Journal of Economic Dynamics & Control 145 (2022) 104546

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Smooth Transition Simultaneous Equation Models☆

Anjeza Kadilli[a,*], Jaya Krishnakumar[b]

[a] Geneva School of Business Administration, HES-SO, University of Applied Sciences and Arts Western Switzerland, Campus Batelle, Rue de la Tambourine 17, Carouge 1227, Geneva, Switzerland

[b] Geneva School of Economics and Management, University of Geneva, Bvd. du Pont d'Arve 40, CH-1211 Geneva 4, Switzerland

ARTICLE INFO

Article history:
Received 23 December 2020
Revised 8 October 2022
Accepted 11 October 2022
Available online 18 October 2022

JEL classification:
C34
C51
G00

Keywords:
Implicit reduced form
Market comovement
NL2SLS estimation method
Nuisance parameters
Simulations
Simultaneity testing

ABSTRACT

This paper proposes a generalization of the nonlinear simultaneous equation model of Pesaran and Pick (2007) by modelling the comovement between the two endogenous variables as a smooth function of the magnitude of the endogenous variable rather than a step function. The threshold and the speed at which a shock is transmitted are estimated with the other parameters of the model. We investigate the properties of an accurate estimation method which takes into account endogeneity, and a testing procedure for simultaneity in the presence of nuisance parameters under the null hypothesis. We study the conditions on the parameters that ensure the uniqueness of the implicit reduced form of the model. We apply this methodology to the comovement between the sovereign and banking sectors of nine developed countries.

1. Introduction

This article proposes a generalization of the nonlinear simultaneous equation model of Pesaran and Pick (2007) by introducing a smooth transition mechanism in place of a regime shift and estimating the parameters by nonlinear two-stage least squares. We also derive a test for simultaneity in the presence of nuisance parameters, adapting a procedure developed in the framework of smooth transition regression (STR) models (Luukkonen et al., 1988). Further, we discuss the theoretical conditions under which the system of equations has a unique solution.

In Pesaran and Pick's (2007) framework the dependence between the two endogenous variables is modeled by the presence of a dummy variable in each structural equation, which accounts for the effect of one endogenous variable on another

☆ The authors are particularly grateful to Laurent Barras, Christoph Bertsch, Gerda Cabej, Ines Chaieb, Carlos de Porres, Alexandre Jeanneret, Christophe Hurlin, Henri Loubergé, Dietmar Maringer, Hashem Pesaran, Andreas Pick, Erik von Schedvin, for their valuable comments and insights. This paper was enriched from comments and suggestions from the participants of the following conferences and seminars: Young Swiss Economists Meeting, International Association for Applied Econometrics Conference, Sveriges Riksbank, Institute for New Economic Thinking at the University of Southern California, CUSO Young Researcher Conference, Geneva School of Economics and Management, Swiss National Bank, University of Orléans. Part of this research was conducted when the first author was visiting the Research Department of Sveriges Riksbank and the Institute for New Economic Thinking at the University of Southern California. The authors would like to thank the referees for their valuable comments and suggestions.

* Corresponding author.
E-mail addresses: anjeza.kadilli@hesge.ch (A. Kadilli), jaya.krishnakumar@unige.ch (J. Krishnakumar).

https://doi.org/10.1016/j.jedc.2022.104546
0165-1889/

when the value of the former exceeds an unknown threshold. This modeling presents a limitation since the dependence structure is insensitive to the magnitude of the shock above the threshold. Moreover, the model assumes that there is a jump in the transmission of shocks. The main aim of this paper is to add flexibility to the dependence structure by modeling it as a smooth function of the magnitude of the shock (i.e. the value of the endogenous variable). Following a broad literature on STR models, we define the endogenous variable by a logistic function.

The threshold and the smoothness parameters are estimated jointly with the other parameters of the model using heuristic methods.[1]

The nonlinearity of the model in the parameters and in the endogenous variables has a number of consequences concerning the estimation method, the testing procedure for simultaneity and the equilibrium of the system. To perform the nonlinear two-stage least squares estimation procedure, we use low-order polynomials of the excluded exogenous variables as instruments for the logistic function, under appropriate identification conditions (Kelejian, 1971). The test for simultaneity contains unidentified parameters under the null hypothesis, in which case the usual test statistics have non-standard distributions. We adapt a testing procedure developed by Luukkonnen et al. (1988) in the framework of STR models. A polynomial approximation around a fixed value of the smoothness parameter is used to avoid the nuisance parameter issue.

The presence of dummy variables in Pesaran and Pick's model introduces multiple solutions in the system. This is known in the literature as an incoherency issue with important drawbacks such as the unsuitability of the model for prediction and the unknown efficiency of the parameters (e.g. Blundell and Smith, 1994; Gourieroux et al., 1980). In our framework, conditions on the parameters can be set to ensure a unique equilibrium of the system, and thus a unique reduced form. The nonlinearity of the system does not allow an explicit reduced form, but an implicit one can be obtained using numerical methods.

We conduct a series of Monte Carlo simulations to assess the properties of the estimation method and the tests for simultaneity in this novel setting. The model is able to produce excessive skewness and strong dependence between the endogenous variables, although the error terms are i.i.d. Gaussian and not correlated between the equations. Conditioning on the smoothness and on the threshold parameters, the estimation method provides highly accurate results for the dependence coefficient and the parameters related to the common and to the equation-specific factors. When those parameters are unknown, the results are less precise, as expected, but generally improve as the sample size increases. The tests for simultaneity are easy to implement and have good size and power properties which improve with the sample size.

One potential use of this model is the investigation of comovement between different asset classes. Various approaches are adopted to quantify market comovement.[2] A commonly used one is the correlation analysis (Bekaert et al., 2009; Forbes and Rigobon, 2002). Other methods are VAR models (Dees et al., 2007; Ehrmann et al., 2011), latent factor models (Dungey and Martin, 2007), probability analysis (Eitrheim and Teräsvirta, 1996), extreme value analysis (Longin and Solnik, 2001), quantile regressions (Cappiello et al., 2014), jump-diffusion models (Aït Sahalia et al., 2014).

Several of the above approaches fail to capture important features of market comovement such as (i) the potentially simultaneous occurrence across different markets (i.e. endogeneity), (ii) the nonlinear transmission of shocks, (iii) the magnitude at which a shock is transmitted and its smooth or abrupt nature, (iv) the indirect dependence implied by common observed and unobserved factors which can be of local or global nature. We develop a framework that introduces flexibility in all these directions.

We apply our methodology to study the comovement between the sovereign and the banking sectors in the main Euro Area countries, Switzerland, the United Kingdom and the United States. Our results show that comovement strengthens in distressed times. This is especially the case for countries with a riskier banking sector and fragile public finances. For countries with a solid banking sector and public finances, the sectors seem to serve as a diversification opportunity to each other. The data reveal that the transmission of shocks is rather abrupt and occurs for shocks of moderate to high magnitude.

The outline of the paper is as follows. In Section 2 we present the smooth transition simultaneous equation model and its main characteristics. The estimation method and the tests for simultaneity are described in Sections 3 and 4, respectively. The empirical investigation of the comovement between the sovereign and the banking sectors is discussed in Section 5. In Section 6 we conclude with a brief discussion on further research.

2. The model

2.1. The two-equation threshold SEM

The two-equation nonlinear threshold simultaneous equation model (that we abbreviate by T-SEM) developed in Pesaran and Pick (2007) can be written as follows:

$$y_{1,t} = \boldsymbol{\delta}_1' \boldsymbol{z}_t + \boldsymbol{\alpha}_1' \boldsymbol{x}_{1,t} + \beta_1 \mathbb{1}(y_{2,t} > c_2) + u_{1,t} \tag{1}$$

[1] The difficulty of a precise estimation of the smoothness and of the threshold parameters is a well known issue in the STR literature (e.g. González and Teräsvirta (2006); Maringer and Meyer (2008); Schleer (2015)). In our framework another difficulty is added since at the first step of the estimation procedure, the logistic function is approximated by a polynomial. To assess the accuracy of the parameter estimation for each equation, we proceed in three steps: (i) fix both the above-mentioned coefficients, (ii) fix only the smoothness parameter and (iii) estimate both of them with the other parameters of each equation.

[2] For a review of these methods, their advantages and limitations see Forbes (2012).

$$y_{2,t} = \boldsymbol{\delta}_2'\boldsymbol{z}_t + \boldsymbol{\alpha}_2'\boldsymbol{x}_{2,t} + \beta_2 \mathbb{1}(y_{1,t} > c_1) + u_{2,t} \tag{2}$$

where $y_{i,t}$ for $i = 1, 2$ and $t = 1, \ldots, T$, can be a return, an interest rate, a spread, $\boldsymbol{z}_t$ and $\boldsymbol{x}_{i,t}$ are p- and k_i-vectors of exogenous / predetermined common and equation-specific variables, respectively. $\boldsymbol{z}_t$ contains a column of ones associated with the constant term along with other common exogenous factors. $\boldsymbol{x}_{i,t}$ can contain lags of the dependent variable, leading to a dynamic system. $\boldsymbol{\delta}_i$, $\boldsymbol{\alpha}_i$, β_i and c_j are a set of unknown parameters. $u_{i,t}$ is $i.i.d(0, \sigma^2_{u_i})$ and the correlation between the error terms is denoted by ρ. Even though $corr(u_{1,t}, u_{2,t})$ is assumed to be constant, $corr(y_{1,t}, y_{2,t})$ is time-varying.

$\mathbb{1}(y_{j,t} > c_j)$ introduces the direct dependence from market j to market i, when market j is hit by a negative[3] shock. As shown by a large literature, negative shocks are transmitted to a much larger extent compared to positive shocks. That is why we focus on the former. The variable takes the value 1 if $y_{j,t} > c_j$ and 0 if $y_{j,t}$ is below this threshold c_j. The parameter c_j can be interpreted as the level of $y_{j,t}$ which disentangles a normal regime ($y_{j,t} < c_j$) from a crisis regime ($y_{j,t} > c_j$). This is a common interpretation in threshold regression (TR) models. The magnitude of the comovement is measured by β_i, which is our parameter of interest that we name *dependence parameter.*

Note that the model also controls for indirect dependence between y_1 and y_2 which intervenes through the set of common observed variables $\boldsymbol{z}_t$, and through the correlation between u_1 and u_2, ρ, which controls for common unobserved variables. The presence of $\boldsymbol{z}_t$ and non-zero correlation between the error terms are not necessary to have direct comovement between the endogenous variables. On the contrary, the equation-specific variables are necessary to identify these two channels since functions of $\boldsymbol{x}_{j,t}$ serve as instruments for the endogenous variable $\mathbb{1}(\cdot)$.

2.2. The two-equation smooth transition SEM

We replace the indicator function Pesaran and Pick (2007)'s setting by a continuous and bounded function between 0 and 1, which allows for a smooth transmission of shocks. The shock-dependent variable tends to 0 for $y_{j,t}$ far below the threshold parameter and to 1 for $y_{j,t}$ far above the threshold parameter. Further, the transmission of shocks does not only depend on its magnitude, but also on its speed; smooth or abrupt. An unknown smoothness parameter to be estimated from the data plays this role. We name this setting *smooth transition simultaneous equation model* (ST-SEM) and write the two-equation system as follows:

$$y_{1,t} = \boldsymbol{\delta}_1'\boldsymbol{z}_t + \boldsymbol{\alpha}_1'\boldsymbol{x}_{1,t} + \beta_1 G(y_{2,t}; \gamma_1, c_2) + u_{1,t} \tag{3}$$

$$y_{2,t} = \boldsymbol{\delta}_2'\boldsymbol{z}_t + \boldsymbol{\alpha}_2'\boldsymbol{x}_{2,t} + \beta_2 G(y_{1,t}; \gamma_2, c_1) + u_{2,t} \tag{4}$$

where $G(y_{j,t}; \gamma_i, c_j)$ is the dependence variable, γ_i is the smoothness parameter and c_j the location parameter, which plays the same role as before.

Many functional forms could be envisaged for $G(y_{j,t}; \gamma_i, c_j)$. In a large literature concerned with smooth transition regression (STR) models, the most commonly employed function is the logistic function (e.g. van Dijk et al. (2002); Teräsvirta (1994); Teräsvirta et al., 2010), which is written as follows:

$$G(y_{j,t}; \gamma_i, c_j) = \left\{1 + e^{-\gamma_i (y_{j,t} - c_j)}\right\}^{-1} \quad \gamma_i > 0, \quad i, j = 1, 2, \quad i \neq j \tag{5}$$

The cumulative function of the logistic distribution has interesting properties such as continuity and monotonicity. It also has continuous and bounded first and second derivatives with respect to $y_{j,t}$ for $\gamma_i < +\infty$. The logistic function is an appropriate alternative for more complicated and computationally-heavy functions, such as the normal distribution with no explicit cumulative function and involving multiple integrals (Chan and Tong, 1986).

Theoretically, the ST-SEM nests the T-SEM model as a special case when $\gamma_i \to +\infty$, with the logistic function converging to a Heaviside step function. As we will discuss in Section 2.3, for increasing values of the smoothness parameter, the system is more likely to contain multiple solutions. Moreover, as shown in Fig. 1, the logistic function is almost a step function for moderate values of γ_i. Hence, we focus on finite values of this parameter. If γ_i happens to be large, then we should consider a step function instead.

Fig. 1 (left) shows how the steepness of the logistic function changes with the values of γ_i. The function is flat for values of γ_i not far from 0, and converges to a constant as γ_i tends to 0 ($G(y_{j,t}; 0, c_j) = 0.5$). For $\gamma_i > 0$, the function rapidly converges to a step function. In fact, one can see that even for a moderate value of $\gamma_i = 10$, the logistic function is *very close* to the step function, and for $\gamma_i = 100$, it is *literally* a step function for the naked eye. As it will be discussed below, this property presents an important issue for the precise estimation of γ_i. As in Pesaran and Pick (2007), the parameter c_j can be interpreted as the threshold between a normal and a crisis regime.

The right panel of Fig. 1 shows 100 realizations of the logistic function from simulations using the ST-SEM for different values of $\gamma_i = [0.1,\ 0.5,\ 1.5,\ 3.5]$, $\delta_i = \alpha_i = \beta_i = c_j = 1$, $u_{i,t} \sim i.i.d.\mathcal{N}(0, 1)$ and $corr(u_{1,t}, u_{2,t}) = 0.5$. The variation of the function for $\gamma_i = 0.1$ and even for $\gamma_i = 0.5$ is very low. This brings about collinearity issues with the constant term. Thus, for the simulation of the model and for the estimation, the minimum value of γ_i should not be very close to 0. For smoothness

[3] For simplicity, we assume that a negative shock is associated with high values of y. If a negative shock is associated with very low values of y, then taking $-y$ leads to the same model.

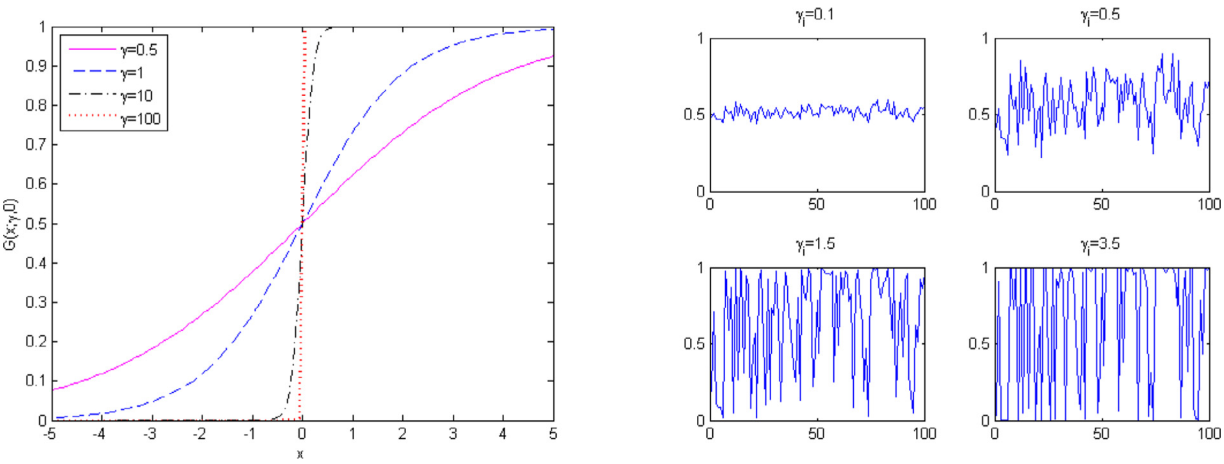

Fig. 1. Logistic distribution for different values of γ. *Notes*: Logistic distribution for different values of γ and $c = 0$: $G(x;\gamma,c) = \left(1 + e^{\gamma(x-c)}\right)^{-1}$.

parameters as low as 1.5, the series of the logistic function exhibits a much higher variation and for $\gamma_i = 3.5$, the function roughly switches back and forth between 0 and 1.

2.3. Reduced form of the model

Given the nonlinearity of the system, an explicit reduced form cannot be obtained. To use this model, we need to make sure that a unique value of $y_{i,t}$ is associated with any admissible value of $\boldsymbol{\zeta}_i = [\boldsymbol{\delta}_i', \boldsymbol{\alpha}_i', \beta_i, \gamma_i, c_j]'$, $\mathbf{h}_{i,t} = [\mathbf{z}_t', \mathbf{x}_{i,t}', G(y_{j,t};\gamma_i, c_j)]'$ and $u_{i,t}$, i.e. an implicit reduced form is well defined. The conditions on the parameters $\boldsymbol{\zeta}_i$ that ensure the uniqueness of the solution are called *coherency conditions*. An incoherent system can be thought of as not fully specified and a major consequence is that it cannot be used to predict $y_{i,t}$. Moreover, if the system is incoherent, the efficiency of the estimator of the dependence parameter β_i is unknown. Further, very often the parameters of incoherent models are rather set- than point-identified. Examples of articles where these conditions are discussed for different specifications of SEM are Blundell and Smith (1994); Gourieroux et al. (1980); Lewbel (2007) and Pick (2007).[4]

In our framework, the existence of a unique solution requires that the spectral radius[5] of the Jacobian matrix of the system be smaller than one. The Jacobian matrix $\nabla H(\boldsymbol{y}_t)$ of the two-equation ST-SEM can be written as follows:

$$\nabla H(\boldsymbol{y}_t) = \begin{pmatrix} 0 & G(y_{2,t};\gamma_1, c_2)[1 - G(y_{2,t};\gamma_1, c_2)]\beta_1\gamma_1 \\ G(y_{1,t};\gamma_2, c_1)[1 - G(y_{1,t};\gamma_2, c_1)]\beta_2\gamma_2 & 0 \end{pmatrix}$$

The spectral radius is given by the following formula:

$$\rho(\nabla H(\boldsymbol{y}_t)) = \left|\sqrt{G(y_{1,t};\gamma_1, c_2)\,[1 - G(y_{1,t};\gamma_1, c_2)]\,G(y_{2,t};\gamma_2, c_1)\,[1 - G(y_{2,t};\gamma_2, c_1)]\,\gamma_1\beta_1\gamma_2\beta_2}\right|$$

The condition highly depends on the combinations of γ_i and β_i values. The product of the first four terms is bounded with a maximum value of $1/2^4$ which corresponds to $G(y_{j,t};\gamma_i, c_j) = 1/2$. In this case the product $|\gamma_1\beta_1\gamma_2\beta_2|$ has to be smaller than 16 such that $\rho(\nabla H(\boldsymbol{y}_t)) < 1$. This is a sufficient but a non-necessary condition because if $G(y_{j,t};\gamma_i, c_j)$ is close to 0 or close to 1 for either of the equations, the product of the first four elements approaches 0, such that $|\gamma_1\beta_1\gamma_2\beta_2|$ can be higher than 16 as far as the spectral radius remains smaller than 1. Note that the product of the first four elements highly depends on γ_i and β_i, both directly and indirectly through $G(y_{j,t};\gamma_i, c_j)$ and $y_{j,t}$.

Fig. 2 shows examples of the graphical solution of our two-equation system for $\beta_i = 1$ and for different values of the parameter γ_i. On the left with $\gamma_i = 2.5$, the system has a unique solution. On the right, which corresponds to $\gamma_i = 10$, the system has three solutions. Putting a negative β_1 in the equation of y_1 would lead to a unique solution again. In summary, the number of solutions depends on the value of γ_i, β_i, $G(y_{j,t};\gamma_i, c_j)$ and $y_{j,t}$. Our system is coherent most likely for (but

[4] The incoherency issue also appears in Pesaran and Pick's model. The authors discuss the solution of this binary system and the possibility of multiple equilibria. For the values of $\boldsymbol{z}_t$, $\boldsymbol{x}_{i,t}$ and $u_{i,t}$ for which two solutions are possible, an index d_t which follows a Bernoulli distribution is introduced. The latter takes the value 1 if the favorable solution occurs (non-crisis times) and the value 0 for the unfavorable solution (crisis times). The reduced form depends on the value of d_t.

[5] $\rho(A)$ denotes the spectral radius of a matrix $A \in \mathbb{R}^{n\times n}$ and is defined by $\rho(A) = \max_{1\le i\le n} |\lambda_i|$ where $\lambda_1, \ldots, \lambda_n$ are all the eigenvalues of A. If $\rho(A) < 1$ this is equivalent to stating that the matrix A is convergent or that $\lim_{k\to\infty} A^k x = 0$, for $\forall x \in \mathbb{R}^n$. A square matrix $A \in \mathbb{R}^{n\times n}$ is convergent if $||A^k|| \to 0$ as $k \to \infty$, which is equivalent to $(A^k)_{i,j} \to 0$ as $k \to \infty$ for all i, j.

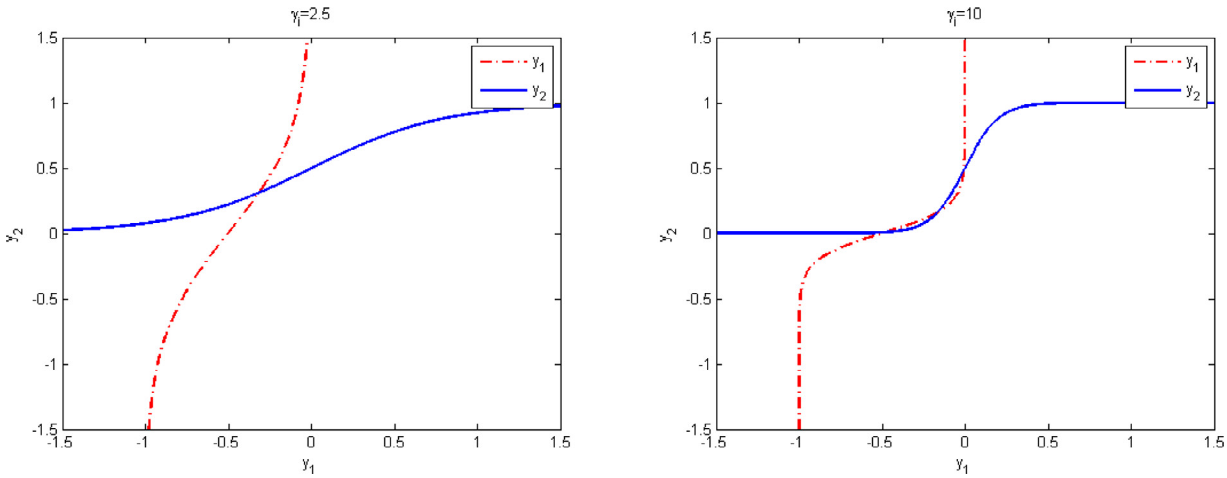

Fig. 2. Graphical solution of a two-equation system. *Notes*: $y_1 = -1 + \{1 + e^{-2.5y_2}\}^{-1}$ and $y_2 = \{1 + e^{-2.5y_1}\}^{-1}$ (left), $y_1 = -1 + \{1 + e^{-10y_2}\}^{-1}$ and $y_2 = \{1 + e^{-10y_1}\}^{-1}$ (right).

not limited to) low values of the smoothness parameter, that is why we focus on them for our Monte Carlo simulations. We solve the system numerically using Fixed Point and Newton's methods.

3. Estimation method and simulations

3.1. Estimation method

As the dependence variable $G(y_{j,t}; \gamma_i, c_j)$ is a nonlinear function of the endogenous variable $y_{j,t}$ and of the unknown parameters γ_i and c_j, we propose a nonlinear two-stage least squares procedure (NL2SLS) for estimating our system based on Kelejian (1971) and Amemiya (1974). Those authors show that the parameters of nonlinear SEM models can be consistently estimated by a NL2SLS procedure for each equation of the system. Kelejian (1971) performs a first stage regression of the nonlinear endogenous variable on a polynomial of the instruments, and notes that the accuracy of the estimator should improve as the order d of the polynomial increases. As mentioned earlier, the presence of exogenous equation-specific variables is necessary to ensure identification. Kelejian (1971) and Amemiya (1974) show the consistency of the estimates for the overall set of parameters and their asymptotic normality under usual regularity conditions. Appendix A.1 gives the limiting distribution of the ST-SEM.

Below we describe the NL2SLS estimation procedure for each equation of our two-equation system. Let us denote by $\boldsymbol{\psi}_i = [\boldsymbol{\delta}_i', \boldsymbol{\alpha}_i', \beta_i]'$ the set of parameters which appear linearly in Equation i, by $\mathbf{h}_{i,t} = [\boldsymbol{z}_t', \boldsymbol{x}_{i,t}', G(y_{j,t}; \gamma_i, c_j)]'$ the set of explanatory variables in this same equation for $i = 1, 2$, $i \neq j$, and by $\mathbf{w}_t = [\boldsymbol{z}_t', \boldsymbol{x}_{1,t}', \boldsymbol{x}_{2,t}']'$ the overall set of the exogenous variables of the system.

Step 1: Take initial values of γ_i and c_j denoted by $\hat{\gamma}_i$ and $\hat{c}_j$ and compute $\tilde{G}_{i,t} = G(y_{j,t}; \hat{\gamma}_i, \hat{c}_j)$.
Step 2: For the first stage of the NL2SLS, regress $\tilde{G}_{i,t}$ on a d-order polynomial of $\mathbf{w}_t$ using OLS method, and denote the fitted value by $\hat{\tilde{G}}_{i,t}$.
Step 3: For the second stage of the NL2SLS regress, using OLS, $y_{i,t}$ on $\hat{\mathbf{h}}_{i,t} = [\boldsymbol{z}_t', \boldsymbol{x}_{i,t}', \hat{\tilde{G}}_{i,t}]'$ to obtain an estimate of $\boldsymbol{\psi}_i$ denoted by $\hat{\boldsymbol{\psi}}_i$, compute the sum of squared residuals $SSR_i = \sum_{t=1}^{T} \hat{u}_{i,t}^2$.
Step 4: Repeat Steps 1 to 3, using an updating procedure for γ_i and c_j, until SSR_i is minimized.

The literature concerned with STAR models points out the difficulty of a precise estimation of the smoothness parameter γ_i, especially for sufficiently large values. Indeed, as γ_i increases the logistic function rapidly converges to a step function, and large variations of this parameter lead to slight changes in the shape of the function (see Fig. 1). To improve the estimator for γ_i, a large number of observations is required around c_j (see van Dijk et al., 2002). An accurate estimation of γ_i is even more difficult in our case because of the approximation of $G(y_{j,t}; \gamma_i, c_j)$ by a polynomial. Nevertheless, (Teräsvirta, 1994) and (Teräsvirta et al., 2010) argue that (i) a lack of significance of γ_i does not suggest that the model is linear, and (ii) a precise estimation of this parameter is not necessary for an accurate estimation of the other parameters of the model. In practice, γ_i is often bounded to a range of values not very close to zero and not too large; for instance, not higher than 100 (e.g. Schleer, 2015).

In threshold regression (TR) models the distribution of c_j is non-standard (see Hansen, 2000). In smooth transition regression (STR) models (with known γ_i), Andrews and Cheng (2013) show that the distribution of c_j depends on the value of $|\beta_i|$ and is also non-standard under no identification or weak identification. The parameter c_j is very often winsorized in the quantile range [15% – 85%] (e.g. Andrews (1993); González and Teräsvirta (2006)), or [20% – 80%] (e.g. Hansen (1996)). Indeed, for extreme values of c_j the logistic function only takes values that are either close to 0 or close to 1, which raises a collinearity issue with the constant term, and ill-behavior of the covariance matrix and of the test statistics.

We use the differential evolution (DE)[6] method to minimize the objective function with respect to γ_i and c_j, following a recent article by Maringer and Meyer (2008) who employ heuristic methods in STAR models. The DE heuristic approach was first developed by Storn and Price (1997) for the minimization of possibly nonlinear and non-differentiable functions. Many studies show that the DE method outperforms other single-agent or multi-agent methods (e.g. Maringer and Meyer (2008); Schleer (2015)).

The DE method has several advantages compared to gradient-based methods: (i) it is able to escape local minima through the presence of random components, (ii) it does not require a sophisticated initial solution. Compared to other heuristic methods, DE needs very little tuning. Crucial parameters are (i) the population size, which should be large enough to reflect diversity, (ii) the number of simulations to allow the algorithm to converge, (iii) the scaling factor F and (iv) the cross-over probability π. Based on the study by Maringer and Meyer (2008), we set $F = 0.5$, $\pi = 0.8$, the population size to $10D$, with D the dimension of the optimization problem, and 500 iterations.

3.2. Simulation results

We now study the properties of the estimation method presented above, using the below data-generating process and proceeding in three steps. First, we fix γ_i and c_j, then only γ_i is fixed and finally all the parameters of each equation are estimated jointly. For ease of interpretation, the distributions are centered around 0.

3.2.1. Data-generating process

For the Monte Carlo simulations, we generate the data using symmetric parameters for both equations of the system. This will allow us to only focus on the first equation, with symmetric results for the second equation. We use the following data-generating process:

$$y_{i,t}^{(r)} = \delta_{0,i} + \delta_{1,i} z_t^{(r)} + \alpha_i x_{i,t}^{(r)} + \beta_i G(y_{j,t}^{(r)}; \gamma_i, c_j) + u_{i,t}^{(r)} \quad i, j = 1, 2 \quad i \neq j \quad t = 1, \ldots, T$$

$$G(y_{j,t}^{(r)}; \gamma_i, c_j) = \left\{ 1 + e^{\gamma_i (y_{j,t}^{(r)} - c_j)} \right\}^{-1} \quad \gamma_i > 0$$

with $r = 1, \ldots, R$, the number of replications $R = 5,000$, and the number of observations $T = [100, 500]$. The error term follows a normal distribution $u_{i,t}^{(r)} \sim i.i.d.\mathcal{N}(0, 1)$ and the correlation between the error terms of the two equations takes three different values $\rho = corr(u_{1,t}, u_{2,t}) = [0, 0.5, 0.8]$.

We set $\delta_{0,i} = \delta_{1,i} = \alpha_i = 1$, while we try several values for the dependence parameter $\beta_i = [0, 0.3, 0.5, 1, 1.5]$ in order to carefully check the properties of the model. Given the sensitivity of the model with respect to the smoothness parameter, we test $\gamma_i = [0.5, 1.5, 2.5]$. The smallest value is close to the lower bound of 0, and given the low variation of the logistic function in this case, collinearity with the constant term may occur. On the other hand, the highest value is set such that the coherency of the model and numerical convergence of the system are guaranteed. Nevertheless, this should not be restrictive since a logistic function with $\gamma_i = 2.5$ is very close to a step function.

The location parameter is set at $c_j = 1$ and in the estimation procedure the values it can take are constrained to a range corresponding to the 20$^{\text{th}}$ and 80$^{\text{th}}$ quantiles of the distribution of $y_{j,t}$. The parameter c should not be far in the tails of the distribution because the dependence variable will have a disproportionately high number of values close to 0 or close to 1. In both cases, identification issues arise. Given that the generation of the values of $y_{j,t}$ depends on the value of c_j, the latter cannot be set in advance as a precise quantile of the former. Ex-post computations show that $c_j = 1$ is around the median. We also performed analyses with $c_j = 1.5$ with little change in the results.

The explanatory variables for each equation follow the autoregressive processes written below:

$$z_t^{(r)} = \xi z_{t-1}^{(r)} + \nu_t, \quad x_{i,t}^{(r)} = \phi_i x_{i,t-1}^{(r)} + (1 - \phi_i^2)^{1/2} \varepsilon_{i,t}^{(r)}$$

with the error term written as:

$$\varepsilon_{1,t}^{(r)} = \theta \varepsilon_{2,t}^{(r)} + \eta_t^{(r)}$$

and $\nu_t^{(r)}, \varepsilon_{2,t}^{(r)}, \eta_t^{(r)} \sim i.i.d.\mathcal{N}(0, 1)$, $\xi = \phi_i = [0, 0.5]$ and $\theta = corr(\varepsilon_{1,t}, \varepsilon_{2,t}) = [0, 0.3]$. If θ is different from 0, then the equation-specific variables are correlated, violating the identification condition. In empirical studies, it happens very frequently that the equation-specific variables are not independent of each other. Hence, it is useful checking the properties of the estimators and of the test statistics in this case.

Following this data-generating process, in Table 1 and Fig. 3 we show characteristics of the distribution of $y_{i,t}$, performing similar exercises to Pesaran and Pick (2007), Table 1 (page 1252) and Figs. 2, 3, 4 (pages 1253–1254). We find very

[6] We have also used gradient-based methods, grid search, threshold acceptance and simulated annealing. We found out that the DE method outperforms the others. This result is similar to Schleer (2015). A comparison of these methods in our framework could be the subject of a separate study.

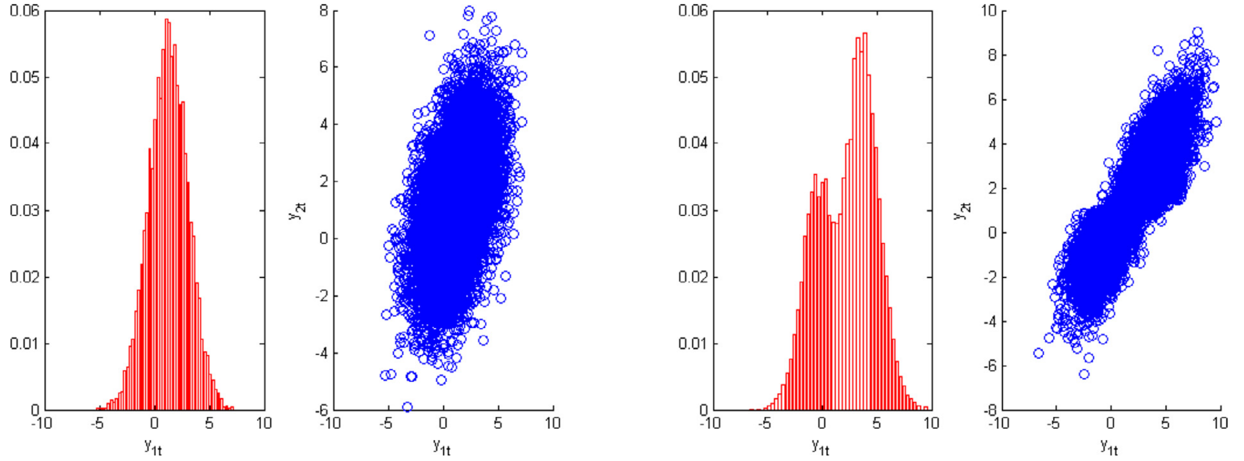

Fig. 3. Distribution of $y_{1,t}$ and scatterplot of $y_{1,t}$, $y_{2,t}$. $\delta_{0,i} = \delta_{1,i} = \alpha_i = 1$, $c_j = 1$, $T = 10000$ Left panel: $\beta_i = 0.5$, $\gamma_i = 1$, $\rho = 0$, $\phi_i = \xi = \theta = 0$. Right panel: $\beta_1 = 2$, $\beta_2 = 0.5$, $\gamma_1 = 6$, $\gamma_2 = 1$, $\rho = 0.8$, $\phi_i = \xi = 0.5$, $\theta = 0.3$.

similar patterns compared to the above article, notably (i) a non-zero correlation between $y_{1,t}$ and $y_{2,t}$ even when the error terms are not correlated, and (ii) bimodal distributions for $y_{i,t}$ associated with high values of the dependence and of the smoothness parameters.

Table 1 shows summary statistics for $y_{i,t}$ using different parameter combinations. It is important to note that even when $\rho = 0$, the correlation between $y_{1,t}$ and $y_{2,t}$ is almost always above 0.5 with a maximum value of 0.75. Thus, the analyses of comovement based only on the correlation between the error terms may be inaccurate since absence of correlation in the error term does not imply no linkages between markets. Naturally, $corr(y_{1,t}, y_{2,t})$ increases as β_i and ρ increase, with a value as high as 0.87 for $\beta_i = 1.5$ and $\rho = 0.8$. The model is also capable of producing skewness, which is a desirable property, especially when modeling financial series.

Fig. 3 shows the distribution of $y_{1,t}$ and a scatter plot between $y_{1,t}$ and $y_{2,t}$ for two combinations of β_i and γ_i: left panel $\beta_i = 0.5$, $\gamma_i = 1$, right panel $\beta_1 = 2$, $\gamma_1 = 6$, $\beta_2 = 0.5$ and $\gamma_2 = 1$. In the left panel, the distribution of $y_{1,t}$ has a regular shape and so does the cloud of points in the scatter plot. In the right panel the distribution of $y_{1,t}$ is bimodal. A further analysis showed that the bimodality is visible as γ_i becomes larger, while for low values of γ_i and large β_i the distribution is well-behaved. If bimodality is present, it will strengthen as the correlation between the error terms ρ becomes larger.

3.2.2. Known smoothness and location parameters

Table 2 shows the bias and the root mean squared error (RMSE) for the dependence parameter β_1 with degree of the polynomial $d = 3$. We performed this analysis for the degree of the polynomial d from 1 to 6. To save space, we only show the results for $d = 3$. The first column of the table shows the different values taken by $\beta_i = [0, 0.3, 0.5, 1, 1.5]$. In the left panel there is no structure in the regressors, while in the right panel, correlation and autocorrelation are introduced. We use $T = 100$ and $T = 500$.

We can see that the bias becomes smaller as the correlation between the error terms decreases, and as the sample size or the smoothness parameter increases. A similar pattern can be observed for the RMSE which is low for $T = 500$ and $\gamma_i > 0.5$. Overall, we find that low degrees of the polynomial have better properties in terms of bias and RMSE, than higher-order polynomials. The simplest case with $d = 1$ produces very accurate results. Indeed, as the order of the polynomial increases, the number of parameters to be estimated during the first stage of the estimation procedure becomes larger.

Fig. 4 shows the distribution of the dependence parameter $\beta_1 = 1$ for $T = 100$ in the left panel and for $T = 500$ in the right panel. In the upper panel there is no structure in the regressors and no correlation between the error terms, while in the lower panel the explanatory variables have an AR(1) structure, are correlated between equations, and so are the error terms. The distributions are symmetric except for the lower left panel which has a longer left tail. The variance is considerably lower for the larger sample.

In our framework, it is not only important to accurately estimate the dependence parameter, but also the parameters related to the common and to the equation-specific variables. The simulation results show that the parameters related to z_t and $x_{i,t}$ have symmetric and narrow distributions; the distributions of the constant term and of the variance of the residuals are somewhat flatter, with a longer right tail for the latter.

In general, there is evidence that the NL2SLS estimation procedure is able to provide accurate estimates for all the parameters of each equation of the ST-SEM. In empirical applications, γ_i could be set to reasonable values, while c_j could be fixed to a quantile in the tails of the distribution of y_j. Low order polynomials seems to be a good choice to approximate $G(y_{j,t}; \gamma_i, c_j)$, especially for small samples.

Table 1
Moments of the distribution of y_1 from the simulation results.

$\boldsymbol{\beta_1}$	$\bar{y}_1$	$\sigma(y_1)$	Kurt	Skew	Corr	$\bar{y}_1$	$\sigma(y_1)$	Kurt	Skew	Corr
	$\gamma_1 = 1$					$\gamma_1 = 2$				
$\boldsymbol{\xi = 0,\ \phi_i = 0,\ \theta = 0}$										
$\boldsymbol{\rho = 0}$										
0.5	1.27	1.80	3.0	-0.06	0.47	1.26	1.83	2.88	-0.01	0.49
1	1.57	1.90	2.88	-0.07	0.59	1.59	1.94	2.86	-0.03	0.61
1.5	2.00	2.00	2.83	-0.12	0.65	2.03	2.07	2.73	-0.16	0.67
$\boldsymbol{\rho = 0.5}$										
0.5	1.26	1.81	2.90	−0.04	0.60	1.26	1.86	2.85	0.01	0.63
1	1.60	1.92	2.88	−0.08	0.69	1.61	1.99	2.78	−0.08	0.71
1.5	1.98	2.03	2.78	−0.15	0.75	1.99	2.16	2.64	−0.22	0.77
$\boldsymbol{\rho = 0.8}$										
0.5	1.28	1.83	2.92	−0.03	0.69	1.25	1.86	2.95	−0.01	0.71
1	1.60	1.95	2.86	−0.05	0.76	1.62	2.03	2.75	−0.08	0.78
1.5	1.90	2.06	2.82	−0.14	0.80	2.00	2.15	2.63	−0.20	0.81
$\boldsymbol{\xi = 0.5,\ \phi_i = 0.5,\ \theta = 0.3}$										
$\boldsymbol{\rho = 0}$										
0.5	1.31	1.92	2.94	0.032	0.59	1.26	1.98	2.88	−0.05	0.62
1	1.55	2.05	2.87	−0.01	0.67	1.58	2.10	2.78	−0.03	0.70
1.5	1.93	2.15	2.80	−0.09	0.74	2.00	2.27	2.70	−0.17	0.75
$\boldsymbol{\rho = 0.5}$										
0.5	1.25	1.94	2.85	−0.03	0.72	1.24	1.99	2.85	−0.01	0.73
1	1.56	2.08	2.84	−0.02	0.77	1.62	2.13	2.76	−0.05	0.79
1.5	2.02	2.21	2.71	−0.13	0.82	1.95	2.30	2.56	−0.19	0.83
$\boldsymbol{\rho = 0.8}$										
0.5	1.24	1.95	3.00	−0.04	0.78	1.22	1.98	2.79	−0.03	0.79
1	1.60	2.08	2.84	−0.07	0.83	1.60	2.13	2.73	−0.10	0.84
1.5	1.94	2.24	2.74	−0.13	0.86	2.01	2.35	2.52	−0.21	0.87

Note: This table presents summary statistics for y_1: $\bar{y}_1$ the sample average, $\sigma(y_1)$ the standard error, 'Kurt' the kurtosis of y_1, 'Skew' the skewness and 'Corr' the correlation between y_1 and y_2. Given that the system is symmetric, we present the results only for y_1. The data-generating process is presented in Section 3.2.1 with $T = 10000$.

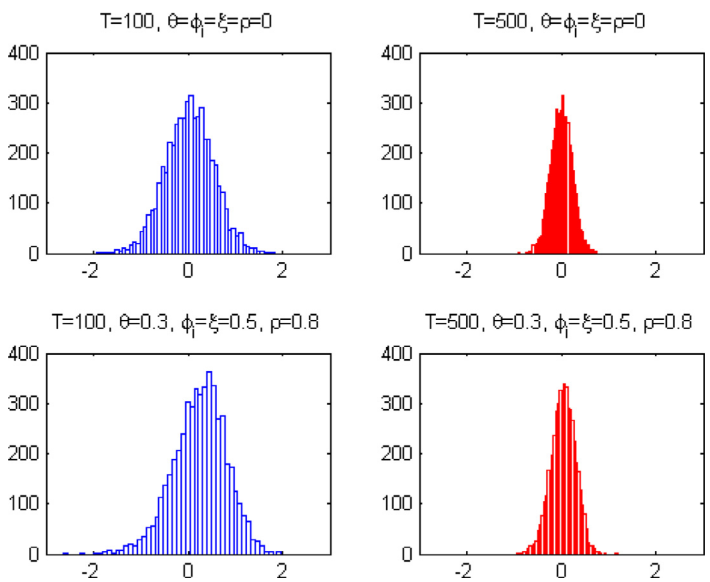

Fig. 4. Distribution of β_1 for *known* γ and *c*. *Notes*: This figure shows the distribution of β_1 obtained from simulations for the following parameter combinations: $\beta_i = 1$, $\gamma_i = 2.5$, $d = 3$, and 5000 simulations. The parameter is centered around 0.

3.2.3. *Known smoothness parameter and unknown location parameter*

Table 3 displays the bias and the RMSE for the dependence parameter β_1 when *only* γ_i is fixed. Given that the numerical procedure with 5000 simulations is time-consuming, we only present the results for $\beta_1 = 1$. We have tried other values of β_i for a smaller number of simulations with similar conclusions. The structure of the table remains the same as above,

Table 2
Bias and root mean squared error for β_1 with *known* γ and c, $d = 3$.

	$\xi = 0, \phi_i = 0, \theta = 0$						$\xi = 0.5, \phi_i = 0.5, \theta = 0.3$					
β_1	T=100			T=500			T=100			T=500		
γ	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5
$\rho = 0$												
Bias												
0	0.006	−0.004	−0.003	0.001	−0.002	0.000	−0.006	−0.005	0.003	0.004	−0.002	0.001
0.3	0.008	−0.004	0.012	0.003	−0.002	−0.004	0.002	−0.001	0.012	0.002	−0.002	−0.002
0.5	0.022	0.010	0.005	0.014	0.002	0.005	0.030	0.007	0.011	0.017	0.001	0.004
1	0.045	0.051	0.030	0.010	0.008	0.001	0.045	0.061	0.049	0.012	0.011	0.004
1.5	0.060	0.040	0.049	−0.004	0.007	0.006	0.051	0.055	0.056	−0.007	0.009	0.011
RMSE												
0	0.942	0.525	0.455	0.419	0.235	0.206	1.025	0.575	0.514	0.442	0.253	0.223
0.3	0.936	0.530	0.472	0.423	0.235	0.214	1.023	0.582	0.522	0.444	0.255	0.235
0.5	0.951	0.537	0.474	0.420	0.235	0.214	1.026	0.585	0.527	0.447	0.260	0.236
1	0.949	0.554	0.512	0.414	0.247	0.226	1.025	0.612	0.565	0.444	0.270	0.248
1.5	0.991	0.582	0.532	0.428	0.261	0.245	1.055	0.634	0.579	0.455	0.282	0.267
$\rho = 0.5$												
Bias												
0	0.221	0.099	0.082	0.042	0.017	0.016	0.240	0.111	0.101	0.049	0.018	0.018
0.3	0.224	0.098	0.102	0.044	0.017	0.012	0.244	0.116	0.116	0.048	0.019	0.016
0.5	0.240	0.116	0.099	0.057	0.022	0.021	0.280	0.129	0.118	0.063	0.022	0.023
1	0.266	0.163	0.134	0.052	0.029	0.020	0.296	0.183	0.166	0.057	0.034	0.024
1.5	0.286	0.162	0.163	0.038	0.030	0.028	0.305	0.185	0.181	0.038	0.033	0.033
RMSE												
0	0.956	0.526	0.456	0.419	0.235	0.206	1.034	0.576	0.517	0.442	0.253	0.223
0.3	0.949	0.536	0.481	0.425	0.237	0.216	1.037	0.590	0.534	0.445	0.257	0.237
0.5	0.969	0.547	0.487	0.423	0.238	0.219	1.049	0.598	0.541	0.449	0.264	0.241
1	0.968	0.577	0.533	0.418	0.253	0.233	1.048	0.633	0.592	0.448	0.278	0.257
1.5	1.010	0.605	0.567	0.431	0.270	0.258	1.071	0.663	0.619	0.458	0.291	0.282
$\rho = 0.8$												
Bias												
0	0.349	0.160	0.134	0.067	0.028	0.025	0.389	0.180	0.159	0.076	0.030	0.029
0.3	0.354	0.162	0.157	0.069	0.028	0.022	0.390	0.188	0.180	0.075	0.031	0.027
0.5	0.371	0.183	0.159	0.082	0.034	0.032	0.430	0.205	0.187	0.090	0.035	0.035
1	0.399	0.233	0.204	0.077	0.043	0.033	0.448	0.260	0.243	0.084	0.049	0.038
1.5	0.423	0.242	0.240	0.063	0.045	0.043	0.459	0.272	0.266	0.066	0.048	0.049
RMSE												
0	0.974	0.530	0.461	0.420	0.235	0.206	1.050	0.580	0.522	0.443	0.254	0.224
0.3	0.966	0.542	0.491	0.427	0.238	0.218	1.060	0.600	0.546	0.447	0.258	0.239
0.5	0.989	0.558	0.499	0.426	0.241	0.222	1.071	0.611	0.556	0.452	0.266	0.245
1	0.993	0.596	0.552	0.421	0.258	0.239	1.075	0.651	0.614	0.451	0.283	0.264
1.5	1.032	0.625	0.597	0.434	0.276	0.267	1.096	0.687	0.651	0.460	0.298	0.293

Note: Bias and the *root mean squared error* (RMSE) of the dependence coefficient β_1 for a model with *known* γ_i and c_j. The system is symmetric, that is why we present the results only for the first equation. We generate 5000 simulations from the data generating process in Section 3.2.1 with $d = 3$.

except that the first column gives the order of the polynomial.[7] The bias is negligible for $d = 1$, but it rapidly increases from $d = 2$, especially for $T = 100$. Again, the bias is considerably lower as the sample size increases with $\gamma_i = 2.5$, and it deteriorates with ρ. Generally, the RMSE reaches lower values for higher-order polynomials.

Fig. 5 shows the distribution of β_1 for $d = 1$ (left) and $d = 3$ (right). Compared to the case where c_j is known, the distributions are somewhat flatter but they have a nice shape for $T = 500$. In the small sample case, the left tail is longer and has a bump. The distributions of the other parameters are comparable to the case where c is known.

Fig. 6 (left) shows the distribution of c_2 for different parameter combinations. As discussed previously, an accurate estimation of this parameter is difficult to be obtained. Although, the distribution does not seem to be normal, our results are comparable to the results of Schleer (2015). In our framework, the optimization is more complex since it depends on the accuracy of the approximation of the logistic distribution by a polynomial in the first step of the NL2SLS.

3.2.4. Unknown smoothness and location parameters

The estimation results with an unknown γ_i are virtually the same as when γ_i is known, that is why we do not comment further on them. This finding is in line with the argument of Teräsvirta et al., 2010 that a precise estimation of this parameter is not necessary for an accurate estimation of the other parameters of the model.

[7] Given the high RMSE seen previously, we do not consider the value of $\gamma_i = 0.5$.

Table 3
Bias and root mean squared error for β_1 with *known* γ and *unknown* c.

	$\beta_1 = 1$		$\xi = 0, \phi_i = 0, \theta = 0$				$\xi = 0.5, \phi_i = 0.5, \theta = 0.3$			
			T=100		T=500		T=100		T=500	
		γ	1.5	2.5	1.5	2.5	1.5	2.5	1.5	2.5
$\rho = 0$	*Bias*									
	$d = 1$		0.040	0.086	0.041	0.065	0.054	0.116	0.046	0.067
	$d = 2$		0.194	0.194	0.083	0.075	0.177	0.179	0.076	0.057
	$d = 3$		0.235	0.202	0.093	0.081	0.216	0.226	0.085	0.072
	$d = 4$		0.236	0.212	0.091	0.076	0.258	0.228	0.085	0.067
	$d = 5$		0.263	0.230	0.089	0.074	0.246	0.233	0.084	0.078
	$d = 6$		0.266	0.251	0.103	0.074	0.272	0.251	0.094	0.079
	RMSE									
	$d = 1$		0.660	0.685	0.288	0.302	0.758	0.796	0.326	0.337
	$d = 2$		0.683	0.637	0.277	0.260	0.774	0.695	0.297	0.276
	$d = 3$		0.666	0.614	0.284	0.259	0.732	0.677	0.293	0.270
	$d = 4$		0.650	0.596	0.273	0.252	0.727	0.651	0.290	0.266
	$d = 5$		0.643	0.582	0.278	0.248	0.699	0.622	0.283	0.265
	$d = 6$		0.619	0.568	0.279	0.254	0.681	0.610	0.288	0.263
$\rho = 0.5$	*Bias*									
	$d = 1$		0.014	0.034	0.045	0.045	0.017	0.028	0.040	0.041
	$d = 2$		0.249	0.241	0.093	0.086	0.246	0.224	0.083	0.069
	$d = 3$		0.341	0.034	0.114	0.096	0.361	0.314	0.104	0.092
	$d = 4$		0.439	0.403	0.128	0.103	0.455	0.405	0.125	0.105
	$d = 5$		0.521	0.453	0.145	0.122	0.543	0.473	0.137	0.117
	$d = 6$		0.584	0.510	0.164	0.138	0.611	0.532	0.158	0.129
	RMSE									
	$d = 1$		0.698	0.685	0.300	0.300	0.773	0.805	0.335	0.329
	$d = 2$		0.703	0.672	0.287	0.275	0.791	0.770	0.305	0.279
	$d = 3$		0.710	0.685	0.289	0.272	0.780	0.713	0.303	0.286
	$d = 4$		0.728	0.668	0.292	0.270	0.784	0.716	0.304	0.285
	$d = 5$		0.755	0.680	0.302	0.276	0.806	0.723	0.310	0.284
	$d = 6$		0.788	0.701	0.305	0.279	0.842	0.742	0.318	0.290
$\rho = 0.8$	*Bias*									
	$d = 1$		0.008	−0.003	0.026	0.026	-0.023	-0.025	0.022	0.021
	$d = 2$		0.266	0.238	0.101	0.084	0.268	0.258	0.094	0.074
	$d = 3$		0.429	0.381	0.125	0.110	0.447	0.377	0.118	0.092
	$d = 4$		0.538	0.495	0.148	0.130	0.568	0.514	0.146	0.118
	$d = 5$		0.660	0.586	0.178	0.150	0.685	0.629	0.167	0.137
	$d = 6$		0.753	0.672	0.201	0.172	0.809	0.694	0.191	0.160
	RMSE									
	$d = 1$		0.708	0.681	0.297	0.294	0.803	0.809	0.341	0.334
	$d = 2$		0.721	0.682	0.296	0.275	0.812	0.767	0.309	0.285
	$d = 3$		0.742	0.682	0.298	0.275	0.810	0.735	0.310	0.286
	$d = 4$		0.779	0.720	0.308	0.284	0.834	0.775	0.315	0.289
	$d = 5$		0.843	0.764	0.314	0.290	0.891	0.810	0.324	0.293
	$d = 6$		0.899	0.814	0.331	0.300	0.971	0.857	0.333	0.302

Note: Bias and of the *root mean squared error* (RMSE) for the dependence parameter β_1 with *unknown* location parameter c_1 and *known* smoothness parameter γ_1. We focus on the first equation, but the second equation is symmetric. The true value of $\beta_1 = 1$ with 5000 simulations. Refer to Section 3.2.1 for the data generating process.

Fig. 6 (right) displays the distribution of γ_1 for $d = 1$. The values of γ_1 do not explode and the mean of the simulations is around 2.64, quite close to the true value of 2.5. The only constraint that we put on γ_i in the estimation procedure is that it has to be greater than 1 in order to avoid collinearity issues with the error term. For higher orders of the polynomial, the algorithm selects very frequently values close to the lower bound.

These results corroborate our above discussion that a precise estimation of the smoothness parameter is difficult due to a flat objective function with respect to this parameter. Studies in a time series framework such as Lundbergh and Teräsvirta (2006) and Andrews and Cheng (2013) fix the value of this parameter to 300 and 10, respectively. Schleer (2015) obtains similar results that the ones presented here.

4. Testing

4.1. Testing in the presence of nuisance parameters

In our framework, testing for dependence between the two endogenous variables translates into testing for simultaneity of a particular form. Let us focus on the first equation of the two-equation ST-SEM given by Eq. 3. The test for dependence

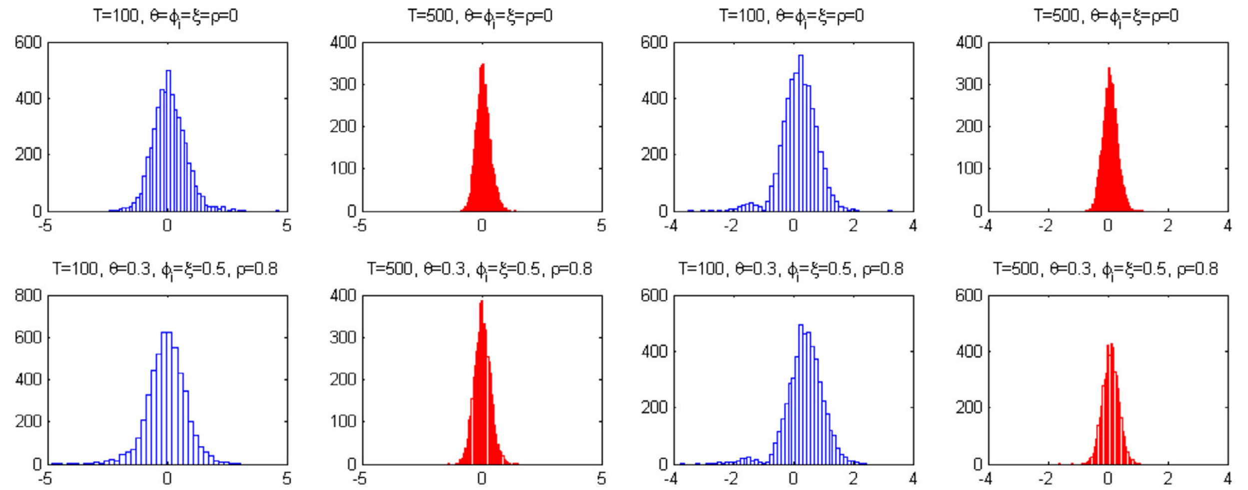

Fig. 5. Distribution of β_1 with *known* γ and *unknown c*. *Notes*: Centered distribution of β_1 obtained from 5000 simulations with the following parameter combinations: $\beta_i = 1$, $\gamma_i = 2.5$, $d = 1$ (left), $d = 3$ (right).

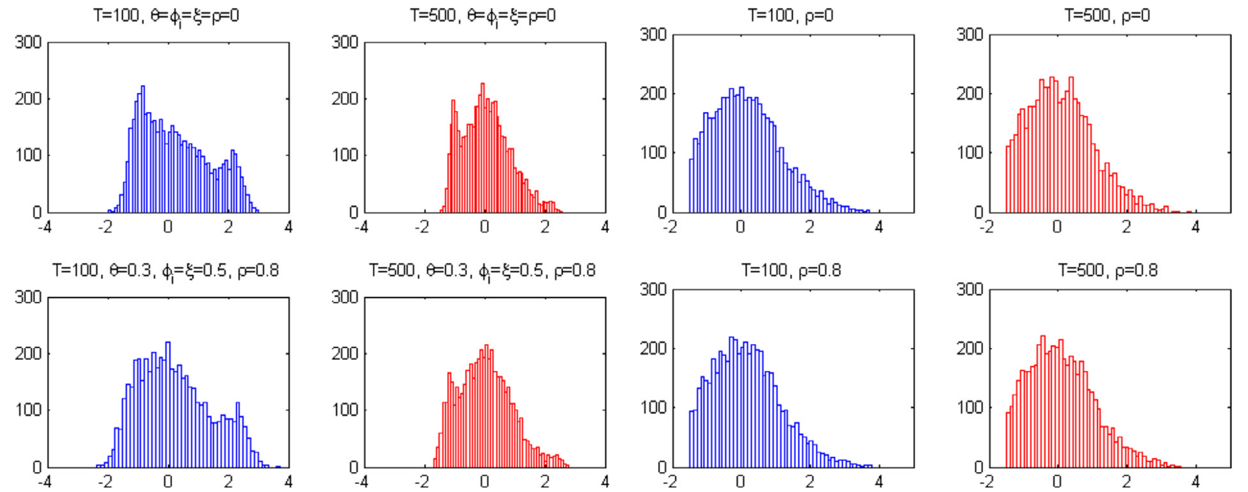

Fig. 6. Distribution of c_2 (left, $d = 3$, *known* γ_1) and distribution of γ_1 (right, $d = 1$). *Notes*: Centered distribution of c_2 (left, with known γ_1 and $d = 3$) and distribution of γ_1 (right, $d = 1$) obtained from simulations with the following parameter combinations: $\beta_i = 1$, $c_j = 1$, $\gamma_i = 2.5$ and 5000 simulations.

is equivalent to testing the null hypothesis that β_1 is equal to 0 ($H_0 : \beta_1 = 0$) against the alternative $H_a : \beta_1 \neq 0$. Under H_0, the smoothness parameter γ_1 and the threshold parameter c_2 are not identified and can take any value. The null hypothesis can alternatively be set as $H_0' : \gamma_1 = 0$, in which case the logistic function becomes a constant. In this case, neither β_1, nor c_2 are identified. The unknown parameters under the null hypothesis are *nuisance parameters*. Their presence under the null hypothesis renders the distribution of conventional tests non-standard. The analytical expressions of these tests are often unavailable and simulation techniques must be used (van Dijk et al., 2002).

Each of the above two null hypotheses leads to a different testing strategy, both being studied in a time-series context. The first class of tests based on $H_0 : \beta_1 = 0$ was initiated by Davies (1977, 1987) and taken over by Andrews (1993); Andrews and Ploberger (1994) and Hansen (1996, 2000). Andrews (1993); Davies (1977, 1987) and Andrews and Ploberger (1994) develop supremum, average and exponential average LM test statistics to solve the identification problem. Hansen (1996, 2000) derived the asymptotic distribution of the tests for TR models and suggested a sampling method to obtain critical values. The second class of tests based on $H_0' : \gamma_1 = 0$ versus $H_a' : \gamma_1 > 0$, first proposed by Luukkonen et al. (1988) (LST), sidesteps the nuisance parameters through a Taylor-series expansion of first and third order[8] around $\gamma = 0$. The resulting test statistic has a standard χ^2 distribution and is widely used in STR models.

[8] The test based on the third-order Taylor series expansion is found to be more powerful in models where the intercept changes regime, which is the case in our model.

González and Teräsvirta (2006) examine both types of tests in a Monte Carlo study and come to the conclusion that their performance is comparable, although average tests have slightly better size and power properties compared to the LST test. Nevertheless the latter test performs as good as the most frequently used supremum test, while having the advantage of being easy to compute. In addition, Teräsvirta (1994) argues that the LST test can be seen as a general test of nonlinearity since it has power against other nonlinear additive models which yield the same auxiliary regression.

We adapt the LST test to our framework in the presence of an endogenous variable in the logistic function. Applying the first-order Taylor approximation to $G(y_{2t};\gamma_1,c_2)$ around $\gamma_1=0$ yields the following auxiliary regression:[9]

$$y_{1,t}=\tilde{\boldsymbol{\delta}}_1'\boldsymbol{z}_t+\boldsymbol{\alpha}_1'\boldsymbol{x}_{1,t}+\tilde{\beta}_1 y_{2,t}+u^*_{1,t} \tag{6}$$

where $\tilde{\boldsymbol{\delta}}_1=\left[(\delta_{1,1}-\frac{1}{4}\beta_1\gamma_1c_2)\ \boldsymbol{\delta}_1^{-\prime}\right]'$, with $\boldsymbol{\delta}_1^-$ equal to $\boldsymbol{\delta}_1$ minus the constant term $\delta_{1,1}$, $\tilde{\beta}_1=\frac{1}{4}\beta_1\gamma_1$ and $u^*_{1,t}=\beta_1R_{1,1}+u_{1,t}$. The above model is free of nuisance parameters and the standard asymptotic theory is applicable since under the null $u^*_{1,t}=u_{1,t}$. Note that this first-order Taylor approximation leads to a test for simultaneity within a linear SEM model. The null hypothesis of $\gamma_1=0$ can be equivalently set as $\tilde{\beta}_1=0$. The alternative hypothesis of $\tilde{\beta}_1\neq 0$ necessarily implies that $\gamma_1>0$ and $\beta_1\neq 0$.

Similarly, the third-order Taylor expansion can be written as:

$$y_{1,t}=\tilde{\boldsymbol{\delta}}_1'\boldsymbol{z}_t+\boldsymbol{\alpha}_1'\boldsymbol{x}_{1,t}+\tilde{\beta}_{1,1}y_{2,t}+\tilde{\beta}_{1,2}y^2_{2,t}+\tilde{\beta}_{1,3}y^3_{2,t}+u^*_{1,t} \tag{7}$$

where $\tilde{\boldsymbol{\delta}}_1=[(\delta_{1,1}-\frac{1}{4}\gamma_1c_2+\frac{1}{48}\gamma_1^3c_2^3)\ \boldsymbol{\delta}_1^{-\prime}]'$, with $\boldsymbol{\delta}_1^-$ defined as above, $\tilde{\beta}_{1,1}=\beta_1\gamma_1(\frac{1}{4}-\frac{1}{16}\gamma_1^2c_2^2)$, $\tilde{\beta}_{1,2}=\frac{1}{16}\beta_1\gamma_1^3c_2$, $\tilde{\beta}_{1,3}=-\frac{1}{48}\beta_1\gamma_1^3$ and $u^*_{1,t}=\beta_1R_{3,1}+u_{1,t}$. The null hypothesis of no simultaneity can be written as $H_0:\tilde{\beta}_{1,1}=\tilde{\beta}_{1,2}=\tilde{\beta}_{1,3}=0$, while under the alternative at least one of these three parameters is different from zero. In this case, β_1 and γ_1 must be both non-zero since they are multiplicative.

Wooldridge (2002) pages 98–99, shows that the conventional LM statistic for 2SLS does not have a known limiting distribution and can even become negative. He proposes a modified version of this test that has a χ^2 square asymptotic distribution. Below we give the details of the implementation procedure of the LST test.

(i) Regress $y_{1,t}$ on $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}']'$ using the OLS[10] estimation procedure under the assumption of exogenous regressors. Take the residual $\hat{u}_{1,t}$, $t=1,\ldots,T$ and compute the residual sum of squares $\hat{SSE}_0=\sum_{t=1}^T\hat{u}^2_{1,t}$.

(ii) Regress $y_{1,t}$ on $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}'\ y_{2,t}]'$ (for the third-order expansion the set of regressors becomes $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}'\ y_{2,t}\ y^2_{2,t}\ y^3_{2,t}]'$) using the 2SLS method and $\mathbf{w}_t$ (squared and cubic values for the third-order expansion) as set of instruments. Take the residual denoted by $\hat{u}^*_{1,t}$, and compute the residual sum of squares $SSE_1=\sum_{t=1}^T\hat{u}^{*2}_{1,t}$.

(iii) Compute $\hat{SSE}_1$ which is the sum of squared residuals from the unrestricted second-stage regression of $y_{1,t}$ on $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}'\ \hat{y}_{2,t}]'$, with $\hat{y}_{2,t}$ the fitted value from the first stage regression of $y_{2,t}$ on $\mathbf{w}_t$ (third-order approximation: $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}'\ \hat{y}_{2,t}\ \hat{y^2}_{2,t}\ \hat{y^3}_{2,t}]'$).

(iv) Compute the test statistic as: $\text{LM}=T(\hat{SSE}_0-\hat{SSE}_1)/SSE_1$.

The above LM test asymptotically follows a χ^2_k, with $k=1$ for the first-order Taylor expansion and $k=3$ for the third-order. In small samples, the F-test statistic has better size properties. In addition, Teräsvirta (1994) documents that the F-test has power against the limiting case where the dependence variable is a dummy. Note that the testing procedure does not require the estimation of the nonlinear ST-SEM, which can be computationally costly.

4.2. Testing: Simulation results

In this part we present simulation results related to the tests for simultaneity with null hypothesis $\beta_i=0$ against the alternative $\beta_i\neq 0$, for different parameter combinations. For known smoothness and location parameters of the logistic function, standard test procedures can be applied. When these parameters are unknown, the LST test described above is used. We set completely symmetric parameters for both equations, that is why we only show the results for the first equation.

4.2.1. Known parameters of the logistic function

We comment on the simulation results of the tests for a system with known γ_i and c_j. The observed value of the test is compared to the 95% quantile of the t distribution with the corresponding degrees of freedom. In Table 4 the degree of the polynomial used to approximate the logistic function in the first stage of the NL2SLS is $d=1$.

The size of the test is given by the first row of each horizontal block with different values of the parameter ρ. For $\rho=0$ the empirical size is very close to the nominal size of 5%. The results appear to be marginally better for $T=500$. As ρ becomes larger, the tests start being slightly oversized, especially for $T=100$ and for low values of γ. On the contrary, the specifications with autocorrelated regressors and $corr(x_{1,t};x_{2,t})=0.3$ do not produce sensibly different results.

[9] The details of the Taylor approximation are given in Appendix A.2.

[10] In the test suggested by Wooldridge (2002), $\hat{SSE}_0$ is the restricted sum of squared residuals from the *second* stage of **2SLS** for $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}']'$ on a valid set of instruments (here $\mathbf{w}_t$). Given that $[\boldsymbol{z}_t'\ \boldsymbol{x}_{1,t}']'$ is assumed to be an exogenous set of variables, the fitted value of these regressors from the first-stage 2SLS is equal to their true value, and consequently OLS is equivalent to 2SLS.

Table 4
Size and power of the test for simultaneity (β_1) with *known* γ and c, $d = 1$.

	$\xi = 0, \phi_i = 0, \theta = 0$						$\xi = 0.5, \phi_i = 0.5, \theta = 0.3$					
β_1	T=100			T=500			T=100			T=500		
γ	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5
$\rho = 0$												
0	0.046	0.049	0.043	0.050	0.050	0.045	0.046	0.047	0.041	0.051	0.048	0.050
0.3	0.094	0.126	0.140	0.181	0.342	0.405	0.090	0.127	0.127	0.166	0.300	0.351
0.5	0.131	0.239	0.272	0.328	0.661	0.728	0.124	0.204	0.237	0.291	0.600	0.652
1	0.271	0.524	0.580	0.755	0.986	0.991	0.251	0.460	0.506	0.704	0.966	0.973
1.5	0.461	0.735	0.761	0.961	0.999	1.000	0.421	0.659	0.666	0.941	0.997	0.996
$\rho = 0.5$												
0	0.062	0.063	0.058	0.054	0.057	0.054	0.060	0.062	0.058	0.056	0.055	0.055
0.3	0.117	0.150	0.162	0.195	0.347	0.405	0.109	0.151	0.147	0.177	0.308	0.355
0.5	0.154	0.257	0.287	0.335	0.642	0.704	0.150	0.230	0.257	0.300	0.587	0.631
1	0.293	0.509	0.553	0.734	0.973	0.976	0.276	0.453	0.495	0.688	0.946	0.951
1.5	0.460	0.690	0.711	0.944	0.998	0.997	0.430	0.625	0.617	0.920	0.991	0.989
$\rho = 0.8$												
0	0.070	0.070	0.068	0.059	0.061	0.057	0.071	0.070	0.064	0.060	0.060	0.059
0.3	0.127	0.163	0.174	0.200	0.351	0.405	0.120	0.165	0.161	0.185	0.311	0.355
0.5	0.168	0.269	0.298	0.338	0.630	0.689	0.165	0.242	0.267	0.306	0.577	0.613
1	0.303	0.503	0.541	0.725	0.963	0.966	0.287	0.449	0.487	0.677	0.933	0.933
1.5	0.460	0.662	0.677	0.934	0.996	0.993	0.435	0.609	0.593	0.909	0.986	0.982

Note: Rejection frequency of the null hypothesis of no dependence (H_0: $\beta_i = 0$) using standard testing procedures (we perform the F test). *d=1* denotes the degree of the polynomial in the first step of the NL2SLS method. We fix the value of γ_i and c_j and use the data-generating process in Section 3.2.1. The results are presented only for Eq. 1 since they are symmetric. We run 5000 simulations.

The power of the test is shown from the second to the fifth row of each horizontal block with $\beta_i = [0.3,\ 0.5,\ 1,\ 1.5]$. The power is quite weak for $\beta_i = 0.3$, but rapidly improves as β_i increases. The power is also weak if $\gamma_i = 0.5$, but it is around the double already for $\gamma_i = 1.5$. The results are sensibly better for longer time series with figures which double or almost triple when switching from $T = 100$ to $T = 500$. The power is very close to 1 for $\beta_i \geq 1$, $T = 500$ and $\gamma_i > 0.5$.

In summary, to obtain good size properties, a linear approximation of the logistic function is sufficient, while the power slightly improves as the degree of the polynomial increases. It is important to note the crucial role of the smoothness parameter. We also obtained the results for $d = 2, \ldots, 6$, which we do not show here to save space. While the power of the test improves, the test is highly oversized, especially for $T = 100$. Given that the test performs well for low-order polynomials, we think that in practice it is not necessary to go further than a polynomial of degree 3.

4.2.2. Unknown parameters of the logistic function

We now discuss the testing results in the case where the smoothness and the location parameters are unknown. In Table 5, an auxiliary regression is constructed based on the first-order Taylor series approximation of the logistic function. The structure of the table is the same as above. The tests have very good size properties for any parameter combinations. Similarly as above, the power improves with the number of observations and with the smoothness parameter. Introducing a structure in the regressors seems to have a slight negative impact, although this is not always the case. The correlation in the error term does not seem to have an impact, neither on the size nor on the power. In Table 6 we perform the same exercise with the only difference that the dependence variable is approximated by a third-order expansion. The results are less precise compared to a linear approximation; the test is undersized and the power weakens.

Given that the parameters of the two-equation system are completely symmetric, when testing $H_0 : \beta_1 = 0$, the value of β_2 is also equal to 0. We also examined how the properties of the test for $H_0 : \beta_1 = 0$ depend on the value β_2. For this reason, we computed the size of the test for β_1 with $\beta_2 = [0,\ 0.3,\ 0.5,\ 1,\ 1.5]$ and find very good results for all the parameter combinations.

In a nutshell, the testing procedure proposed by Luukkonnen et al. (1988) works also very well in a model where endogeneity is present. In our framework, the simplest linear approximation seems to perform better than the heavier third-order approximation. In all cases, the test is simple and does not require the estimation of the ST-SEM which is computationally more demanding.

5. Comovement: Sovereign and banking sectors

We use the ST-SEM to investigate the comovement between the sovereign and the banking sectors of nine developed countries, namely Germany, France, Greece, Italy, Portugal, Spain, Switzerland, the United Kingdom and the United States.

Table 5
Size and power of the test for simultaneity (β_1) with *unknown* γ and c – *first-order Taylor approximation.*

	$\xi=0, \phi_i=0, \theta=0$						$\xi=0.5, \phi_i=0.5, \theta=0.3$					
β_1	T=100			T=500			T=100			T=500		
γ	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5
$\rho=0$												
0	0.050	0.051	0.041	0.048	0.052	0.046	0.050	0.045	0.041	0.047	0.046	0.056
0.3	0.068	0.095	0.124	0.112	0.305	0.394	0.056	0.087	0.115	0.096	0.279	0.374
0.5	0.087	0.192	0.251	0.260	0.679	0.780	0.099	0.172	0.205	0.227	0.614	0.714
1	0.190	0.527	0.594	0.686	0.994	0.997	0.184	0.489	0.518	0.655	0.988	0.997
1.5	0.384	0.743	0.759	0.944	1.000	1.000	0.334	0.695	0.689	0.927	1.000	1.000
$\rho=0.5$												
0	0.050	0.051	0.045	0.054	0.052	0.041	0.050	0.048	0.046	0.062	0.044	0.053
0.3	0.074	0.115	0.126	0.128	0.311	0.394	0.073	0.115	0.137	0.125	0.284	0.349
0.5	0.108	0.212	0.258	0.236	0.644	0.740	0.109	0.193	0.230	0.243	0.600	0.702
1	0.224	0.487	0.549	0.671	0.985	0.994	0.218	0.458	0.491	0.630	0.974	0.986
1.5	0.400	0.677	0.698	0.928	1.000	1.000	0.365	0.633	0.646	0.919	0.998	0.999
$\rho=0.8$												
0	0.045	0.050	0.047	0.041	0.049	0.046	0.052	0.052	0.046	0.055	0.054	0.054
0.3	0.080	0.120	0.146	0.141	0.321	0.358	0.092	0.105	0.148	0.126	0.315	0.361
0.5	0.105	0.220	0.268	0.258	0.644	0.718	0.123	0.214	0.238	0.261	0.598	0.679
1	0.242	0.472	0.536	0.670	0.979	0.983	0.228	0.448	0.465	0.614	0.956	0.970
1.5	0.382	0.670	0.670	0.906	0.999	0.999	0.376	0.618	0.579	0.896	0.998	0.996

Note: Rejection frequency of the null hypothesis of no simultaneity ($H_0 : \beta_i = 0$) using standard testing procedures (F tests) *after* approximating the logistic function by a *first-order* Taylor expansion around $\gamma_i = 0$. The auxiliary regression estimated by 2SLS is Eq. 6. The results are presented only for Eq. 1 since they are symmetric. We run 2000 simulations.

Table 6
Size and power of the test for (β_1) with *unknown* γ and c – *third-order Taylor approximation.*

	$\xi=0, \phi_i=0, \theta=0$						$\xi=0.5, \phi_i=0.5, \theta=0.3$					
β_1	T=100			T=500			T=100			T=500		
γ	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5	0.5	1.5	2.5
$\rho=0$												
0	0.016	0.015	0.015	0.036	0.046	0.035	0.014	0.010	0.016	0.038	0.034	0.035
0.3	0.026	0.038	0.038	0.088	0.182	0.238	0.024	0.026	0.040	0.066	0.175	0.225
0.5	0.023	0.074	0.074	0.148	0.500	0.633	0.024	0.065	0.075	0.142	0.464	0.580
1	0.074	0.291	0.291	0.519	0.978	0.992	0.079	0.255	0.341	0.492	0.958	0.983
1.50	0.185	0.545	0.545	0.879	1.000	0.999	0.178	0.509	0.550	0.859	0.999	1.000
$\rho=0.5$												
0	0.025	0.021	0.019	0.036	0.035	0.042	0.023	0.020	0.020	0.038	0.040	0.044
0.3	0.035	0.055	0.070	0.091	0.201	0.273	0.034	0.062	0.072	0.094	0.194	0.270
0.5	0.058	0.117	0.131	0.176	0.504	0.587	0.050	0.094	0.134	0.167	0.459	0.563
1	0.136	0.328	0.365	0.535	0.953	0.975	0.116	0.305	0.324	0.492	0.943	0.966
1.5	0.244	0.501	0.555	0.843	0.996	0.997	0.227	0.536	0.550	0.813	0.993	0.994
$\rho=0.8$												
0	0.026	0.032	0.030	0.030	0.035	0.041	0.032	0.030	0.023	0.034	0.038	0.031
0.3	0.050	0.083	0.083	0.083	0.217	0.286	0.051	0.074	0.088	0.096	0.219	0.257
0.5	0.069	0.149	0.160	0.160	0.492	0.577	0.081	0.140	0.161	0.178	0.467	0.530
1	0.165	0.331	0.386	0.386	0.929	0.958	0.138	0.327	0.366	0.494	0.924	0.942
1.5	0.289	0.515	0.540	0.540	0.993	0.992	0.264	0.499	0.536	0.804	0.990	0.987

Note: Rejection frequency of the null hypothesis of no simultaneity ($H_0 : \beta_i = 0$) using standard testing procedures (F tests) *after* approximating the logistic function by a *third-order* Taylor expansion around $\gamma_i = 0$. The auxiliary regression estimated by NL2SLS is Eq. 7. The results are presented only for Eq. 1 since they are symmetric. We run 2000 simulations.

5.1. Modeling sovereign and bank returns

During the great financial and sovereign debt crises the within- and cross-country interactions between the sovereign and the banking sectors intensified as showed by several studies (e.g. Kallestrup et al. (2016)). The nature of this comovement was rather simultaneous (contagion) than lagged in time (i.e. spillovers). Acharya et al. (2014) design bank bailouts by the sovereign as the main mechanisms through which shocks were transmitted between these two markets. Government interventions to rescue distressed banks resulted in a dramatic increase of risk in the sovereign sector and considerable losses, liquidity and insolvency issues for banks which owned a large share of sovereign bonds. This mechanism continued like a loop. Other channels of market comovement which do not necessarily require common factors between markets, are wake-up calls (e.g. Ahnert and Bertsch (2017)), loss of confidence, general financial panic or herd behavior (e.g. Bae et al. (2003); Corsetti et al. (2005)).

Table 7
Summary statistics for each country.

	Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
Banking variables									
Annual returns (%)	−3.88	6.03	−10.83	−1.41	−7.14	0.14	−0.04	−0.45	3.52
Dividend yield (%)	2.26	4.00	1.62	3.80	2.30	4.71	2.50	3.99	2.60
VIX (%)	–	–	–	–	–	–	–	–	19.7
Sovereign variables									
Sovereign yield (%)	2.74	3.01	7.42	3.90	4.61	3.72	1.64	3.41	3.53
Fisc.bal/ GDP (%)	−0.8	−3.5	−6.3	−4.0	−3.8	−2.3	0.1	−0.5	−3.3
Debt / GDP (%)	67.8	78.6	137.0	114.4	91.0	67.7	48.4	60.8	82.1
S&P rating	AAA	AA+	BB+	A	A-	AA-	AAA	AAA	AAA
Correlation									
Banks and yield	0.20	0.09	−0.39	−0.19	−0.28	−0.02	0.09	0.09	−0.05
Banks and diff. yield	0.48	0.33	−0.21	−0.28	−0.37	−0.12	0.44	0.44	0.50

Data source: Datastream and Standard & Poor's. *Notes:* This table shows the average for each country of the dependent and of the explanatory variables. The data set contains nine developed countries with monthly data from January 1999 to November 2019. The rating categories of Standard & Poor's were translated into a numerical scale with the best rating AAA taking the value 22 and the worst rating D taking the value 1. In the last row "diff" stands for annual difference.

We explain the annual return of banks (BR_t) at the country level by the dividend yield (DY_t), the implied volatility of the S&P500 (VIX_t) and the yearly difference of the 10-year sovereign yield for the corresponding country, as shown in Eq. 8. Referring to Eq. 9, the annual change in the 10-year sovereign yield (SY_t) is explained by the debt to GDP ratio (DB_t), the fiscal balance to GDP ratio (FB_t), the sovereign rating[11] (RT_t) and the banking sector annual returns (BR_t).

$$\text{BR}_t = \delta_1 + \alpha_{1,1}\text{DY}_{t-k_1} + \alpha_{1,2}\text{VIX}_{t-k_1} + \beta_1 G(\text{SY}_t; \gamma_1, c_2) + u_{1,t} \tag{8}$$

$$\text{SY}_t = \delta_2 + \alpha_{2,1}\text{DB}_{t-k_2} + \alpha_{2,2}\text{FB}_{t-k_2} + \alpha_{2,3}\text{RT}_{t-k_2} + \beta_2 G(\text{BR}_t; \gamma_2, c_1) + u_{2,t} \tag{9}$$

While it is admitted that the banking sector and the sovereign are highly interrelated, it is not straightforward to guess the sign of the impact of one market on the other market. If the public sector is hit by a negative shock – say a major political event – bond prices should go down, thus leading to a higher yield to maturity. If the shock is big enough to put at risk the banking sector (through different channels), the banking price index should drop, thus leading to lower returns and a negative relationship between bank returns and sovereign yield. We could also observe the opposite effect, as banks could serve as a diversification opportunity, especially if their sovereign bond holdings are low.

Similarly, a shock originating from the banking sector could drag sovereign bond prices, especially in the countries where the banking sector represents a large share of the economic activity, or when governments have limited intervention tools to alleviate the shock. The resulting comovement has a negative sign. In many cases however, investors will use government bonds as a safe haven, leading to a positive relationship between bank returns and sovereign yield. In addition, while our model controls for important domestic variables, it disregards international factors able to modify the country relationship between banks and the sovereign.

While we do not claim that the explanatory variables are strictly exogenous, we do believe that the weak exogeneity assumption is satisfied. Moreover, we integrate them within the system with lags k_1 and k_2. We have tried several lags and have concluded that for banks the explanatory power is reduced as the lag increases, showing that the transmission lag is very short. On the other hand, the fiscal variables and the sovereign rating keep their impact on the change of 10-year sovereign yield even for further lags, indicating that more time is needed for the structural reforms to be incorporated in sovereign bond prices. As a result, we have retained $k_1 = 1$ and $k_2 = 6$ months.

5.2. Data description

Our dataset extends over more than 20 years, from January 1999 to November 2019. The period under study covers the dot-com bubble, the financial and the Euro Area sovereign debt crises, characterized by high volatility. Indeed, the VIX registered historical peaks in 2008, 2010 and 2012. More recently, the volatility has also substantially increased since the start of the trade dispute between the United States and China in 2018.

Table 7 collects summary statistics by country on the dependent and on the explanatory variables. During the period considered, the annual returns of the banking sector have been significantly positive only for France and for the United States. The worst performance during the last 20 years was registered in Greece. The latter has also registered the largest cost of public funding and the worst average rating from the S&P. In early 2012, the sovereign yield reached almost 40% which constrained Greece to request a bailout from the Troika. The most healthy public finances can be found in Switzerland, where

[11] We consider the sovereign rating from Standard & Poor's that we convert into a numerical scale from 1 (D rating) to 22 (AAA rating). Given that during the period under study, there is no rating change for Germany, Switzerland, United Kingdom and only one change for the United States, this variable is excluded when investigating these countries.

the government can get public financing with a negative yield beyond 30 years of maturity. While some countries remain riskier than others, the cost of public funding has been dramatically reduced with the highly accommodative monetary policies, especially in the Euro Area and in Switzerland where negative interest rates prevail.

The last part of the table shows the correlation between annual returns of banks and the 10-year sovereign yield. As explained in Section 5.1 the correlation could theoretically be positive or negative. In our dataset, the correlation appears to be negative for half of the countries and is quite small in magnitude for all the countries. This correlation strengthens if we consider the yearly changes of the sovereign yield instead of its level. Changes are most subject to market forces, rather than the trend, which follows monetary policy decisions. Fig. A.1 in the appendix shows for each country the evolution over time of annual bank returns and the yearly difference of the 10-year sovereign yield. Bank returns have a high volatility and a large magnitude. The variation is much smaller for the sovereign sector, with the exception of Greece.

5.3. Estimation results

Table 8 shows the estimation results for the ST-SEM model, both with fixed and unknown γ_i and c_j, and a linear specification estimated by OLS which does not correct for potential endogeneity. Although a linear model is not directly comparable to our nonlinear specification and ignores the endogeneity issue, it provides an indication of the results obtained from commonly used models. In the top panel, the OLS estimation results suggest that the sovereign and banking sectors are strongly linked. For the most fragile countries in the Euro Area the sign of the coefficients is negative, in line with the above hypotheses.

We use the test developed in Section 4 to test for simultaneity and present the results in Table A.1 in the appendix. Using first-order Taylor series expansion, the null hypothesis of no dependence is strongly rejected for most of the countries. Similar results are found with the third-order Taylor expansion, indicating that not only are the sovereign and the banking sectors of major developed economies correlated, but this dependence is nonlinear in nature. Before estimating the ST-SEM we check the relevence of the instrumental variables and present the results in Table A.3 in the appendix. The null hypothesis is stongly rejected with a single exception: instruments for $y_{2,t}$ in the United States. We estimate the ST-SEM based on this evidence.

While the impact of a linear specification is constant over time, in ST-SEM the impact of one market on the other is the product of the parameter β_i and the logistic function. Throughout the three different ST-SEM specifications – (i) γ_i and c_j fixed, (ii) only γ_i fixed, (iii) neither γ_i nor c_j is fixed – the estimated coefficients appear to be consistent in sign and magnitude. In general, the results show that countries with high governance scores, such Germany, Switzerland, the United States and the United Kingdom register a positive comovement between the variables under investigation. On the opposite, the comovement turns to negative or insignificant for countries with weaker fundamentals such as Italy and Spain. For Greece and Portugal the results are mixed.

We first fix γ_i at 3 and c_j at the thirtieth percentile of the distribution of the transition variable $y_{j,t}$. In a second step, we estimate the scale parameter c_j with the other parameters of the model. The estimated value of c_j is mostly in the extremes, but not always at the lower end of the distribution of $y_{j,t}$. Finally, we add γ_i to the estimation procedure. We find that the estimated value of γ_i is in general close to 20 and suggests an abrupt transmission of shocks. As highlighted in the methodological framework, a precise estimation of c_j, and in particular of γ_i, is difficult. In our empirical framework, the estimation of the latter alters in some occasions the estimation of β_i.

Fig. A.2 in the appendix shows the evolution of the logistic function for both equations in Italy and in the United Kingdom. Since the financial crisis, both countries have gone through important economic, financial and political events. In Italy, the sovereign debt crisis that followed the financial crisis reflected the large public debt burden. More recently, the bank Monte dei Paschi di Sienna was bailed out in 2017 and was followed by political turmoil. In the United Kingdom, the Brexit process has been a source of instability and insecurity for the government and a major risk for the banking sector. The value of the logistic function is close to 1 around the above-mentioned negative shocks. A striking feature is that banking sector variations are transmitted more abruptly to the sovereign sector. This in line with the intrinsic volatility of the banking sector and with the role of the government as a guarantee for stability.

The goodness of fit of the first equation of the system is quite high, with the R-squared ranging between 30% and 55%. The dividend yield and the VIX have a negative impact on bank returns (estimation results presented in Table A.5 in the appendix). This result is consistent through countries and is highly significant. For the second equation, the R-squared is lower and roughly ranges between 15% and 35%. Higher debt to GDP ratio, deterioration of the fiscal balance, and lower rating should be all associated with larger sovereign yields. While the impact of these variables is significant for most countries, the sign of the coefficients does not always satisfy the above hypotheses.

We also examine the model with consumer price inflation (CPI) as a common exogenous variable to both equations. Higher inflation triggers in general a price correction for stocks and bonds leading to lower returns and higher yields. The simultaneity test (Table A.2) and the weak identification test (Table A.4) are in line with the above conclusions. The main estimation results in Table A.6 are comparable with the results in Table 8. The estimation results for the control variables, including the CPI inflation with one-month lag, are in Table A.7. For bank returns the impact of CPI is negative, but significant only for a few countries. For sovereign yields the impact is positive and strongly significant for most of the countries. The R-squared improves significantly.

Table 8
Estimation results for the banking and for the sovereign sectors.

	Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
				Linear-model estimation results (OLS)					
β_1	18.77	14.17	−0.54	−5.87	−5.68	−3.81	14.04	10.39	8.44
σ_1	(2.83)	(2.30)	(0.43)	(1.61)	(1.16)	(1.59)	(1.98)	(1.77)	(1.47)
R_1^2	0.40	0.48	0.30	0.48	0.34	0.40	0.56	0.49	0.54
β_2	0.009	0.007	−0.057	−0.007	−0.029	−0.002	0.010	0.013	0.017
σ_2	(0.001)	(0.001)	(0.011)	(0.002)	(0.004)	(0.002)	(0.002)	(0.001)	(0.002)
R_2^2	0.26	0.19	0.22	0.21	0.37	0.31	0.19	0.26	0.26
				Nonlinear ST-SEM model – *known γ and c*					
β_1	54.25	20.67	49.44	−41.52	40.17	−21.37	43.64	83.13	64.79
σ_1	(17.79)	(9.53)	(13.08)	(13.05)	(15.41)	(7.36)	(13.95)	(14.90)	(15.85)
γ_1	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_2	−0.60	−0.56	−1.18	−0.53	−0.82	−0.50	−0.53	−0.48	−0.55
$Q(c_2)$	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
R_1^2	0.32	0.41	0.37	0.47	0.29	0.41	0.48	0.56	0.53
β_2	0.93	0.32	-8.74	0.19	−1.48	0.03	1.19	1.15	1.11
σ_2	(0.16)	(0.16)	(1.84)	(0.22)	(0.60)	(0.17)	(0.24)	(0.19)	(0.19)
γ_2	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_1	−24.66	−8.98	−52.19	−21.91	−32.06	−18.59	−13.56	−10.61	−6.14
$Q(c_1)$	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
R_2^2	0.17	0.08	0.21	0.15	0.20	0.30	0.18	0.19	0.13
				Nonlinear ST-SEM model – *known γ and unknown c*					
β_1	89.84	28.42	53.16	−79.57	53.32	−22.77	41.20	87.89	75.01
σ_1	(37.37)	20.37	(15.87)	(18.64)	(14.33)	(10.67)	(11.15)	(29.19)	(32.75)
γ_1	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_2	0.27	0.29	−1.65	0.65	−1.48	−0.77	0.32	−0.69	0.18
$Q(c_2)$	0.79	0.80	0.22	0.80	0.17	0.21	0.80	0.21	0.68
R_1^2	0.36	0.42	0.37	0.52	0.31	0.41	0.50	0.56	0.55
β_2	0.85	0.34	−12.09	0.42	−2.71	−0.42	1.56	1.19	1.18
σ_2	(0.18)	(0.17)	(1.65)	(0.31)	(0.51)	(0.47)	(0.33)	(0.18)	(0.20)
γ_2	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_1	−17.65	−13.35	−63.08	16.05	−5.78	17.67	13.79	−16.89	−2.08
$Q(c_1)$	0.36	0.27	0.15	0.66	0.50	0.73	0.71	0.21	0.37
R_2^2	0.18	0.09	0.30	0.16	0.28	0.31	0.26	0.21	0.14
				Nonlinear ST-SEM model – *unknown γ and c*					
β_1	66.73	29.06	52.94	−83.72	59.38	−21.94	31.46	105.64	47.63
σ_1	(17.78)	(8.39)	(14.99)	(63.24)	(22.66)	(7.62)	(11.08)	(198.38)	(20.43)
γ_1	19.74	20.65	18.59	2.60	17.21	18.09	20.40	2.00	12.84
c_2	0.21	0.21	−1.48	0.65	−1.42	−0.08	0.32	−0.65	0.00
$Q(c_2)$	0.76	0.76	0.24	0.80	0.18	0.54	0.80	0.23	0.60
R_1^2	0.39	0.44	0.38	0.52	0.34	0.42	0.50	0.56	0.56
β_2	0.89	0.36	−12.08	0.43	−2.71	−0.46	1.61	1.20	1.20
σ_2	(0.16)	(0.17)	(1.92)	(0.32)	(0.59)	(0.32)	(0.32)	(0.20)	(0.37)
γ_2	20.45	19.05	2.05	18.28	19.71	19.40	20.94	17.42	18.11
c_1	−17.55	−13.32	−63.08	15.91	−5.80	17.42	13.55	−18.51	−1.98
$Q(c_1)$	0.36	0.27	0.15	0.66	0.50	0.73	0.71	0.20	0.37
R_2^2	0.18	0.09	0.30	0.16	0.28	0.31	0.26	0.21	0.15

Notes: The table shows the estimation results for Eqs. 8 and 9. We present the estimated value of β_i and in parentheses its standard deviation σ_i. γ_i and c_j are first fixed and then estimated with the other parameters of the system. $Q(c_j)$ corresponds to the quantile of the distribution of y_j at which c_j is fixed or estimated. R_i^2 is the R-squared for each equation. In the top panel we show the estimation results of linear model estimated by OLS.

6. Conclusion

In this paper we develop a new econometric framework within a smooth transition simultaneous equation model to measure the comovement between two endogenous random variables. The model is a generalization of Pesaran and Pick's (2007) threshold SEM, which is included as a special case. Comovement occurs through a smooth function which is dependent on the magnitude of the endogenous variable. The framework also allows for an indirect channel of dependence controlling for constant links between the endogenous variables, which is modelled through common observed and unobserved variables. The presence of equation-specific exogenous variables is necessary to disentangle between direct and indirect market comovements.

We develop and adapt a valid estimation method and a testing procedure for simultaneity that we examine through Monte Carlo simulations. An advantage of our specification compared to Pesaran and Pick's model is that conditions can be set on the parameters in order to obtain a unique solution of the system. Models with multiple reduced forms – explicit or implicit – are known in the literature as incoherent. They are unsuitable for forecasting and the properties of their estimators are unknown. We suggest numerical procedures to solve the system of equations.

The Monte Carlo simulations show that the parameters of the model are correctly estimated when the smoothness and the threshold parameters of the dependence variable are known. When all the parameters are unknown, the results deteriorate for small samples, but highly improve for larger samples. Although we obtain good results for some parameter combinations, an accurate estimation of the threshold parameter, and especially of the smoothness parameter proves difficult. This is a well known issue in the literature – perhaps not sufficiently explored – caused by a flat objective function. The tests for simultaneity which control for the presence of nuisance parameters are easy to implement and compute, and have good size and power properties.

As an illustration, we study the comovement between the sovereign and the banking sectors for the main Euro Area countries, Switzerland, the United Kingdom and the United States. The period under study includes major events such as the financial crisis, the Euro Area debt crisis and the Brexit vote. Our results show that market interdependence has significantly strengthened during distressed times. This is especially the case for countries with weak fundamentals and / or with a fragile banking sector.

As a future line of research, several extensions of this methodology are envisaged. First, the study of the properties of a multi-equation system is required for the investigation of comovement in a group of countries or markets. Second, a skewness parameter should be introduced in the transition function between different regimes in order to take into account a potentially asymmetric transmission of shocks. To further increase the flexibility, the threshold parameter could be modeled as time-varying. The methodology can be used to study the dependence structure in various markets such as stocks, bonds, interest rates, exchange rates.

Data availability

Data will be made available on request.

Appendix A

A1. Limiting distribution of the estimated parameters

We consider the ith equation of a two-equation ST-SEM which is written as follows:

$$y_{it} = f(\boldsymbol{h}_{it}, \boldsymbol{\zeta}_i) + u_{it}, \quad f(\boldsymbol{h}_{it}, \boldsymbol{\zeta}_i) = \boldsymbol{\delta}_i' \boldsymbol{z}_t + \boldsymbol{\alpha}_i' \boldsymbol{x}_{it} + \beta_i G(y_{jt}; \gamma_i, c_j)$$

Table A.1
Simultaneity test results for the principal model.

		Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
		First-order Taylor series approximation								
F-test	Eq. 1	1.4	2.1	4.6	0.4	11.5	10.7	4.0	16.2	4.6
	Eq. 2	30.7	17.6	5.5	1.7	13.1	4.9	7.8	41.6	26.5
p-value	Eq. 1	0.244	0.151	0.033	0.548	0.001	0.001	0.045	0.000	0.033
	Eq. 2	0.000	0.000	0.020	0.198	0.000	0.028	0.006	0.000	0.000
		Third-order Taylor series approximation								
F-test	Eq. 1	1.0	5.0	3.2	6.1	10.0	6.3	4.6	14.8	7.1
	Eq. 2	12.4	1.0	1.9	1.2	7.1	4.4	18.1	10.7	11.2
p-value	Eq. 1	0.375	0.002	0.025	0.001	0.000	0.000	0.004	0.000	0.000
	Eq. 2	0.000	0.391	0.136	0.293	0.000	0.005	0.000	0.000	0.000

Notes: Simultaneity test results – F-test and p-value – of first- and third-order Taylor series approximation for the empirical application presented in Eqs. 8 and 9.

Table A.2
Simultaneity test results for the model augmented with inflation.

		Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
		First-order Taylor series approximation								
F-test	Eq. 1	0.5	10.4	12.2	6.9	18.1	8.6	0.8	18.2	4.0
	Eq. 2	28.0	23.7	3.7	1.5	8.8	4.5	11.5	35.8	30.0
p-value	Eq. 1	0.490	0.001	0.001	0.009	0.000	0.004	0.384	0.000	0.046
	Eq. 2	0.000	0.000	0.057	0.222	0.003	0.034	0.001	0.000	0.000
		Third-order Taylor series approximation								
F-test	Eq. 1	5.9	9.2	7.0	7.0	8.3	3.9	6.6	15.7	7.0
	Eq. 2	11.2	2.5	4.1	2.1	7.6	4.1	15.9	14.9	11.4
p-value	Eq. 1	0.001	0.000	0.000	0.000	0.000	0.009	0.000	0.000	0.000
	Eq. 2	0.000	0.057	0.008	0.102	0.000	0.007	0.000	0.000	0.000

Notes: Simultaneity test results – F-test and p-value – of first- and third-order Taylor series approximation for the empirical application augmented with consumer price inflation. See notes under Table A.6.

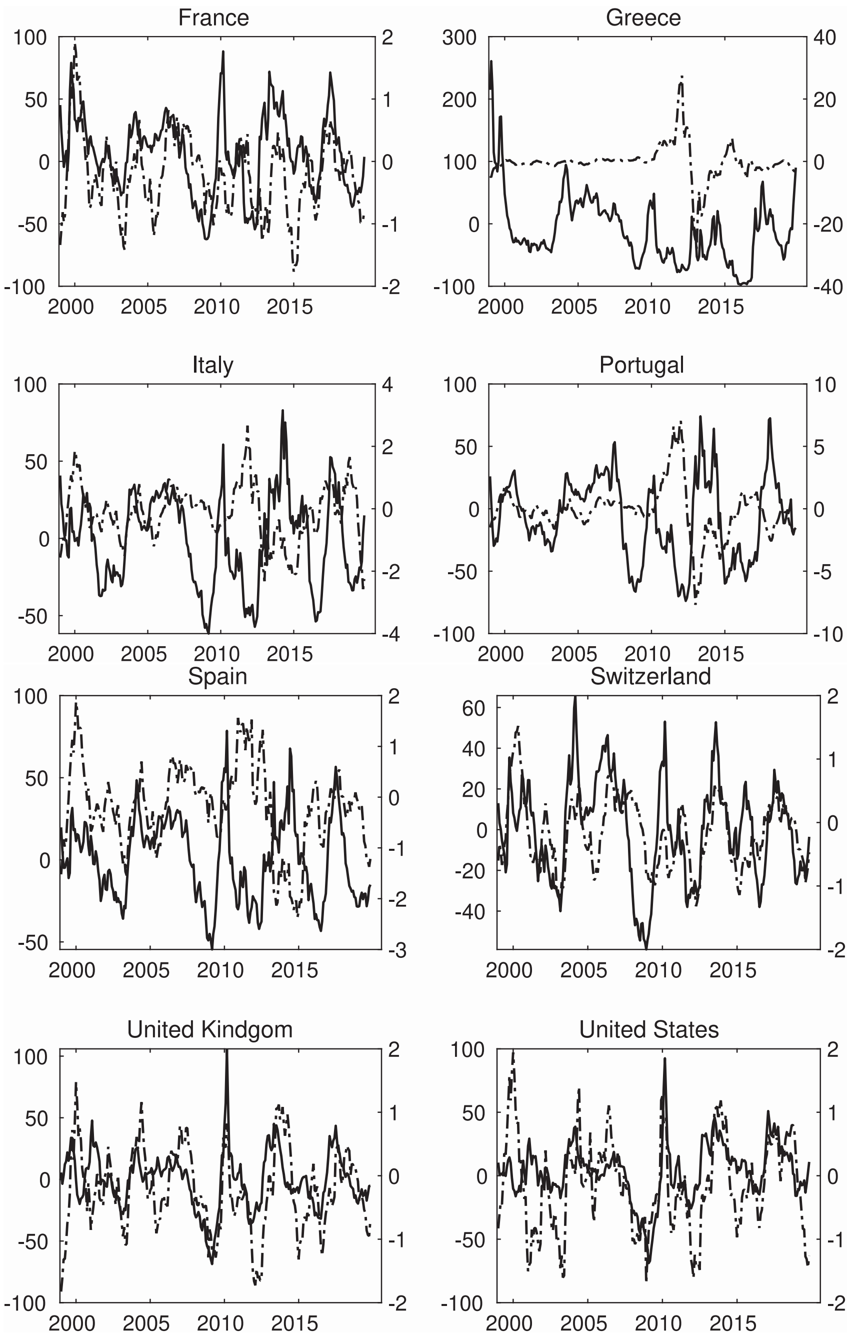

Fig. A.1. Banks' performance and sovereign yield for each country. *Notes*: The solid line shows the annual return (in percentage) of the banking sector for each country (left scale). The dotted line shows the annual difference (in percentage) of the sovereign yield (right scale).

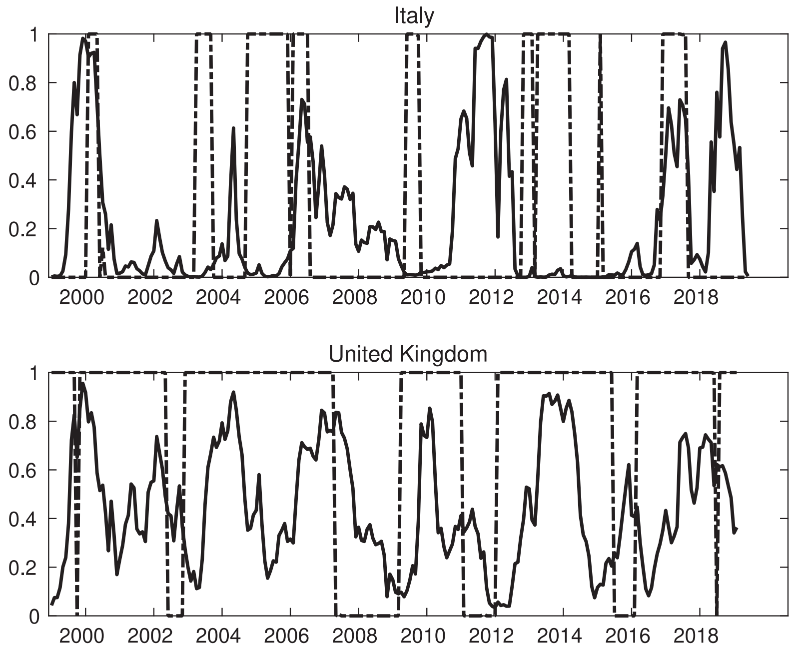

Fig. A.2. Transition functions. *Notes*: The graphic presents the logistic function for Eq. 8 (solid line) and Eq. 9 (dashed line). Two countries are selected: Italy (top pannel) and the United Kingdom (bottom pannel).

Table A.3
Weak identification test for the principal model.

		Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
Eq. 1	χ^2-test	11.4	29.7	18.5	45.5	26.8	128.6	14.5	25.6	1.8
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41
	F-test	5.7	9.9	6.2	15.2	8.9	42.9	7.2	12.8	0.9
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42
Eq. 2	χ^2-test	124.3	137.8	113.6	204.4	132.4	286.1	145.3	227.1	232.3
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	F-test	62.1	68.9	56.8	102.2	66.2	143.1	72.6	113.5	116.2
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes: Top panel: $y_{2,t} = \boldsymbol{\lambda}_2 \boldsymbol{z}_t + \boldsymbol{\phi}_2 \boldsymbol{x}_{1,t} + \boldsymbol{\psi}_2 \boldsymbol{x}_{2,t} + e_{2,t}$. Bottom panel: $y_{1,t} = \boldsymbol{\lambda}_1 \boldsymbol{z}_t + \boldsymbol{\phi}_1 \boldsymbol{x}_{2,t} + \boldsymbol{\psi}_1 \boldsymbol{x}_{1,t} + e_{1,t}$. Test $H_0 : \boldsymbol{\psi}_i = 0$ using the estimates from the first stage of the 2SLS procedure. To compute the p-values, the value of the test is compared to the critical values of χ^2 and F distributions at 95%. See notes under Table 8.

Table A.4
Weak identification test for the model augmented with inflation.

		Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
Eq. 1	χ^2-test	17.2	34.4	7.1	36.5	32.3	118.3	11.8	19.9	8.7
	p-value	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.01
	F-test	8.6	11.5	2.4	12.2	10.8	39.4	5.9	10.0	4.3
	p-value	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.01
Eq. 2	χ^2-test	121.8	160.0	116.3	198.1	128.6	295.2	145.9	228.0	233.3
	p-value	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	F-test	60.9	80.0	58.2	99.1	64.3	147.6	73.0	114.0	116.6
	p-value	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Notes: Top panel: $y_{2,t} = \boldsymbol{\lambda}_2 \boldsymbol{z}_t + \boldsymbol{\phi}_2 \boldsymbol{x}_{1,t} + \boldsymbol{\psi}_2 \boldsymbol{x}_{2,t} + e_{2,t}$. Bottom panel: $y_{1,t} = \boldsymbol{\lambda}_1 \boldsymbol{z}_t + \boldsymbol{\phi}_1 \boldsymbol{x}_{2,t} + \boldsymbol{\psi}_1 \boldsymbol{x}_{1,t} + e_{1,t}$. Test $H_0 : \boldsymbol{\psi}_i = 0$ using the estimates from the first stage of the 2SLS procedure. To compute the p-values, the value of the test is compared to the critical values of χ^2 and F distributions at 95%. See notes under Table A.6.

$$G(y_{jt};\gamma_i,c_j)=\left\{1+e^{-\gamma_i(y_{jt}-c_j)}\right\}^{-1}\quad \gamma_i>0\quad i\neq j\quad i=1,2\quad t=1,\ldots,T$$

with $\boldsymbol{\zeta}_i=[\boldsymbol{\delta}_i',\boldsymbol{\alpha}_i',\beta_i,\gamma_i,c_j]'$ the overall set of parameters, and $\mathbf{h}_{it}=[\mathbf{z}_t',\ \mathbf{x}_{it}',\ G(y_{jt};\gamma_i,c_j)]'$ the explanatory variables. The NL2SLS estimator of $\boldsymbol{\zeta}_i$ denoted by $\hat{\boldsymbol{\zeta}}_i$ is the value of $\boldsymbol{\zeta}_i$ that minimizes the following equation:

$$\hat{\boldsymbol{\zeta}}_i=\text{argmin}\,(\boldsymbol{y}_i-f_i)'(Q(Q'Q)^{-1}Q')(\boldsymbol{y}_i-f_i)$$

with $\boldsymbol{y}_i=[y_{i1}\,y_{i2}\,\ldots,y_{iT}]'$, $f_i=[f(\boldsymbol{h}_{i1},\boldsymbol{\zeta}_i),\ldots,f(\boldsymbol{h}_{iT},\boldsymbol{\zeta}_i)]'$ and Q a $T\times K$ matrix containing low-order polynomials of the overall set of the exogenous parameters of the model.

Following Kelejian (1971) and Amemiya (1974), under usual regularity conditions, the limit in probability of $\hat{\boldsymbol{\zeta}}_i$ converges to its true value and the limiting distribution is normal.

$$\sqrt{T}(\hat{\boldsymbol{\zeta}}_i-\boldsymbol{\zeta}_i)\sim\mathcal{N}\left\{0,\ \sigma_i^2\text{plim}\left[\frac{1}{T}\frac{\partial f_i'}{\partial\boldsymbol{\zeta}_i}Q(Q'Q)^{-1}Q'\frac{\partial f_i)}{\partial\boldsymbol{\zeta}_i}\right]^{-1}\right\}$$

with

$$\frac{\partial f(\boldsymbol{h}_{it},\boldsymbol{\zeta}_i)}{\partial\boldsymbol{\zeta}_i}=\begin{pmatrix}\frac{\partial f(\boldsymbol{h}_{it},\boldsymbol{\zeta}_i)}{\partial\boldsymbol{\delta}_i}\\ \frac{\partial f(\boldsymbol{h}_{it},\boldsymbol{\zeta}_i)}{\partial\boldsymbol{\alpha}_i}\\ \frac{\partial f(\boldsymbol{h}_{it},\boldsymbol{\zeta}_i)}{\partial\boldsymbol{\beta}_i}\\ \frac{\partial f(\boldsymbol{h}_{it},\boldsymbol{\zeta}_i)}{\partial\gamma_i}\\ \frac{\partial f(\boldsymbol{h}_{it},\boldsymbol{\zeta}_i)}{\partial c_j}\end{pmatrix}=\begin{pmatrix}\boldsymbol{z}_t\\ \boldsymbol{x}_{it}\\ G(y_{jt};\gamma_i,c_j)\\ \beta_i G(y_{jt};\gamma_i,c_j)\,[1-G(y_{jt};\gamma_i,c_j)]\,(y_{jt}-c_j)\\ -\beta_i G(y_{jt};\gamma_i,c_j)\,[1-G(y_{jt};\gamma_i,c_j)]\,\gamma_i\end{pmatrix}$$

A2. Derivation of the tests for contagion

Below we give the details of the derivation of the tests for contagion based on the procedure of Luukkonnen et al. (1988) and Wooldridge (2002). Consider the following two-equation system:

$$y_{1t}=\boldsymbol{\delta}_1'\boldsymbol{z}_t+\boldsymbol{\alpha}_1'\boldsymbol{x}_{1t}+\beta_1G(y_{2t};\gamma_1,c_2)+u_{1t}\tag{A.1}$$

$$y_{2t}=\boldsymbol{\delta}_2'\boldsymbol{z}_t+\boldsymbol{\alpha}_2'\boldsymbol{x}_{2t}+\beta_2G(y_{1t};\gamma_2,c_1)+u_{2t}\tag{A.2}$$

$$G(y_{jt};\gamma_i,c_j)=\left\{1+e^{-\gamma_i(y_{jt}-c_j\sigma_{j,t-1})}\right\}^{-1}\quad\gamma_i>0,\quad i,j=1,2,\quad i\neq j\tag{A.3}$$

A2.1. First-order Taylor expansion

We compute the derivative with respect to γ_i of the logistic function in Eq. A.3 and evaluate it at $\gamma_i=0$.

$$\frac{\partial G(y_{jt};\gamma_i,c_j)}{\partial\gamma_i}=\frac{e^{\gamma_i(y_{jt}-c_j)}}{\left(1+e^{\gamma_i(y_{jt}-c_j)}\right)^2}(y_{jt}-c_j)=G(y_{jt};\gamma_i,c_j)\big[1-G(y_{jt};\gamma_i,c_j)\big](y_{jt}-c_j)$$

$$\left.\frac{\partial G(y_{jt};\gamma_i,c_j)}{\partial\gamma_i}\right|_{\gamma_i=0}=\frac{1}{4}(y_{jt}-c_j)$$

We evaluate the logistic function at $\gamma_i=0$ and then subtract the value obtained from the logistic function $G(y_{jt};\gamma_i,c_j)$. We denote this new function by $\tilde{G}(y_{jt};\gamma_i,c_j)$, whose value in $\gamma_i=0$ is equal to 0. This reparametrization is one of the conditions set by Luukkonnen et al. (1988) and is used only for the derivation of the test.

$$G(y_{jt};\gamma_i,c_j)|_{\gamma_i=0}=\frac{1}{2}$$

$$\tilde{G}(y_{jt};\gamma_i,c_j)=G(y_{jt};\gamma_i,c_j)-G(y_{jt};\gamma_i,c_j)|_{\gamma_i=0}$$

The first-order Taylor expansion of $\tilde{G}(y_{jt};\gamma_i,c_j)$ is written as follows, with R_{1i} the remainder:

$$\tilde{G}(y_{jt};\gamma_i,c_j)=-\frac{1}{4}c_j\gamma_i+\frac{1}{4}\gamma_iy_{jt}+R_{1i}$$

Replacing the above approximation in the i^{th} equation of the ST-SEM yields the following auxiliary regression:

$$y_{it}=\tilde{\delta}_i'z_t+\alpha_i'x_{it}+\tilde{\beta}_iy_{jt}+u_{it}^*$$

with $\tilde{\delta}_i=\left[(\delta_{i,1}-\frac{1}{4}\beta_i\gamma_ic_j)\ \delta_i^{-\prime}\right]'$, with δ_i^- equals to δ_i minus the constant term δ_{i1}, $\tilde{\beta}_i=\frac{1}{4}\beta_i\gamma_i$ and $u_{it}^*=\beta_iR_{1i}+u_{it}$.

Table A.5
Estimated coefficients of the control variables for the principal model.

	Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
Linear-model estimation results (OLS)									
$\alpha_{1,1}$	−4.05	−8.13	12.13	−6.47	1.41	−4.29	−9.71	−7.0	−6.90
$\sigma_{1,1}$	(1.04)	(0.94)	(1.95)	(0.80)	(1.24)	(0.63)	(0.98)	(0.80)	(0.92)
$\alpha_{1,2}$	−0.95	−1.08	−3.28	−1.08	−2.06	−1.36	−1.58	−0.54	−1.12
$\sigma_{1,2}$	(0.26)	(0.19)	(0.41)	(0.19)	(0.33)	(0.17)	(0.14)	(0.16)	(0.13)
δ_1	28.63	62.30	25.22	43.68	25.85	46.35	57.50	40.21	44.67
σ_{δ_1}	(4.25)	(4.28)	(7.09)	(3.85)	(5.25)	(4.12)	(3.97)	(3.39)	(3.30)
$\alpha_{2,1}$	−0.02	0.01	0.08	0.04	0.02	0.03	0.002	−0.002	−0.0004
$\sigma_{2,1}$	(0.004)	(0.01)	(0.04)	(0.01)	(0.02)	(0.01)	(0.01)	(0.002)	(0.003)
$\alpha_{2,2}$	0.07	0.11	−0.48	0.21	−0.26	−0.05	0.0013	0.04	0.018
$\sigma_{2,2}$	(0.04)	(0.05)	(0.10)	(0.04)	(0.07)	(0.02)	(0.06)	(0.02)	(0.02)
$\alpha_{3,2}$	–	0.26	0.72	0.23	0.37	0.33	–	–	–
$\sigma_{3,2}$	–	(0.14)	(0.22)	(0.06)	(0.17)	(0.05)	–	–	–
δ_2	0.94	−5.92	−23.88	−7.69	−9.31	−8.42	−0.26	−0.06	−0.11
σ_{δ_2}	(0.29)	(3.44)	(7.76)	(2.31)	(4.66)	(1.47)	(0.38)	(0.12)	(0.18)
Nonlinear ST-SEM model – *known γ and c*									
$\alpha_{1,1}$	−3.87	−8.43	9.17	−5.19	−5.29	−4.55	−10.22	−6.26	−6.07
$\sigma_{1,1}$	(1.12)	(0.98)	(2.37)	(1.03)	(1.73)	(0.67)	(1.05)	(1.15)	(1.22)
$\alpha_{1,2}$	−0.80	−1.07	−3.94	−1.16	−1.51	−1.30	−1.36	0.19	−0.25
$\sigma_{1,2}$	(0.31)	(0.21)	(0.47)	(0.22)	(0.40)	(0.18)	(0.21)	(0.28)	(0.33)
δ_1	−15.72	46.82	6.00	68.75	6.14	60.91	22.76	−31.38	−19.32
σ_{δ_1}	(15.85)	(8.15)	(9.66)	(8.22)	(10.85)	(6.28)	(12.72)	(14.36)	(17.64)
$\alpha_{2,1}$	−0.03	0.01	0.11	0.04	0.03	0.03	0.03	0.001	0.001
$\sigma_{2,1}$	(0.01)	(0.01)	(0.04)	(0.01)	(0.02)	(0.01)	(0.01)	(0.002)	(0.003)
$\alpha_{2,2}$	0.10	0.11	−0.50	0.23	−0.26	−0.05	0.23	0.05	0.002
$\sigma_{2,2}$	(0.05)	(0.07)	(0.10)	(0.05)	(0.08)	(0.02)	(0.10)	(0.02)	(0.018)
$\alpha_{3,2}$	–	0.24	0.93	0.24	0.39	0.35	–	–	–
$\sigma_{3,2}$	–	(0.16)	(0.23)	(0.06)	(0.18)	(0.06)	–	–	–
δ_2	0.94	−5.88	−23.41	−8.21	−8.78	−9.16	−2.33	−1.00	−0.96
σ_{δ_2}	(0.33)	(4.31)	(7.79)	(2.49)	(5.13)	(1.72)	(0.69)	(0.21)	(0.22)
Nonlinear ST-SEM model – *known γ and unknown c*									
$\alpha_{1,1}$	−4.28	−7.88	8.93	−5.61	−5.69	−4.66	−9.24	−6.33	−5.95
$\sigma_{1,1}$	(1.12)	(1.16)	(2.41)	(1.12)	(1.59)	(0.93)	(1.06)	(1.17)	(1.52)
$\alpha_{1,2}$	−0.56	−1.07	−3.86	−1.26	−1.51	−1.26	−1.53	0.16	−0.30
$\sigma_{1,2}$	(0.48)	(0.24)	(0.50)	(0.27)	(0.38)	(0.29)	(0.21)	(0.29)	(0.35)
δ_1	−7.06	50.61	−0.09	62.97	−8.36	63.68	41.88	−41.24	−2.49
σ_{δ_1}	(25.92)	(8.31)	(21.43)	(7.04)	(15.27)	(19.46)	(8.89)	(37.20)	(14.50)
$\alpha_{2,1}$	−0.03	0.01	0.17	0.05	0.02	0.03	−0.01	−0.003	−0.0002
$\sigma_{2,1}$	(0.01)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)	(0.02)	(0.005)	(0.0037)
$\alpha_{2,2}$	0.12	0.12	−0.65	0.22	−0.25	−0.05	0.05	0.01	−0.0044
$\sigma_{2,2}$	(0.06)	(0.07)	(0.11)	(0.06)	(0.09)	(0.02)	(0.09)	(0.04)	(0.02)
$\alpha_{3,2}$	–	0.27	1.28	0.27	0.38	0.30	–	–	–
$\sigma_{3,2}$	–	(0.17)	(0.29)	(0.06)	(0.19)	(0.10)	–	–	–
δ_2	1.12	−6.52	−33.33	−9.28	−7.36	−7.61	−0.32	−0.93	−0.88
σ_{δ_2}	(0.52)	(4.39)	(9.58)	(2.57)	(5.20)	(2.86)	(0.80)	(0.28)	(0.35)
Nonlinear ST-SEM model – *unknown γ and c*									
$\alpha_{1,1}$	−5.21	−7.74	8.57	−5.49	−6.27	−4.35	−8.38	−6.34	−7.80
$\sigma_{1,1}$	(1.31)	(0.99)	(2.61)	(1.31)	(1.51)	(0.70)	(1.65)	(1.66)	(1.65)
$\alpha_{1,2}$	−0.36	−0.96	−3.84	−1.25	−1.48	−1.41	−1.60	0.17	−0.35
$\sigma_{1,2}$	(0.35)	(0.22)	(0.48)	(0.55)	(0.36)	(0.19)	(0.17)	(0.39)	(0.40)
δ_1	−0.17	48.73	0.13	63.96	−13.57	58.51	45.96	−48.74	11.21
σ_{δ_1}	(9.56)	(7.04)	(15.16)	(8.42)	(18.30)	(5.67)	(6.03)	(199.00)	(19.24)
$\alpha_{2,1}$	0.01	0.01	0.16	0.05	0.02	0.02	−0.01	−0.004	−0.0004
$\sigma_{2,1}$	(0.01)	(0.01)	(0.05)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
$\alpha_{2,2}$	0.13	0.13	−0.64	0.22	−0.25	−0.05	0.06	0.01	−0.01
$\sigma_{2,2}$	(0.07)	(0.07)	(0.13)	(0.06)	(0.12)	(0.02)	(0.08)	(0.04)	(0.02)
$\alpha_{3,2}$	–	0.28	1.27	0.27	0.38	0.29	–	–	–
$\sigma_{3,2}$	–	(0.18)	(0.30)	(0.07)	(0.25)	(0.07)	–	–	–
δ_2	1.12	−6.88	−32.87	−9.32	−7.35	−7.39	−0.37	−0.89	−0.87
σ_{δ_2}	(0.38)	(4.62)	(9.98)	(2.66)	(6.87)	(1.84)	(0.67)	(0.39)	(0.41)

Notes: Complement of the estimation results from Eqs. 8 and 9 presented in Table 8.

Table A.6
Main estimation results of augmented model with consumer price inflation.

	Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
					Linear-model estimation results (OLS)				
β_1	19.00	18.12	−1.42	−5.25	−5.93	−2.73	15.84	10.39	8.31
σ_1	(2.86)	(2.25)	(0.41)	(1.74)	(1.17)	(1.76)	(2.22)	(1.77)	(1.46)
R_1^2	0.41	0.55	0.43	0.48	0.34	0.41	0.56	0.49	0.55
β_2	0.01	0.01	−0.05	−0.005	−0.03	−0.002	0.01	0.01	0.02
σ_2	(0.001)	(0.001)	(0.011)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.002)
R_2^2	0.28	0.28	0.28	0.27	0.49	0.37	0.43	0.27	0.26
					Nonlinear ST-SEM model – *known γ and c*				
β_1	57.96	43.93	2.50	−12.23	35.62	−10.58	60.99	75.69	70.24
σ_1	(18.24)	(9.90)	(15.29)	(13.82)	(14.44)	(8.15)	(20.66)	(12.38)	(16.56)
γ_1	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_2	−0.60	−0.56	−1.18	−0.53	−0.82	−0.50	−0.53	−0.48	−0.55
$Q(c_2)$	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
R_1^2	0.32	0.47	0.39	0.46	0.29	0.40	0.49	0.56	0.55
β_2	0.83	0.31	−6.31	0.57	−1.01	−0.11	1.07	1.07	0.98
σ_2	(0.14)	(0.14)	(1.58)	(0.22)	(0.55)	(0.16)	(0.20)	(0.15)	(0.17)
γ_2	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_1	−24.66	−8.98	−52.19	−21.91	−32.06	−18.59	−13.56	−10.61	−6.14
$Q(c_1)$	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
R_2^2	0.16	0.15	0.26	0.27	0.34	0.37	0.38	0.24	0.12
					Nonlinear ST-SEM model – *known γ and unknown c*				
β_1	93.34	49.80	−64.41	−48.01	51.69	−13.68	58.91	76.46	71.98
σ_1	(29.56)	(16.42)	(9.28)	(14.23)	(16.78)	(9.39)	(14.53)	(12.68)	(18.35)
γ_1	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_2	0.27	0.29	1.79	0.65	−1.51	−0.77	0.32	−0.10	0.11
$Q(c_2)$	0.79	0.80	0.85	0.80	0.16	0.21	0.80	0.61	0.65
R_1^2	0.38	0.48	0.53	0.49	0.31	0.41	0.51	0.57	0.57
β_2	0.80	0.38	−9.36	0.61	−2.57	−0.18	1.41	1.15	1.06
σ_2	(0.18)	(0.15)	(1.81)	(0.23)	(0.50)	(0.22)	(0.20)	(0.18)	(0.18)
γ_2	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
c_1	−17.56	−14.42	−63.08	−18.95	−5.88	−13.46	−21.63	−16.63	−2.11
$Q(c_1)$	0.36	0.26	0.15	0.32	0.50	0.36	0.22	0.21	0.37
R_2^2	0.18	0.15	0.32	0.27	0.43	0.37	0.43	0.26	0.13
					Nonlinear ST-SEM model – *unknown γ and c*				
β_1	68.39	46.47	−65.06	−51.03	57.89	−16.42	38.04	97.04	50.40
σ_1	(20.71)	(8.38)	(9.31)	(73.48)	(20.30)	(25.39)	(15.66)	(96.30)	(13.85)
γ_1	14.84	16.37	20.45	2.64	18.35	2.05	19.00	2.00	18.89
c_2	0.27	0.22	1.08	0.65	-1.44	−0.77	0.32	−0.03	-0.23
$Q(c_2)$	0.79	0.76	0.84	0.80	0.18	0.21	0.80	0.62	0.48
R_1^2	0.40	0.50	0.54	0.49	0.33	0.41	0.51	0.57	0.58
β_2	0.80	0.40	−9.38	0.63	−2.58	−0.19	1.44	1.16	1.06
σ_2	(0.13)	(0.17)	(1.85)	(0.23)	(0.55)	(0.26)	(0.25)	(0.19)	(0.24)
γ_2	17.73	15.78	3.82	18.41	19.36	17.18	18.64	18.51	16.28
c_1	−17.52	−18.11	−63.08	−18.78	−5.64	−13.88	−21.43	−16.54	−2.20
$Q(c_1)$	0.36	0.25	0.15	0.33	0.50	0.35	0.22	0.21	0.36
R_2^2	0.18	0.16	0.32	0.28	0.43	0.37	0.44	0.26	0.14

Notes: See notes under Table 8. We add to Eqs. 8 and 9 the consumer price inflation (CPI) with one month lag ($k_1 = 1$) as a common explanatory variable to check the robustness of the results. The augmented system becomes:
$\text{BR}_t = \delta_1 + \mu_1 \text{CPI}_{t-k_1} + \alpha_{1,1}\text{DY}_{t-k_1} + \alpha_{1,2}\text{VIX}_{t-k_1} + \beta_1 G(\text{SY}_t; \gamma_1, c_2) + u_{1,t}$
$\text{SY}_t = \delta_2 + \mu_2 \text{CPI}_{t-k_1} + \alpha_{2,1}\text{DB}_{t-k_2} + \alpha_{2,2}\text{FB}_{t-k_2} + \alpha_{2,3}\text{RT}_{t-k_2} + \beta_2 G(\text{BR}_t; \gamma_2, c_1) + u_{2,t}$

Table A.7
Coefficients of the control variables for the model augmented with inflation.

	Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
				Linear-model estimation results (OLS)					
μ_1	−1.02	−10.50	9.13	−1.33	2.26	−1.47	−2.41	−0.24	1.87
σ_{μ_1}	(1.86)	(1.75)	(1.43)	(1.46)	(1.87)	(1.05)	(1.38)	(0.97)	(0.86)
$\alpha_{1,1}$	−4.04	−7.54	7.80	−6.51	0.67	−4.82	−9.36	−6.95	−7.86
$\sigma_{1,1}$	(1.04)	(0.89)	(1.89)	(0.80)	(1.38)	(0.74)	(0.99)	(0.80)	(1.01)
$\alpha_{1,2}$	−0.94	−1.17	−3.09	−1.04	−1.89	−1.31	−1.50	−0.54	−1.07
$\sigma_{1,2}$	(0.26)	(0.18)	(0.37)	(0.20)	(0.35)	(0.17)	(0.14)	(0.16)	(0.13)
δ_1	29.94	78.47	11.57	45.38	20.72	51.30	56.57	40.60	42.07
σ_{δ_1}	(4.87)	(4.83)	(6.77)	(4.27)	(6.73)	(5.42)	(3.98)	(3.75)	(3.49)
μ_2	0.12	0.25	1.40	0.31	0.69	0.17	0.36	0.04	0.02
σ_{μ_2}	(0.04)	(0.05)	(0.34)	(0.07)	(0.10)	(0.04)	(0.04)	(0.04)	(0.04)
$\alpha_{2,1}$	−0.01	0.001	0.10	0.04	0.06	0.03	−0.03	−0.001	0.0001
$\sigma_{2,1}$	(0.004)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.002)	(0.003)
$\alpha_{2,2}$	0.14	0.08	−0.30	0.16	−0.15	−0.07	−0.12	0.05	0.02
$\sigma_{2,2}$	(0.04)	(0.05)	(0.10)	(0.04)	(0.07)	(0.02)	(0.05)	(0.02)	(0.02)
$\alpha_{3,2}$	–	0.10	0.48	0.16	0.55	0.29	–	–	–
$\sigma_{3,2}$	–	(0.13)	(0.22)	(0.06)	(0.15)	(0.05)	–	–	–
δ_2	0.53	−2.69	−25.79	−7.67	−16.65	−8.32	1.29	−0.15	−0.18
σ_{δ_2}	(0.31)	(3.30)	(7.44)	(2.22)	(4.32)	(1.40)	(0.35)	(0.15)	(0.24)
				Nonlinear ST-SEM model – *known γ and c*					
μ_1	−2.74	−11.54	7.23	−1.58	−1.49	−1.15	−5.90	−0.61	2.84
σ_{μ_1}	(2.26)	(2.11)	(2.06)	(2.15)	(2.39)	(1.25)	(3.03)	(1.31)	(1.17)
$\alpha_{1,1}$	−3.83	−8.07	9.08	−6.45	−4.35	−4.80	−9.59	−6.34	−7.41
$\sigma_{1,1}$	(1.12)	(0.93)	(1.99)	(0.99)	(1.70)	(0.76)	(1.09)	(1.08)	(1.38)
$\alpha_{1,2}$	−0.73	−1.09	−3.49	−1.05	−1.65	−1.30	−1.05	0.12	−0.06
$\sigma_{1,2}$	(0.32)	(0.19)	(0.41)	(0.22)	(0.43)	(0.17)	(0.32)	(0.25)	(0.35)
δ_1	−15.30	47.71	18.76	54.55	12.00	57.79	6.44	−23.61	−29.33
σ_{δ_1}	(15.04)	(7.27)	(8.43)	(7.05)	(10.22)	(5.78)	(18.94)	(12.22)	(18.75)
μ_2	0.17	0.22	1.52	0.45	0.74	0.17	0.33	0.06	0.04
σ_{μ_2}	(0.05)	(0.05)	(0.33)	(0.08)	(0.11)	(0.04)	(0.04)	(0.04)	(0.04)
$\alpha_{2,1}$	−0.02	0.001	0.13	0.04	0.07	0.03	-0.01	0.001	0.002
$\sigma_{2,1}$	(0.005)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.002)	(0.003)
$\alpha_{2,2}$	0.17	0.08	-0.30	0.16	−0.14	-0.07	0.06	0.064	0.004
$\sigma_{2,2}$	(0.05)	(0.06)	(0.10)	(0.05)	(0.07)	(0.02)	(0.09)	(0.02)	(0.02)
$\alpha_{3,2}$	–	0.09	0.59	0.13	0.59	0.29	–	–	–
$\sigma_{3,2}$	–	(0.16)	(0.22)	(0.07)	(0.17)	(0.06)	–	–	–
δ_2	0.31	−2.50	−25.40	−7.92	−17.23	−8.34	−0.61	−1.09	−1.07
σ_{δ_2}	(0.34)	(4.04)	(7.27)	(2.52)	(4.87)	(1.63)	(0.63)	(0.22)	(0.27)
				Nonlinear ST-SEM model – *known γ and unknown c*					
μ_1	−2,19	−10.17	10.98	1.13	−1.83	−0.96	−4.91	−0.72	−0.02
σ_{μ_1}	(2.83)	(2.49)	(1.64)	(1.96)	(2.31)	(1.39)	(2.48)	(1.29)	(1.88)
$\alpha_{1,1}$	-4.27	-7.10	0.82	-6.18	-4.83	-4.83	-8.36	-6.30	-6.06
$\sigma_{1,1}$	(1.13)	(1.03)	(2.98)	(1.02)	(1.66)	(0.99)	(1.15)	(1.06)	(1.86)
$\alpha_{1,2}$	-0.50	-1.11	-2.16	-1.22	-1.66	-1.28	-1.30	0.12	-0.27
$\sigma_{1,2}$	(0.45)	(0.22)	(0.43)	(0.28)	(0.42)	(0.25)	(0.27)	(0.25)	(0.36)
δ_1	-5.66	57.79	13.24	55.10	-3.45	60.26	32.78	-10.67	-3.69
σ_{δ_1}	(18.89)	(6.82)	(6.34)	(5.59)	(15.55)	(14.98)	(11.45)	(12.35)	(16.35)
μ_2	0.18	0.23	1.41	0.46	0.68	0.18	0.35	0.11	0.04
σ_{μ_2}	(0.05)	(0.06)	(0.41)	(0.08)	(0.12)	(0.04)	(0.07)	(0.05)	(0.04)
$\alpha_{2,1}$	−0.02	0.004	0.17	0.04	0.06	0.03	−0.01	−0.002	0.001
$\sigma_{2,1}$	(0.01)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)	(0.02)	(0.004)	(0.003)
$\alpha_{2,2}$	0.20	0.09	−0.43	0.15	−0.14	−0.07	0.06	0.04	−0.002
$\sigma_{2,2}$	(0.07)	(0.06)	(0.11)	(0.05)	(0.11)	(0.02)	(0.10)	(0.04)	(0.02)
$\alpha_{3,2}$	–	0.13	0.90	0.13	0.56	0.28	–	–	–
$\sigma_{3,2}$	–	(0.16)	(0.31)	(0.07)	(0.21)	(0.07)	–	–	–
δ_2	0.46	−3.59	−32.96	−8.05	−14.65	−7.84	−0.98	−1.19	−0.99
σ_{δ_2}	(0.65)	(4.21)	(8.98)	(2.55)	(5.99)	(2.18)	(0.78)	(0.30)	(0.33)
				Nonlinear ST-SEM model – *unknown γ and c*					
μ_1	0.09	−10.83	10.11	1.36	−2.59	−0.85	−2.48	−0.63	0.94
σ_{μ_1}	(2.09)	(2.24)	(1.68)	(3.50)	(2.36)	(1.55)	(3.53)	(1.54)	(2.48)
$\alpha_{1,1}$	−5.20	−6.87	2.05	−6.09	−5.08	−4.77	−7.76	−6.14	−8.51
$\sigma_{1,1}$	(1.30)	(0.97)	(3.91)	(1.18)	(1.63)	(1.17)	(1.32)	(1.26)	(1.63)
$\alpha_{1,2}$	−0.47	−0.98	−1.94	−1.23	−1.70	−1.28	−1.50	0.11	0.005
$\sigma_{1,2}$	(0.35)	(0.21)	(0.63)	(0.35)	(0.41)	(0.28)	(0.26)	(0.28)	(0.52)
δ_1	3.74	57.75	10.14	55.64	−6.34	61.34	42.44	−19.20	−3.36
σ_{δ_1}	(9.97)	(6.18)	(10.14)	(7.88)	(17.41)	(27.40)	(8.47)	(43.22)	(27.06)
μ_2	0.18	0.24	1.41	0.47	0.68	0.18	0.34	0.11	0.04
σ_{μ_2}	(0.05)	(0.05)	(0.44)	(0.08)	(0.12)	(0.04)	(0.07)	(0.06)	(0.05)
$\alpha_{2,1}$	−0.02	0.01	0.17	0.04	0.06	0.03	−0.01	−0.001	0.001

(*continued on next page*)

Table A.7 (*continued*)

	Germany	France	Greece	Italy	Portugal	Spain	Switzerland	UK	US
$\sigma_{2,1}$	(0.01)	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.01)	(0.005)	(0.004)
$\alpha_{2,2}$	0.20	0.10	−0.44	0.15	−0.14	−0.07	0.07	0.04	−0.002
$\sigma_{2,2}$	(0.05)	(0.06)	(0.11)	(0.05)	(0.12)	(0.02)	(0.08)	(0.04)	(0.02)
$\alpha_{3,2}$	–	0.15	0.90	0.13	0.56	0.28	–	–	–
$\sigma_{3,2}$	–	(0.16)	(0.33)	(0.07)	(0.25)	(0.09)	–	–	–
δ_2	0.46	−4.07	−33.15	−8.14	−14.64	−7.78	−1.13	−1.20	−0.99
σ_{δ_2}	(0.42)	(4.24)	(9.79)	(2.53)	(7.11)	(2.61)	(0.65)	(0.38)	(0.34)

Notes: Complement to the estimation results in Table A.6

A2.2. Third-order Taylor expansion

We proceed in the same way for the construction of an auxiliary regression based on the third-order expansion of the logistic function. We compute the second and the third derivative of $G(y_{jt};\gamma_i,c_j)$ and evaluate them at $\gamma_i=0$. Given that the logistic function is odd, the second derivative is 0.

$$\frac{\partial^2 G(y_{jt};\gamma_i,c_j)}{\partial\gamma_i^2}=G(y_{jt};\gamma_i,c_j)\big[1-G(y_{jt};\gamma_i,c_j)\big]\big[1-2G(y_{jt};\gamma_i,c_j)\big](y_{jt}-c_j)^2$$

$$\left.\frac{\partial^2 G(y_{jt};\gamma_i,c_j)}{\partial\gamma_i^2}\right|_{\gamma_i=0}=0$$

$$\frac{\partial^3 G(y_{jt};\gamma_i,c_j)}{\partial\gamma_i^3}=G(y_{jt};\gamma_i,c_j)\big[1-G(y_{jt};\gamma_i,c_j)\big]\big[6G(y_{jt};\gamma_i,c_j)^2-6G(y_{jt};\gamma_i,c_j)+1\big](y_{jt}-c_j)^3$$

$$\left.\frac{\partial^3 G(y_{jt};\gamma_i,c_j)}{\partial\gamma_i^3}\right|_{\gamma_i=0}=-\frac{1}{8}(y_{jt}-c_j)^3$$

The third-order Taylor expansion of $\tilde{G}(y_{jt};\gamma_i,c_j)$ – with R_{3i} the remainder – is given by the following equation.

$$\tilde{G}(y_{jt};\gamma_i,c_j)=-\frac{1}{4}\gamma_i c_j+\frac{1}{48}\gamma_i^3c_j^3+\left(\frac{1}{4}\gamma_i-\frac{1}{16}\gamma_i^3c_j^2\right)y_{jt}+\frac{1}{16}\gamma_i^3c_jy_{jt}^2-\frac{1}{48}\gamma_i^3y_{jt}^3+R_{3i}$$

Finally, the auxiliary regression can be written as:

$$y_{it}=\tilde{\delta}_i'z_t+\alpha_i'x_{it}+\tilde{\beta}_{i,1}y_{jt}+\tilde{\beta}_{i,2}y_{jt}^2+\tilde{\beta}_{i,3}y_{jt}^3+u_{it}^*$$

with $\tilde{\delta}_i=[(\delta_{i,1}-\frac{1}{4}\gamma_ic_j+\frac{1}{48}\gamma_i^3c_j^3)\ \delta_i^{-\prime}]'$, with δ_i^- defined as above, $\tilde{\beta}_{i,1}=\beta_i\gamma_i(\frac{1}{4}-\frac{1}{16}\gamma_i^2c_j^2)$, $\tilde{\beta}_{i,2}=\frac{1}{16}\beta_i\gamma_i^3c_j$, $\tilde{\beta}_{i,3}=-\frac{1}{48}\beta_i\gamma_i^3$ and $u_{it}^*=\beta_iR_{3i}+u_{it}$.

References

Acharya, V.V., Drechsler, I., Schnabl, P., 2014. A pyrrhic victory? bank bailouts and sovereign credit risk. Journal of Finance 69 (6), 2689–2739.
Ahnert, T., Bertsch, C., 2017. A wake-up call theory of contagion. Sveriges Riksbank Working Paper Series No. 294.
Amemiya, T., 1974. The nonlinear two-stage least-squares estimator. J Econom 2 (2), 105–110.
Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61 (4), 821–856.
Andrews, D.W.K., Cheng, X., 2013. Maximum likelihood estimation and uniform inference with sporadic identification failure. J Econom 173 (1), 36–56.
Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62 (6), 1383–1414.
Bae, K., Karolyi, G.A., Stulz, R., 2003. A new approach to measuring financial contagion. Review of Financial Studies 16 (3), 717–763.
Bekaert, G., Hodrick, R.J., Zhang, X., 2009. International stock return comovements. Journal of Finance 64 (6), 2591–2626.
Blundell, R., Smith, R.J., 1994. Coherency and estimation in simultaneous models with censored or qualitative dependent variables. J Econom 64 (1–2), 355–373.
Cappiello, L., Gérard, B., Kadareja, A., Manganelli, S., 2014. Measuring comovements by regression quantiles. Journal of Financial Econometrics 12 (4), 645–678.
Chan, K.S., Tong, H., 1986. On estimating thresholds in autoregressive models. Journal of Time Series Analysis 7 (3), 179–190.
Corsetti, G., Pericoli, M., Sbracia, M., 2005. Some contagion, some interdependence: more pitfalls in tests of financial contagion. J Int Money Finance 24 (8), 1177–1199.
Davies, R.B., 1977. Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 64 (2), 247–254.
Davies, R.B., 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 74 (1), 33–43.
Dees, S., di Mauro, F., Pesaran, M.H., Smith, V., 2007. Exploring the international linkages of the euro area: a global VAR analysis. Journal of Applied Econometrics 22 (1), 1–38.
Teräsvirta, T., Tjøstheim, D., Granger, C.W.J., 2010. Modeling Nonlinear Economic Time Series. Oxford University Press.
van Dijk, D., Teräsvirta, T., Franses, P.H., 2002. Smooth transition autoregressive models – a survey of recent developments. Econom Rev 21 (1), 1–47.
Dungey, M., Martin, V.L., 2007. Unravelling financial market linkages during crises. Journal of Applied Econometrics 22 (1), 89–119.
Ehrmann, M., Fratzscher, M., Rigobon, R., 2011. Stocks, bonds, money markets and exchange rates: measuring international financial transmission. Journal of Applied Econometrics 26 (6), 948–974.
Eitrheim, O., Teräsvirta, T., 1996. Testing the adequacy of smooth transition autoregressive models. J Econom 74 (1), 59–75.
Forbes, K., Rigobon, R., 2002. No contagion, only interdependence: measuring stock market comovements. Journal of Finance 57 (5), 2223–2261.
Forbes, K.J., 2012. The "Big C": Identifying and mitigating contagion. Working paper.

González, A., Teräsvirta, T., 2006. Simulation-based finite sample linearity test against smooth transition models. Oxf Bull Econ Stat 68 (1), 797–812.
Gourieroux, C., Laffont, J.J., Monfort, A., 1980. Coherency conditions in simultaneous linear equation models with endogenous switching regimes. Econometrica 48 (3), 675–695.
Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64 (2), 413–430.
Hansen, B.E., 2000. Sample splitting and threshold estimation. Econometrica 68 (3), 575–603.
Kallestrup, R., Lando, D., Murgoci, A., 2016. Financial sector linkages and the dynamics of bank and sovereign credit spreads. Journal of Empirical Finance 38 (part A), 374–393.
Kelejian, H.H., 1971. Two-stage least squares and econometric systems linear in parameters but nonlinear in the endogenous variables. J Am Stat Assoc 66 (334), 373–374.
Lewbel, A., 2007. Coherency and completeness of structural models containing a dummy endogenous variable. Int Econ Rev (Philadelphia) 48 (4), 1379–1392.
Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. Journal of Finance 56 (2), 649–676.
Lundbergh, S., Teräsvirta, T., 2006. A time series model for an exchange rate in a target zone with applications. J Econom 131 (1–2), 579–609.
Luukkonnen, R., Saikkonen, P., Teräsvirta, T., 1988. Testing linearity against smooth transition autoregressive models. Biometrika 75 (3), 491–499.
Maringer, D.G., Meyer, M., 2008. Smooth transition autoregressive models – new approaches to the model selection problem. Nonlinear Dynamical Methods and Time Series Analysis 12 (1). Article 5
Pesaran, M.H., Pick, A., 2007. Econometric issues in the analysis of contagion. Journal of Economic Dynamics and Control 31 (4), 1245–1277.
Pick, A., 2007. Financial contagion and tests using instrumental variables. Netherlands Central Bank, working paper.
Aït Sahalia, Y., Laeven, R.J.A., Pelizzon, L., 2014. Mutual excitation in eurozone credit CDS. J Econom 183 (2), 151–167.
Schleer, F., 2015. Finding starting-values for the estimation of vector STAR models. Econometrics 3 (1), 65–90.
Storn, R., Price, K., 1997. Differential evolution - a simple and efficient heuristic and global optimization over continuous spaces. J. Global Optim. 11 (4), 341–359.
Teräsvirta, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. J Am Stat Assoc 89 (425), 208–218.
Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press.