avoid repeated inversions of the matrix $(\boldsymbol{I} - \lambda\boldsymbol{W})$, they linearize the spatial parameters around the non-spatial parameter values that are obtained from a standard (non-spatial) probit or logit model. GMM studies do not specify the distribution function of the error terms, and therefore do not solve the multi-

[5] Klier and McMillen (2008) use the same technique to estimate a spatial lag logit model. Following these two studies, Diallo and Geniaux (2011) propose a GMM estimator for a logit model with both a spatially lagged dependent variable and a spatially autocorrelated error term. Flores-Lagunes and Schnier (2012) develop an estimator for their so-called Tobit type II model. These studies criticize the Bayesian MCMC and ML estimation methods for relying on the potentially inaccurate assumption of normally distributed errors. Instead, they assume that the individual error terms v_i are i.i.d. with mean zero and variance σ^2.

dimensional integration problem. They take into account that the diagonal elements of the covariance matrix are different from one unit to another by scaling the explanatory variables $\boldsymbol{X}_i$ of each unit i by σ_i, $\boldsymbol{X}_1/\sigma_i$, where σ_i represents the ith diagonal element of the covariance matrix of the error term $\boldsymbol{\Omega}_p = [(\boldsymbol{I} - p\boldsymbol{W})'(\boldsymbol{I} - p\boldsymbol{W})]^{-1}$ with $p = \lambda$ in case of the spatial error model and $p = \rho$ in case of the spatial lag model. However, they do no take into account that the off-diagonal elements of this matrix are non-zero too. Consequently, they overrule the basic notion underlying spatial econometric models in general and spatial discrete-response models in particular that units cannot be treated as independent entities. In other words, although these studies are right that the ML and Bayesian methods rely on the potentially inaccurate assumption of normally distributed error terms, they in turn ignore the spatial interaction effects among the error terms.

Our paper adopts a maximum likelihood estimation method. Starting from McMillen (1992), Beron and Vijverberg (2004) developed a Simulated Maximum Likelihood (SML) estimator for the spatial lag probit model. This simulation method is known as Recursive-Importance-Sampling (RIS) and relies on Monte Carlo simulation of truncated multivariate normal distributions, as discussed by Vijverberg (1997). First, a lower-triangular Cholesky matrix of the variance-covariance matrix of the error terms is determined, and then the multidimensional integral in Equation (3) is evaluated. Originally, Vijverberg (1997) considered four different density functions: the logit, normal, t and a transform of the Beta(2,2). Although relatively slow, Beron and Vijverberg (2004) favor the normal distribution for its efficiency. They also point out that the RIS-normal simulator is identical to the Geweke-

Hajivassiliou-Keane (GHK) simulator (Börsch-Supan and Hajivassiliou, 1993, Keane, 1993). The advantage of this estimation method is that it provides a feasible and efficient algorithm to approximate the N-dimensional truncated normal density function needed to maximize the log-likelihood function. A more detailed explanation is given in the next section. Franzese Jr. and Hays (2010) compare the performance of different estimation methods of the spatial lag probit model using Monte-Carlo experiments and find that the RIS simulator produces more efficient estimates of the spatial parameter ρ than Bayesian MCMC. However, the RIS procedure turns out to be computationally intensive and time-consuming. Recently, Pace and LeSage (2012) called the GHK/RIS simulator one of the most effective techniques for computing the N-dimensional truncated normal distribution. They also suggest several sparse matrix algorithms to speed up computation time. Unfortunately, these Matlab routines are not available yet.

3 Transfers at different moments in time

Model

In the Introduction it has been explained that the spatial probit model cannot be fruitfully used to describe the transition from one state to another when this transition for one geographical unit takes place at a different moment in time than for another unit. To deal with this problem we assume that the data are sorted; first the units that are in state 0 at the start of time period t ($\boldsymbol{Y}_t^{0*}$) followed by the units that are already in state 1 at the start of period

t ($\boldsymbol{Y}_t^{1*}$). We initially propose the following spatial probit model

$$\begin{pmatrix} \boldsymbol{Y}_t^{0*} \\ \boldsymbol{Y}_t^{1*} \end{pmatrix} = \rho \begin{pmatrix} \boldsymbol{W}_t^{00} & \boldsymbol{W}_t^{01} \\ \boldsymbol{W}_t^{10} & \boldsymbol{W}_t^{11} \end{pmatrix} \begin{pmatrix} \boldsymbol{Y}_t^{0*} \\ \boldsymbol{Y}_t^{1*} \end{pmatrix} + \begin{pmatrix} \boldsymbol{X}_t^{0} \\ \boldsymbol{X}_t^{1} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \boldsymbol{v}_t^{0} \\ \boldsymbol{v}_t^{1} \end{pmatrix}, \quad (7)$$

where $t = 1, \ldots, T$ is an index for the time dimension. The $N \times N$ matrix $\boldsymbol{W}_t$ describing the spatial arrangement of the units in the sample is partitioned into four submatrices: $\boldsymbol{W}_t^{00}$ expresses spatial relations between the units that are in state 0 at the start of period t; $\boldsymbol{W}_t^{11}$ between the units that are already in state 1 at the start of t; and $\boldsymbol{W}_t^{01}$ and $\boldsymbol{W}_t^{01}$ describe spatial relations of the units in state 0 with the units in state 1 (and vice versa) at the start of period t. Since the number of spatial units in state 0 and 1 may be different from one period to another, these submatrices are time dependent. This is indicated by the subscript t.[6]

If in line with the literature on duration models observations on units after they transferred to state 1 are removed from the sample, only the observation with superscript 0 in the first line of this model remain. If N_t^0 denotes the number of observations that are not yet in state 1 at the start of time period t, the total number of observations to estimate the parameters of this model amounts to $\sum_{t=1}^{T} N_t^0$. Due to the removal of observations that are in state 1, the expected value of the latent variable $\boldsymbol{Y}_t^{1*}$ at the right-hand side is no longer defined. Therefore, we replace the latent variable $\boldsymbol{Y}_t^{1*}$ by the observed variable $\boldsymbol{Z}_t$ which is equivalent to $\boldsymbol{Y}_t^{1}$. Furthermore, since it is reasonable to assume that neighboring units that already transferred to state 1 have

[6]The dimensions of submatrices $\boldsymbol{W}_t^{00}$, $\boldsymbol{W}_t^{01}$, $\boldsymbol{W}_t^{10}$, $\boldsymbol{W}_t^{11}$ are $N_t^0 \times N_t^0$, $N_t^0 \times N_t^1$, $N_t^1 \times N_t^0$ and $N_t^1 \times N_t^1$, respectively, where $N_t = N_t^0 + N_t^1$, for all t.

a different impact than neighboring units that are still in state 0, we allow these two variables to have different coefficients ρ and δ. This yields

$$\boldsymbol{Y}_t^{0*} = \rho \boldsymbol{W}_t^{00} \boldsymbol{Y}_t^{0*} + \delta \boldsymbol{W}_t^{01} \boldsymbol{Z}_t + \boldsymbol{X}_t^0 \boldsymbol{\beta} + \boldsymbol{v}_t^0. \tag{8}$$

In sum, the first variable at the right-hand side, $\boldsymbol{W}_t^{00}\boldsymbol{Y}_t^{0*}$, denotes the endogenous interaction effect with neighboring units that are also in state 0. The second variable at the right-hand side, $\boldsymbol{W}_t^{01}\boldsymbol{Z}_t$, denotes the interaction effect with neighboring units that already transferred to state 1. The first variable, like the dependent variable, is specified in terms of unobserved choices, and the second variable in terms of observed choices. In some studies, there are units that are still in state 0 at the end of the observation period and units that already transferred to state 1 before the start of the observation period. In the first case, the sample is called right-censored; these observations are covered by the first spatial term on the right-hand side of the regression equation. In the second case, the sample is called left-censored; these observations are covered by the second spatial term on the right-hand side of the regression.

Units that did not yet transfer may be affected by neighboring units that also did not yet transfer, and vice versa, as a result of which the right-hand side variable $\boldsymbol{W}^{00}\boldsymbol{Y}_t^{0*}$ needs to be treated as an endogenous explanatory variable. Units that did not yet transfer may also be affected by neighboring units that already transferred. However, since observations on units in time periods after they transferred to state 1 are removed from the sample, units that did already transfer cannot be affected by units that are still in state

0. Consequently, the right-hand side variable $\boldsymbol{W}^{01}\boldsymbol{Z}_t$ may be treated as an exogenous rather than an endogenous explanatory variable.[7] Hence, the parameters in Equation (8) can be estimated similarly to those of a standard spatial lag probit model, Equation (4).

Estimation

Based on strengths and weaknesses of different estimation methods discussed in the previous section, we use the RIS/GHK-simulator for the normal distribution to obtain ML parameter estimates, described in Vijverberg (1997). For this we need to evaluate the N-dimensional integral similar to Equation (3). We explain the mechanisms behind the simulator using a simple example. Assume that $N = 3$, $\boldsymbol{Y} = (1, 0, 1)'$ and that the mean $\boldsymbol{\mu}$ of vector $\boldsymbol{Y}^*$ corresponds to Equation (8) with variance-covariance matrix $\boldsymbol{\Omega}_\rho$, where $\boldsymbol{\Omega}_\rho$ is positive definite. Using the Cholesky decomposition, we can find a lower-triangular matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}\boldsymbol{Q}' = \boldsymbol{\Omega}_\rho$. Taking q_{ij} as elements of $\boldsymbol{Q}$, we have:

$$
\begin{aligned}
y_1^* &= \mu_1 + q_{11}\nu_1 \geq 0 \\
y_2^* &= \mu_2 + q_{21}\nu_1 + q_{22}\nu_2 \leq 0 \\
y_3^* &= \mu_3 + q_{31}\nu_1 + q_{32}\nu_2 + q_{33}\nu_3 \geq 0
\end{aligned}
$$

Note that ν_1 follows a standard normal distribution truncated below by $-\mu_1/q_{11}$. Let z_1 be a draw from this distribution.[8] Conditional on z_1, ν_2

[7] This is the reason why we change the notation from $\boldsymbol{Y}_t^1$ to $\boldsymbol{Z}_t$ in Equation (8).

[8] The procedure of drawing from a truncated standard normal distribution is the following. Let $x \sim \mathcal{N}(0, 1)$ and $x > a$. Then the cumulative distribution function (cdf) of x

follows a standard normal distribution truncated above by

$$\frac{-\mu_2 - q_{21}z_1}{q_{22}}$$

Next, let z_2 be a draw from the distribution of ν_2. Finally, conditional on z_1 and z_2, ν_3 follows a standard normal distribution truncated below by

$$\frac{-\mu_3 - q_{31}z_1 - q_{32}z_2}{q_{33}},$$

and z_3 is a draw from this distribution.

Since z_1, z_2, and z_3 are independently distributed, the mean of $\tilde{p}_r = (1-\Phi(z_1))\Phi(z_2)(1-\Phi(z_3))$ is the joint probability that $y_1^* \geq 0$, $y_2^* \leq 0$ and $y_3^* \geq 0$. If we repeat this procedure R times, then

$$\hat{p} \equiv \frac{1}{R}\sum_{r=1}^{R}\tilde{p}_r$$

is a consistent estimator of the joint probability, known as the Recursive-Importance-Sampling (RIS) estimator (Vijverberg, 1997, Beron and Vijverberg, 2004). The general form of the RIS simulated likelihood for the N-

---

is

$$F(x) = \frac{\Phi(x) - \Phi(a)}{1 - \Phi(a)},$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. If $u \sim \mathcal{U}(0,1)$, then we can draw x by solving $u = F(x)$. This leads to:

$$x = \Phi^{-1}(u(1-\Phi(a)) + \Phi(a)).$$

If $x < a$, then a similar derivation method implies

$$x = \Phi^{-1}((1-u)\Phi(a)).$$

variate case

$$\hat{p} \equiv \frac{1}{R}\sum_{r=1}^{R}\left\{\prod_{j=1}^{N}\Phi(z_{j,r})\right\} \tag{9}$$

is used to find the standard maximum-likelihood estimator of $L(\boldsymbol{\beta}, \rho)$

$$L(\boldsymbol{\beta}, \rho) = \int_{\boldsymbol{Y}_t^{0*}} \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Omega}_\rho|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{v}'\boldsymbol{\Omega}_\rho^{-1}\boldsymbol{v}\right\} d\boldsymbol{v},$$

by the following optimization

$$\min_{\boldsymbol{\beta},\rho}\{-\log L(\boldsymbol{\beta}, \rho)\}$$

subject to $\rho \in (-1/\omega_{min}, 1)$. Note that δ is an element of $\boldsymbol{\beta}$. Finally, we compute the Hessian H of $-\log L(\beta, \rho)$ numerically and calculate standard deviations as the square root of the diagonal elements of H^{-1}.

Lee (2004) and Qu and Lee (2012) show that the ML estimator of respectively the spatial lag and the spatial probit model produces consistent and asymptotically normal estimates, provided that the following regularity conditions are satisfied. $\boldsymbol{W}$ should be a nonnegative matrix of known constants. The diagonal elements are set to zero by assumption. The matrix $(\boldsymbol{I} - \rho\boldsymbol{W})$ should be nonsingular. For a row-normalized $\boldsymbol{W}$, this condition is satisfied as long as ρ is in the interior of $(1/\omega_{min}, 1)$, where ω_{min} denotes the smallest characteristic root of $\boldsymbol{W}$. Furthermore, the row and column sums of the matrices $\boldsymbol{W}$ and $\{(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\}$ should be uniformly bounded in absolute value as N goes to infinity. In a cross-sectional setting, the row and column sums of W should also not diverge to infinity at a rate equal to or faster than the rate of the sample size N before $\boldsymbol{W}$ is row-normalized, but in a panel data setting this condition is not needed, provided that time-period fixed effects are not included (Kelejian et al., 2004).

Direct and indirect effects

It is well-known that the point estimates of the parameter vector $\boldsymbol{\beta}$ in the probit model $\boldsymbol{Y}^* = \boldsymbol{X\beta} + \boldsymbol{v}$ and in the spatial lag model with a continuous dependent variable $\boldsymbol{Y} = \rho\boldsymbol{WY} + \boldsymbol{X\beta} + \boldsymbol{v}$ are not equal to their marginal effects, see respectively Cameron and Trivedi (2005, p.466) and LeSage and Pace (2009, pp.293-297). LeSage et al. (2011) consider the marginal effects of the spatial probit model by combining these two models. When applied to our model set forth in Equation (8), the matrix of partial derivatives of the expected value of $\boldsymbol{Y}$ with respect to the k^{th} explanatory variable of $\boldsymbol{X}$ in unit 1 up to unit N (say x_{ik} for $i = 1, \ldots, N$, respectively) at a particular moment in time t takes the form

$$\begin{pmatrix} \frac{\partial E(\boldsymbol{Y}_t)}{\partial x_{1k}} \cdots \frac{\partial E(\boldsymbol{Y}_t)}{\partial x_{Nk}} \end{pmatrix} = \begin{pmatrix} \frac{\partial E(y_{1t})}{\partial x_{1k}} \cdots \frac{\partial E(y_{1t})}{\partial x_{Nk}} \\ \vdots \ddots \vdots \\ \frac{\partial E(y_{Nt})}{\partial x_{1k}} \cdots \frac{\partial E(y_{Nt})}{\partial x_{Nk}} \end{pmatrix} = \mathrm{diag}(\phi(\boldsymbol{\eta}))(\boldsymbol{I} - \rho\boldsymbol{W}_t^{00})^{-1} I_N \boldsymbol{\beta}_k \tag{10}$$

where $\boldsymbol{\eta} = (\boldsymbol{I} - \rho\boldsymbol{W}_t^{00})^{-1}(\delta\boldsymbol{W}_t^{01}\boldsymbol{Z}_t + \boldsymbol{X}_t^0\boldsymbol{\beta})$ denotes the vector of predicted values of $\boldsymbol{Y}_t^0$.[9] The first matrix on the right-hand side of this equation is a diagonal matrix of order N whose elements ϕ_i represent the probability that the dependent variable takes its observed value, dependent on the observed values of the other units in the sample. For this reason, each observation has

[9] A similar expression applies to the explanatory variable $\boldsymbol{W}_t^{01}\boldsymbol{Z}_t$ with coefficient δ.

its own mean and variance. Define the matrix $\boldsymbol{\Pi}$ as $\boldsymbol{\Pi} = \boldsymbol{\eta}\boldsymbol{\eta}'$, π_{ij} as the $(i,j)^{th}$ element of $\boldsymbol{\Pi}$, $\boldsymbol{\Pi}_{-ii}$ as the $(N-1)\times(N-1)$ matrix that is obtained after removing both row and column i, and $\boldsymbol{\pi}_{-i}$ as the i^{th} row vector and $\boldsymbol{\pi}_{i-}$ as the i^{th} column vector removed from $\boldsymbol{\Pi}$. Then ϕ_i $(i = 1, \ldots, n)$ evaluates the normal probability density function for the observed value of y_i, which is either 0 or 1, with mean $\mu_i + \boldsymbol{\pi}_{-i}\boldsymbol{\Pi}_{-ii}^{-1}(y_i - \mu_i)$ and variance $\pi_{ii} - \boldsymbol{\pi}_{-i}\boldsymbol{\Pi}_{-ii}^{-1}\boldsymbol{\pi}_{i-}$.

The second matrix on the right-hand side is an $N \times N$ matrix whose diagonal elements represent the impact on the dependent variable of unit 1 up to N if the k^{th} explanatory variable in the own unit changes, while its off-diagonal elements represent the impact on the dependent variable if the k^{th} explanatory variable in another unit changes. The first is called a direct effect and the second an indirect or spatial spillover effect. Since both the direct and indirect effects are different for different units in the sample, the presentation of these effects is a problem. If we have N spatial units and K explanatory variables, we obtain K different $N \times N$ matrices of direct and indirect effects. Even for small values of N and K, it may already be rather difficult to report these results compactly. LeSage and Pace (2009) therefore propose to report one direct effect measured by the average of the diagonal elements of the matrix on the right-hand side of Equation (10), and one indirect effect measured by the average of either the row sums or the column sums of the non-diagonal elements of that matrix. Since the numerical magnitudes of these two calculations of the indirect effect are the same, it does not matter which one is used. Usually, the indirect effect is interpreted as the impact of changing a particular element of an exogenous variable on the dependent variable of all other units, which corresponds to the average column effect. In contrast to LeSage et al. (2011) and other studies in the spatial econometrics literature, the right-hand side of our model is not independent of time t. To obtain one summary indicator for the direct effect and for the indirect effect of every explanatory variable in the model, we therefore propose to also average the outcomes over time.

The standard errors and t-values of the direct and indirect effects estimates are more difficult to determine, because they depend on β_k, ρ and the elements of the spatial weights matrix $\boldsymbol{W}_t^{00}$ in a rather complicated way. In order to draw inferences regarding the statistical significance of the direct and indirect effects, LeSage and Pace (2009, p. 39) suggest simulating the distribution of the direct and indirect effects using the variance-covariance matrix implied by the maximum likelihood estimates. If the full parameter vector $\boldsymbol{\theta} = (\rho, \delta, \boldsymbol{\beta}')'$ is drawn D times from $N(\hat{\boldsymbol{\theta}}, \text{AsyVar}(\hat{\boldsymbol{\theta}}))$, the standard deviation of each summary indicator can be approximated by the standard deviation of the mean value over these D draws, and the significance by dividing each summary indicator by the estimated corresponding standard deviation.

4 Illustration

To illustrate our model, we analyze the transition of countries from one type of monetary policy strategy to another. Specifically, we focus on the adoption of inflation targeting, a monetary policy strategy that involves the public announcement of medium-term targets for inflation and a strong commitment to price stability as a final monetary policy objective (Mishkin and

Schmidt-Hebbel, 2001). The decision of countries to adopt inflation targeting is influenced by their own characteristics as well as choices of neighboring or geographically proximate countries that decide to either adopt inflation targeting or use an alternative monetary strategy. Hence, the testable hypotheses are whether endogenous interaction effects affect the probability of countries to adopt inflation targeting and whether the explanatory variables of inflation targeting cause significant spatial spillover effects.

In our analysis, we assume that in each time period (year) a country can be in one of two possible states: state 1 corresponds to the adoption of inflation targeting, while state 0 corresponds to an alternative monetary strategy. We are interested in estimating the probability of transition from non-inflation targeting to inflation targeting, i.e. from state 0 to state 1.

Data description

Our panel dataset is taken from Samarina and de Haan (2013) and consists of 58 countries over the period 1985–2008; 29 countries adopted inflation targeting during this period, whereas 29 countries did not. Table A.1 in the Appendix provides the list of countries in our dataset. We use official adoption dates based on central banks' announcements and follow the 'half-year-rule': if a country adopts inflation targeting in the second half of year t (July-December), the adoption year is $(t + 1)$, otherwise the adoption year is t.

We adopt the spatial weight matrix $\boldsymbol{W}$, in which the elements $w_{ij} = 1$ if country j belongs to the 10 nearest neighbors of country i in the sample, and 0 otherwise. The diagonal elements in $\boldsymbol{W}$ are set to zero. The weight matrix is row-normalized before it is split into submatrices.

The matrix $\boldsymbol{X}$ includes six explanatory variables that are associated with countries' motivation for adopting inflation targeting: inflation, output growth, the exchange rate regime, government debt, financial development and central bank instrument independence. For details see Samarina and de Haan (2013).

Table A.2 in the Appendix describes the explanatory variables and their data sources. The explanatory variables are assumed strictly exogenous and although not reported here do not highly correlate with each other. Unfortunately, the data set is not complete; the percentage of missing observations on the different explanatory variables ranges from 1% to 13% of all observations. In order to have a complete dataset, an imputation technique is used for filling in missing observations.[10]

Estimation results

Table 1 reports the estimation results when the explanatory variables are included without lags.[11] We record the coefficient estimates and their t-statistics for three specifications of the spatial probit model.

Column (1) of Table 1 reports the estimation results for the standard spatial lag probit model when pooling the cross-sectional data over time. This model can be obtained from Equation (4) by adding a subscript t, which

[10] We apply the Expectation-Maximization (EM) algorithm for missing values imputation, which is described in Dempster et al. (1977) and Schafer (1997).

[11] Outcomes in which explanatory variables are included with a one-year lag to avoid potential endogeneity problems are qualitatively similar and hence not reported here.

runs from 1 to T, to the variables and the error terms of that equation. This model is similar to the one employed in Mukherjee and Singer (2008) for their analysis of inflation targeting adoption. We find that the coefficient estimate ρ of the endogenous interaction effects is positive and significant, while Mukherjee and Singer (2008) report a positive but insignificant result. One explanation is that we use data over a longer time period, 1985-2008 versus 1987-2003 in Mukherjee and Singer (2008). The findings for two variables used in their study and ours—exchange rate regime and central bank independence—are comparable, while the result for inflation is different, both in terms of the sign and significance of the estimate.

Table 1: Estimation Results - Spatial Lag Probit

Variables	(1) Standard spatial probit	(2) Our spatial probit	(3) ...with duration variable
ρ	0.222 ***	0.137	0.022
	(4.20)	(1.01)	(0.17)
δ		0.371	−1.525 **
		(0.77)	(−2.52)
Inflation	−8.850 ***	−0.203	−4.018 **
	(−6.98)	(−0.27)	(−2.22)
Output growth	−0.006	−0.042 *	−0.076 ***
	(−0.39)	(−1.87)	(−2.66)
Exchange rate regime	0.950 ***	0.764 ***	0.773 ***
	(9.93)	(4.25)	(3.97)
Government debt	−0.005 ***	−0.003	−0.006 *
	(−3.28)	(−1.10)	(−1.85)
Financial development	−0.253 ***	−0.235	−0.437 **
	(−2.82)	(−1.31)	(−2.14)
Central bank instrument independence	0.550 ***	0.506 ***	0.367 *
	(5.97)	(2.60)	(1.79)
Constant	−0.283	−1.823 ***	−1.963 ***
	(−1.59)	(−5.33)	(−4.26)
Duration of non-IT period			0.082 ***
			(3.41)
Observations	1392	1127	1127
Log-Likelihood	−500.8	−116.0	−105.8

Notes: Table 1 reports coefficient estimates and their t-values (in parentheses). *, ** and *** denote significance at 10%, 5% and 1% significance level, respectively. Column (1) shows the results for a standard spatial lag probit model, column (2) reports estimation results for our spatial probit model, while column (3) extends the model by controlling for temporal dependence.

Column (2) shows estimation results for our spatial probit model with two spatially lagged variables and full set of regressors. The results suggest that the coefficients of both spatial lags are insignificant. Note that the value of the log-likelihood is much higher in our specification.

The estimation of our spatial probit model, similarly to discrete-time duration models, may face the potential temporal dependency problem. This means that the probability of a country to adopt inflation targeting in year t may depend on the duration of the non-inflation targeting (non-IT) period, i.e., the time that has passed from the start of the sample period until the inflation targeting adoption date. Ignoring temporal dependence may lead to inefficiency and inaccurate statistical inference (Beck et al., 1998). To correct for temporal dependence, we follow the approach of Beck et al. (1998) and construct a duration variable that counts the number of years from the start of the sample period until the inflation targeting adoption date.[12]

Column (3) reports the estimation results when the duration variable is added as a regressor to our spatial probit model. Importantly, the coefficient estimate of the duration variable turns out to be significant with a positive sign, implying that the longer is the non-IT period, the more likely are countries to adopt inflation targeting. Hence, this specification of our spatial probit model is preferred.

The coefficient estimate of the endogenous interaction effect in Column (3), ρ, has again a positive sign. This indicates that countries that adopt inflation targeting in the current period, i.e. transfer from state 0 to state 1

[12] Alternatively, one could generate a set of time dummies that mark each non-inflation targeting duration period. However, including 24 dummies leads to a substantial loss of degrees of freedom and a substantial increase of computation time.

in year t, have a positive effect on the probability of other countries to make the same decision in that period. However, in contrast to the pooled model, this result is no longer statistically significant.

The coefficient estimate of the spatial interaction effect with countries that already adopted inflation targeting before period t, δ, is significant with a negative sign. There are several explanations for this negative spatial interaction effect. An intuitive explanation for the negative sign of δ is that more countries adopt inflation targeting as time elapses, as a result of which there are less countries in the sample left that did not transfer yet. Therefore, over time the probability to adopt inflation targeting becomes lower as the number of neighboring countries that can decide to switch, diminishes. Another explanation is that the explanatory variables of countries that did and did not yet adopt inflation targeting take different values and are treated differently. Whereas these explanatory are part of the model for the latter group of countries, they are not for the former group since observations on these countries are removed from the sample. Consequently, the impact these countries have on countries that are still considering inflation targeting only runs through the spatial interaction coefficient, which therefore also captures the effect of any value changes in these explanatory variables.

Additionally, we find that countries with lower output growth, more flexible exchange rate regimes and lower debt are more likely to adopt inflation targeting. Financial system development has a negative significant influence on the probability to adopt inflation targeting, while the estimate for central bank instrument independence is significant with a positive sign. It should be noted that differences between these variables in the pre- and post-inflation targeting periods affect the magnitude of the spatial interaction coefficient too, which therefore should be interpreted as the net effect of all value changes in all explanatory variables between these two periods.

Direct and indirect effects

Table 2 shows the estimates of direct, indirect, and total effects of the three models, as reported in Table 1. They are used to measure the effects of changes in explanatory variables on the probability of a particular country to adopt inflation targeting (*direct effects*), as well as the spatial spillover effects of explanatory variables on neighboring countries (*indirect effects*).

The results show that only two direct effects and not one single spatial spillover effect is significant in our model specifications (2) and (3), while both five direct and five spatial spillover effects appear to be (weakly) significant in the standard spatial probit model (1). Additionally, the estimated effects in model (1) tend to be larger in magnitude (absolute values) than their counterparts in models (2) and (3). Thus, using a standard spatial probit approach to analyze state transfers at different moments in time (on the example of inflation targeting adoption) leads to overly optimistic statistical inference and largely overestimated direct and, above all, spatial spillover effects. Once the model specification is altered, as in our spatial probit with two spatial terms and duration dependence, we do not find any evidence in support of spatial spillover effects.

Table 2: Marginal effects

Variables	Direct effects	Indirect effects	Total effects
(1) *Standard spatial probit*			
Inflation	−1.919 ***	−0.558 ***	−2.477 ***
	(−9.02)	(−2.74)	(−6.75)
Output growth	−0.002	−0.001	−0.002
	(−0.56)	(−0.51)	(−0.55)
Exchange rate regime	0.208 ***	0.062 **	0.270 ***
	(6.87)	(2.27)	(4.80)
Government debt	−0.001 ***	−0.0003**	−0.002 ***
	(−3.00)	(−2.06)	(−2.88)
Financial development	−0.056 ***	−0.016 *	−0.072 ***
	(−2.97)	(−1.91)	(−2.77)
Central bank instrument independence	0.116 ***	0.034 **	0.150 ***
	(5.35)	(2.30)	(4.37)
(2) *Our spatial probit*			
Countries that already adopted IT ($\boldsymbol{Z}_t$)	−0.251	−0.037	−0.288
	(−1.36)	(−0.71)	(−1.35)
Inflation	−0.036	−0.006	−0.042
	(−0.30)	(−0.19)	(−0.29)
Output growth	−0.005	−0.001	−0.006
	(−0.74)	(−0.47)	(−0.72)
Exchange rate regime	0.110	0.018	0.128
	(1.29)	(0.71)	(1.25)
Government debt	0.002	0.0002	0.002
	(0.46)	(0.19)	(0.43)
Financial development	−0.032	−0.005	−0.037
	(−0.91)	(−0.50)	(−0.89)
Central bank instrument independence	0.067	0.011	0.078
	(1.29)	(0.68)	(1.22)
(3) *Our spatial probit with a duration variable*			
Countries that already adopted IT ($\boldsymbol{Z}_t$)	−0.133 *	−0.011	−0.145
	(−1.78)	(−0.47)	(−1.49)
Inflation	−0.274	−0.020	−0.295
	(−1.54)	(−0.46)	(−1.38)
Output growth	−0.006	−0.001	−0.006
	(−1.34)	(−0.41)	(−1.14)
Exchange rate regime	0.058	0.006	0.065
	(1.45)	(0.48)	(1.23)
Government debt	−0.0004	−0.0000	−0.0004
	(−1.11)	(−0.33)	(−0.98)
Financial development	−0.030	−0.002	−0.032
	(−1.44)	(−0.48)	(−1.29)
Central bank instrument independence	0.027	0.003	0.030
	(1.20)	(0.47)	(1.09)
Duration of non-IT period	0.006 *	0.001	0.006
	(1.88)	(0.50)	(1.61)

Notes: Table 2 reports direct, indirect, and total effects with t-values (in parentheses) for models (1), (2), and (3) as in Table 1. *, ** and *** denote significance at 10%, 5% and 1% significance level, respectively.

This also follows from the year when different countries adopted inflation targeting (see Table A.1). For instance, countries that adopted inflation targeting in 2001, namely Hungary, Iceland, Mexico, and Norway, are anything but close neighbors to each other. Conversely, countries that are close neighbors to each other, like Czech Republic, Poland, Hungary, Slovakia, and Romania, adopted inflation targeting in various years, respectively 1998, 1999, 2001, 2005, and 2006. These two examples do not give much evidence in favor of the hypothesis that if one country adopts inflation targeting a neighboring country will also do so. Our model corroborates this finding, whereas the standard spatial probit model that pools the data does not.

5 Conclusion

The spatial probit model can be employed to describe interaction effects among geographical units when the dependent variable takes the form of a binary response variable. Unfortunately, it cannot adequately deal with transitions from one state to another when these transitions take place at different moments in time for different geographical units.

This paper proposes a spatial probit model with two spatially lagged variables, one for units that did not transfer to the other state yet, and one for units that already transferred. Observations on units that made the transfer from one state to the other are removed after the transfer. Hence, units that already made the transfer cannot be affected by units that are still in the original state. This allows us to treat the spatially lagged variable for units that already transferred, as exogenous. The empirical model is estimated by maximum likelihood methods, using the Recursive Importance Sampling simulator to evaluate the truncated multidimensional normal distribution.

We illustrate our approach with a study of the adoption of inflation targeting for a sample of 58 countries over the period 1985–2008. We make a distinction between interaction effects among countries that did not adopt inflation targeting yet and interaction effects of countries that already adopted inflation targeting on countries that did not adopt yet. In contrast to previous studies, we find that the first interaction effect is insignificant, whereas the second interaction effect that has not been considered in previous studies turns out to be negative and significant. In addition, we find no evidence in favor of spatial spillover effects.

Our spatial probit model has various applications in economics, business, and political studies. The approach can also be used to explain contagion of financial crises when countries enter crises at different time. Additionally, studies on the introduction of new brands and firms' entry decisions might also include state transfers at different periods.

Acknowledgements

This paper is the result of our reading group on spatial binary response models. We thank Tom Wansbeek and Paul Bekker for their valuable input in reading and discussing the articles, Badi Baltagi, Chandra Bhat, Peter Egger, and Jim LeSage for helpful discussions and the participants of the 59th Annual North American Meetings of the Regional Science Association International, Ottawa, the KOF Research Seminar, Zurich, the Scottish Economic Society Annual Conference, Perth, 12th International Workshop Spatial Econometrics and Statistics, Orléans, France, the Netherlands Econometric Study Group, Amsterdam, and the 9th Triennial Choice Symposium, Noordwijk, the Netherlands for comments and suggestions.

References

Arnone, M., B. J. Laurens, J.-F. Segalotto, and M. Sommer (2007). Central bank autonomy: Lessons from global trends. *IMF Working Paper* 07/88, International Monetary Fund.

Beck, N., J. N. Katz, and R. Tucker (1998). Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science 42*, 1260–1288.

Beron, K. J., J. C. Murdoch, and W. P. Vijverberg (2003). Why cooperate? Public goods, economic power, and the Montreal Protocol. *The Review of Economics and Statistics 85*, 286–297.

Beron, K. J. and W. P. Vijverberg (2004). Probit in a spatial context: A Monte Carlo analysis. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Chapter 8, pp. 169–195. Berlin: Springer-Verlag.

Bhat, C. R., I. N. Sener, and N. Eluru (2010). A flexible spatially dependent discrete choice model; formulation and application to teenagers' weekday recreational activity participation. *Transportation Research Part B 44*, 903–921.

Bolduc, D., B. Fortin, and S. Gordon (1997). Multinomial probit estimation of spatially interdependent choices: an empirical comparison of two new techniques. *International Regional Science Review 20*, 77–101.

Börsch-Supan, A. and V. A. Hajivassiliou (1993). Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of Econometrics 58*, 347–368.

Cameron, C. A. and P. K. Trivedi (2005). *Microeconometrics. Methods and Applications*. U.S.: Cambridge University Press.

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological) 34*, 187–220.

Cukierman, A., G. P. Miller, and B. Neyapti (2002). Central bank reform, liberalization and inflation in transition economies—an international perspective. *Journal of Monetary Economics 49*, 237–264.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological) 39*, 1–38.

Diallo, A. and G. Geniaux (2011). Spatial logit for large samples with local spatial lag and regional spatial random effects using linearized GMM: an application to land use models. Presented at the 5th World Conference of the Spatial Econometrics Association (SEA), Toulouse, 8 July 2011.

Fleming, M. M. (2004). Techniques for estimating spatially dependent discrete choice models. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Chapter 7, pp. 145–167. Berlin: Springer-Verlag.

Flores-Lagunes, A. and K. E. Schnier (2012). Estimation of sample selection models with spatial dependence. *Journal of Applied Econometrics 27*, 173–204.

Franzese Jr., R. J. and J. C. Hays (2010). Spatial, temporal, and spatiotemporal autoregressive probit models of binary outcomes: estimation, interpretation, and presentation. Presented at the Annual Meeting of the Public Choice Society, 6 March 2009.

Jaimovich, D. and U. Panizza (2010). Public debt around the world: a new data set of central government debt. *Applied Economics Letters 17*, 19–24.

Keane, M. P. (1993). Simulation estimation for panel data models with limited dependent variables. In *Handbook of Statistics, Vol. 11: Econometrics.* Amsterdam: North-Holland.

Kelejian, H. H., I. R. Prucha, and Y. Yuzefovich (2004). Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: large and small sample results. In J. P. LeSage and K. R. Pace (Eds.), *Spatial and Spatiotemporal Econometrics*, pp. 163–198. Amsterdam: Elsevier.

Klier, T. and D. P. McMillen (2008). Clustering of auto supplier plants in the United States: Generalized Method of Moments spatial logit for large samples. *Journal of Business Economics and Statistics 26*, 460–471.

Lee, L.-f. (2004). Asymptotic distributions of Quasi-Maximum Likelihood estimators for spatial autoregressive models. *Econometrica 72*, 1899–1925.

LeSage, J. P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis 32*, 19–35.

LeSage, J. P., K. R. Pace, N. Lam, R. Campanella, and X. Liu (2011). New Orleans business recovery in the aftermath of Hurricane Katrina. *Journal of the Royal Statistical Society A 174*, 1007–1027.

LeSage, J. P. and R. K. Pace (2009). *Introduction to spatial econometrics.* U.S.: Chapman&Hall/CRC.

Levy-Yeyati, E. and F. Sturzenegger (2005). Classifying exchange rate regimes: Deeds vs. words. *European Economic Review 49*, 1603–1635.

McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science 32*, 335–348.

Mishkin, F. S. and K. Schmidt-Hebbel (2001). One decade of inflation targeting in the world: What do we know and what do we need to know? NBER Working Paper 8397, National Bureau of Economic Research.

Mukherjee, B. and D. A. Singer (2008). Monetary institutions, partisanship, and inflation targeting. *International Organization 62*(2), 323–358.

Pace, R. K. and J. P. LeSage (2012). Fast simulated maximum likelihood estimation of the spatial probit model capable of handling large samples. Available at SSRN: http://ssrn.com/abstract=1966039.

Pinkse, J. and M. E. Slade (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics 85*, 125–154.

Qu, X. and L.-f. Lee (2012). LM tests for spatial correlation in spatial model with limited dependent variables. *Regional Science and Urban Economics 42*, 430–445.

Samarina, A. and J. de Haan (2013). Right on target: Exploring the factors leading to inflation targeting adoption. *Contemporary Economic Policy*. DOI: 10.1111/coep.12017.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data. Monographs on Statistics and Applied Probability No 72*. London: Chapman & Hall/CRC.

Smirnov, O. A. (2010). Modeling spatial discrete choice. *Regional Science and Urban Economics 40*, 292–298.

Soetevent, A. R. and P. Kooreman (2007). A discrete choice model with social interactions: with an application to high school teen behavior. *Journal of Applied Econometrics 22*, 599–624.

Vijverberg, W. P. (1997). Monte Carlo evaluation of multivariate normal probabilities. *Journal of Econometrics 76*, 281–307.

Wang, X. and K. M. Kockelman (2009). Bayesian inference for ordered response data with a dynamic spatial-ordered probit model. *Journal of Regional Science 49*, 877–913.

Appendix

Table A.1: Country sample

Inflation targeting countries (29)

Country	Adoption year	Country	Adoption year
Australia	1993	Norway	2001
Brazil	1999	Peru	2002
Canada	1991	Philippines	2002
Chile	1991	Poland	1999
Colombia	2000	Romania	2006
Czech Republic	1998	Slovakia	2005
Finland	1993	South Korea	1998
Ghana	2007	South Africa	2000
Guatemala	2005	Spain	1995
Hungary	2001	Sweden	1993
Iceland	2001	Switzerland	2000
Indonesia	2005	Thailand	2000
Israel	1992	Turkey	2006
Mexico	2001	United Kingdom	1993
New Zealand	1990		

Non-inflation targeting countries (29)

Argentina	Denmark	Ireland	Netherlands
Austria	Egypt	Italy	Pakistan
Belgium	Estonia	Japan	Panama
Bolivia	France	Latvia	Portugal
Bulgaria	Germany	Lithuania	Singapore
China	Greece	Luxembourg	United States
Costa Rica	India	Malaysia	Venezuela
Cyprus			

Source: Samarina and de Haan (2013).

Table A.2: Explanatory variables and data sources

Variable	Description and data sources
Inflation	Annual CPI inflation rate transformed as $\frac{\pi_t/100}{1+\pi_t/100}$. Sources: IFS IMF, Datastream.
Output growth	Annual GDP growth rates, in %. Sources: IFS IMF, WDI&GDF World Bank.
Exchange rate regime	Dummy variable, 1 - floating exchange rate regime, 0 - fixed exchange rate regime. Source: Levy-Yeyati and Sturzenegger (2005).
Government debt	Central government debt as % of GDP. Sources: Datastream, Jaimovich and Panizza (2010)
Financial development	Private credit (domestic credit provided by the banking sector)/GDP. Sources: WDI&GDF World Bank.
Central bank instrument independence	Dummy variable, 1 - central bank is instrument independent, 0 - otherwise. Sources: Cukierman et al. (2002), Arnone et al. (2007), central banks laws.

List of research reports

12001-HRM&OB: Veltrop, D.B., C.L.M. Hermes, T.J.B.M. Postma and J. de Haan, A Tale of Two Factions: Exploring the Relationship between Factional Faultlines and Conflict Management in Pension Fund Boards

12002-EEF: Angelini, V. and J.O. Mierau, Social and Economic Aspects of Childhood Health: Evidence from Western-Europe

12003-Other: Valkenhoef, G.H.M. van, T. Tervonen, E.O. de Brock and H. Hillege, Clinical trials information in drug development and regulation: existing systems and standards

12004-EEF: Toolsema, L.A. and M.A. Allers, Welfare financing: Grant allocation and efficiency

12005-EEF: Boonman, T.M., J.P.A.M. Jacobs and G.H. Kuper, The Global Financial Crisis and currency crises in Latin America

12006-EEF: Kuper, G.H. and E. Sterken, Participation and Performance at the London 2012 Olympics

12007-Other: Zhao, J., G.H.M. van Valkenhoef, E.O. de Brock and H. Hillege, ADDIS: an automated way to do network meta-analysis

12008-GEM: Hoorn, A.A.J. van, Individualism and the cultural roots of management practices

12009-EEF: Dungey, M., J.P.A.M. Jacobs, J. Tian and S. van Norden, On trend-cycle decomposition and data revision

12010-EEF: Jong-A-Pin, R., J-E. Sturm and J. de Haan, Using real-time data to test for political budget cycles

12011-EEF: Samarina, A., Monetary targeting and financial system characteristics: An empirical analysis

12012-EEF: Alessie, R., V. Angelini and P. van Santen, Pension wealth and household savings in Europe: Evidence from SHARELIFE

13001-EEF: Kuper, G.H. and M. Mulder, Cross-border infrastructure constraints, regulatory measures and economic integration of the Dutch – German gas market

13002-EEF: Klein Goldewijk, G.M. and J.P.A.M. Jacobs, The relation between stature and long bone length in the Roman Empire

13003-EEF: Mulder, M. and L. Schoonbeek, Decomposing changes in competition in the Dutch electricity market through the Residual Supply Index

13004-EEF: Kuper, G.H. and M. Mulder, Cross-border constraints, institutional changes and integration of the Dutch – German gas market

13005-EEF: Wiese, R., Do political or economic factors drive healthcare financing privatisations? Empirical evidence from OECD countries

13006-EEF: Elhorst, J.P., P. Heijnen, A. Samarina and J.P.A.M. Jacobs, State transfers at different moments in time: A spatial probit approach www.rug.nl/feb