The estimation of threshold models in price transmission analysis

Friederike Greb, Stephan von Cramon-Taubadel, Tatyana Krivobokova and Axel Munk

The threshold vector error correction model is a popular tool for the analysis of spatial price transmission and market integration. In the literature, the profile likelihood estimator is the preferred choice for estimating this model. Yet, in certain settings this estimator performs poorly. In particular, if the true thresholds are such that one or more regimes contain only a small number of observations, if unknown model parameters are numerous or if parameters differ little between regimes, the profile likelihood estimator displays large bias and variance. Such settings are likely when studying price transmission. For simpler, but related threshold models Greb et al. (2011) have developed an alternative estimator, the regularized Bayesian estimator, which does not exhibit these weaknesses. We explore the properties of this estimator for threshold vector error correction models. Simulation results show that it outperforms the profile likelihood estimator, especially in situations in which the profile likelihood estimator fails. Two empirical applications – a reassessment of the seminal paper by Goodwin and Piggott (2001), and an analysis of price transmission between German and Spanish markets for pork – demonstrate the relevance of the new approach for spatial price transmission analysis.

Key words: Bayesian estimator, market integration, price transmission, spatial arbitrage, TVECM.

When assessing the integration of spatially separated markets, agricultural economists typically analyze the transmission of price shocks between these markets (Fackler and Goodwin 2001). The law of one price (LOP) states that prices for a homogeneous good at different locations should differ by no more than the

---

Friederike Greb is a PostDoc at the Department of Agricultural Economics and Rural Development of Göttingen University. Stephan von Cramon-Taubadel holds the chair of agricultural policy at the same department. Tatyana Krivobokova is a professor at Göttingen University's Courant Research Centre "Poverty, Equity and Growth in Developing Countries" and the Institute for Mathematical Stochastics. Axel Munk is the director of the Institute for Mathematical Stochastics and a Max-Planck Fellow at the Max-Planck Insitute for Biophysical Chemistry. The authors are responsible for any errors.

transaction costs of trading the good between these locations. Otherwise traders will engage in spatial arbitrage, which increases the price at the low-price location and reduces the price at the high-price location until the LOP is restored.

In spatial equilibrium, the manner in which price shocks are transmitted between two locations will therefore depend on the magnitude of the price difference between these locations (Goodwin and Piggott 2001; Stephens et al. 2011). Shocks that increase the price difference so that it exceeds the costs of trade between the two locations will lead to arbitrage and price transmission. However, if the price difference remains less than these transaction costs, arbitrage will not be profitable and there will be no price transmission. The result is referred to in the literature as "regime-dependent" price transmission. Specifically, the spatial equilibrium model described above will lead to three regimes delineated by two threshold values that equal the transaction costs of trade in one and the other direction, respectively. In the outer regimes where the price difference is greater than the transaction costs of trade in the one or the other direction, arbitrage will lead to the transmission of price shocks. If the price difference lies within the "band of inaction" between these transaction costs, prices can evolve independently of one another. The costs of trade between two locations need not be symmetric; for example, river transport might be more expensive going upstream than it is going downstream. Hence, the thresholds that define the boundaries of the spatial price transmission regimes will have opposite signs and possibly different magnitudes.

Threshold vector error correction models (TVECMs) are frequently used to model this regime-dependent spatial price transmission process. TVECMs became popular with Balke and Fomby's (1997) article on threshold cointegration. Goodwin and Piggott's (2001) seminal paper established TVECMs in price transmission analysis, and dozens of applications have followed. As an indication of the ongoing popularity of the TVECM, a search of the AgEconSearch website (www.ageconsearch.umn.edu) on December 20, 2012 with the keywords "price transmission" and "threshold" produced 17 papers posted since 2010.

Typically, and as we explain in greater detail below, thresholds in TVECMs are estimated by maximizing the profile likelihood (Hansen and Seo 2002). However, in many settings, this estimator is biased and has a high variance. Lo and Zivot (2001) and Balcombe, Bailey, and Brooks (2007) acknowledge this problem. Profile likelihood estimates are especially prone to be unreliable in situations characterized by large numbers of unknown model parameters besides the thresholds, when there is little difference between adjoining regimes, and when the location of the thresholds leaves only few observations in one of the regimes (which is inevitable in small samples). These problems are generic and emerge in many econometric settings, but they are particularly acute when profile likelihood is used to estimate TVECMs.

To cope with these shortcomings, several strategies are proposed in the literature. Perhaps the best of these is the modified profile likelihood function introduced by Barndorff-Nielsen (1983). However, the proposed modifications are usually based on regularity assumptions that do not hold for the TVECM. A further weakness of the profile likelihood estimator is that it depends on an arbitrary trimming parameter that ensures that each regime contains a minimum number of observations and, thus, that estimation of the model parameters in that regime is possible. This can be a problematic restriction when modeling spatial price transmission. If market integration is strong, differences in prices between two locations that exceed the transaction cost thresholds – and therefore fall into one of the outer regimes – will be corrected quickly. In this case, there will be few observations in the outer regimes, and a trimming parameter which forces more observations into these regimes will inevitably lead to unreliable estimates of both the threshold values and the model parameters in each regime. Estimation is not necessarily easier if the price data originate from markets that are poorly integrated because in this case the weak price transmission displayed in the outer regimes may be observationally quite similar to the independent price movements in the inner "band of inaction". Finally, the non-differentiability of the TVECM's likelihood function with respect to the thresholds exacerbates computation of its maximum, which can also be a source of imprecise estimates.

These problems with the profile likelihood estimator suggest that there is a need to rethink the estimation of TVECMs in price transmission analysis. In this article we investigate the suitability of an alternative threshold estimator developed for generalized threshold regression models (Greb et al. 2011). Among its advantages, this alternative estimator does not require a trimming parameter. We demonstrate using Monte Carlo experiments that this so-called regularized Bayesian estimator clearly outperforms the profile likelihood estimator not only for generalized threshold regression models, but also specifically for TVECMs, even in settings in which the profile likelihood estimator is highly biased and variable. We also show that although employing the regularized Bayesian estimator is technically easy, careful numerical implementation – even if it is computationally intensive – can be decisive. Of course, it is important to go beyond the demonstration of the superior statistical properties of the regularized Bayesian threshold estimator, and to consider as well its implications for empirical price transmission analysis using TVECMs. Here, it is crucial to interpret not only the estimated threshold parameters, but also the parameters that describe the dynamics of price transmission within each regime. We draw on two empirical applications to illustrate this.

The rest of this article is organized as follows. In the next section, we specify the TVECM, discuss existing threshold estimators and their deficiencies, present the alternative estimator, and comment on computational pitfalls in threshold estimation. Subsequently, we illustrate the performance of the new estimator by means of a simulation study. As empirical applications we first revisit the analysis of spatial market integration for four corn and soybean markets in North Carolina detailed in the seminal contribution by Goodwin and Piggott (2001), and second analyze spatial price transmission between German and Spanish pork markets. The last section concludes.

Theory

We begin this section by specifying the TVECM and discussing the methods that have been used to estimate it. This is followed by a presentation of the regularized Bayesian estimation method that we propose.

The Threshold Vector Error Correction Model

Observations $p_t = (p_{1,t}, p_{2,t})'$, $t = 1 \dots n$, of a two-dimensional time series generated by a TVECM with three different regimes, which are characterized by parameters $\rho_k, \theta_k \in \mathbb{R}^2$ and $\Theta_{km} \in \mathbb{R}^{2\times 2}$ for $k = 1, 2, 3$ and $m = 1, \dots, M$, can be written as

$$(1) \quad \Delta p_t = \begin{cases} \rho_1 \gamma' p_{t-1} + \theta_1 + \sum_{m=1}^{M} \Theta_{1m} \Delta p_{t-m} + \varepsilon_t & , \quad \gamma' p_{t-1} \leq \psi_1 \quad \text{(Regime 1)} \\ \rho_2 \gamma' p_{t-1} + \theta_2 + \sum_{m=1}^{M} \Theta_{2m} \Delta p_{t-m} + \varepsilon_t & , \quad \psi_1 < \gamma' p_{t-1} \leq \psi_2 \quad \text{(Regime 2)} \\ \rho_3 \gamma' p_{t-1} + \theta_3 + \sum_{m=1}^{M} \Theta_{3m} \Delta p_{t-m} + \varepsilon_t & , \quad \psi_2 < \gamma' p_{t-1} \quad \text{(Regime 3).} \end{cases}$$

We assume that p_t forms an $I(1)$ time series with cointegrating vector $\gamma \in \mathbb{R}^2$ and error correction term $\gamma' p_t$. We further assume that the errors denoted by ε_t have expected value $\mathrm{E}(\varepsilon_t) = 0$ and covariance matrix $\mathrm{Cov}(\varepsilon_t) = \Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \in (\mathbb{R}^+)^{2\times 2}$.

We call ψ_1, ψ_2 the threshold parameters and define the threshold parameter space Ψ to include all $\psi = (\psi_1, \psi_2)$ such that $\min(\gamma' p_t) < \psi_1 < \psi_2 < \max(\gamma' p_t)$, where $\min(\gamma' p_t)$ and $\max(\gamma' p_t)$ are, respectively, the lowest and highest values of the error correction term. Although all of the coefficients in equation (1) can vary across regimes, some of them can remain constant.

In the spatial equilibrium setting, $p_{1,t}$ and $p_{2,t}$ are prices at different locations and γ is often taken to equal $(1, -1)'$ so that the error correction term $\gamma' p_t$ measures the difference between p_1 and p_2 at time t. The threshold ψ_1 (ψ_2) corresponds to the transaction costs of trade from location 1 to location 2 (location 2 to location 1). Regimes 1 and 3 are the outer regimes in which the violation of spatial equilibrium leads to arbitrage and price transmission, and regime 2 represents the inner "band of inaction". Not only the estimates of the threshold parameters $\psi = (\psi_1, \psi_2)$ are of economic interest, however. The estimates of ρ_k $(k = 1, 2, 3)$ (often referred to as the "adjustment parameters") are also of interest as they measure the speed with which violations of spatial equilibrium between two locations are corrected in the respective regimes.

To express the model in matrix notation, we define vectors Δp_i and ε_i by stacking the ith components of Δp_t and ε_t, respectively; and $I\left(\gamma' p \leq \psi_1\right)$, $I\left(\psi_1 < \gamma' p \leq \psi_2\right)$, and $I\left(\psi_2 < \gamma' p\right)$ by stacking $I\left(\gamma' p_{t-1} \leq \psi_1\right)$, $I\left(\psi_1 < \gamma' p_{t-1} \leq \psi_2\right)$ and $I\left(\psi_2 < \gamma' p_{t-1}\right)$, respectively. $I(\cdot)$ denotes the indicator function. For observations at n time points, an $n \times d$ matrix X is constructed by stacking rows $x_t' = (\gamma' p_{t-1}, 1, \Delta p_{t-1}', \ldots, \Delta p_{t-M}')$ of length $d = 2M + 2$. $\beta_{i,k}$ is the ith column of the matrix $(\rho_k, \theta_k, \Theta_{k1}, \ldots, \Theta_{kM})'$, $i = 1, 2$ and $k = 1, 2, 3$. With $\mathrm{diag}\left\{I(\cdot)\right\}$ defined as the diagonal matrix with entries $I(\cdot)$ in the diagonal, we can write

$$(2) \quad \Delta p_i = \text{diag}\{I(\gamma' p \leq \psi_1)\} X\beta_{i,1} + \text{diag}\{I(\psi_1 < \gamma' p \leq \psi_2)\} X\beta_{i,2} + \text{diag}\{I(\psi_2 < \gamma' p)\} X\beta_{i,3} + \varepsilon_i = X_1\beta_{i,1} + X_2\beta_{i,2} + X_3\beta_{i,3} + \varepsilon_i$$

for $i = 1, 2$. This leads to a compact representation of model (1),

$$(3) \quad \Delta p = \begin{pmatrix} \Delta p_1 \\ \Delta p_2 \end{pmatrix} = (I_2 \otimes X_1)\beta_1 + (I_2 \otimes X_2)\beta_2 + (I_2 \otimes X_3)\beta_3 + \varepsilon,$$

where $\beta_k' = \left(\beta_{1,k}', \beta_{2,k}'\right)$ for $k = 1, 2, 3$, $I_2 \in \mathbb{R}^{2\times 2}$ denotes the identity matrix, and $X = X_1 + X_2 + X_3$.

A variety of modifications and restrictions of the general TVECM (1) have been implemented in price transmission studies. Lo and Zivot (2001) and Ihle (2010, table 2.1) provide details on a number of important specifications. We limit attention to the general TVECM. Restrictions of the model imply further information about the parameters (or relations between them) and, hence, facilitate estimation. The most general case is thus the most challenging. Although the TVECM can be generalized to include r thresholds and $r + 1$ regimes, we focus on a TVECM with two thresholds and three regimes as this is the version of the TVECM that is grounded in spatial equilibrium theory as outlined above. Generalization is straightforward.

Commonly used threshold estimators

The most frequently used threshold estimator in the econometrics literature is the profile likelihood estimator (Hansen and Seo 2002; Lo and Zivot 2001). According to this method, for each possible pair of the threshold parameters $\psi = (\psi_1, \psi_2)$ the remaining parameters in the likelihood function corresponding to (1) are replaced by their maximum likelihood estimates. The pair of thresholds that maximizes the resulting profile likelihood function is selected as the estimate. More precisely, denoting the log-likelihood function of (1) by $\ell\left(\psi, \beta_1, \beta_2, \beta_3, \Sigma\right)$, the profile likelihood estimator is defined as

$$(4) \qquad \hat{\psi}_{pL} = \arg\max \ell_p(\psi) \qquad \text{with} \qquad \ell_p(\psi) = \ell\left(\psi, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \widehat{\Sigma}\right)$$

and $\hat{\beta}_k$ and $\widehat{\Sigma}$ the maximum likelihood estimates of β_k and Σ. Hence,

$$(5) \qquad \ell_p(\psi) \propto -\left\{\Delta p - (I_2 \otimes X_1)\hat{\beta}_1 - (I_2 \otimes X_2)\hat{\beta}_2 - (I_2 \otimes X_3)\hat{\beta}_3\right\}' \left\{\Delta p - (I_2 \otimes X_1)\hat{\beta}_1 - (I_2 \otimes X_2)\hat{\beta}_2 - (I_2 \otimes X_3)\hat{\beta}_3\right\}$$

and $\hat{\beta}_k = \left\{(I_2 \otimes X_k)'(I_2 \otimes X_k)\right\}^{-1}(I_2 \otimes X_k)'\Delta p$, $k = 1, 2, 3$. Since this (logged) profile likelihood function $\ell_p(\psi)$ is not differentiable with respect to the threshold parameters, the thresholds that maximize it are determined by calculating (5) for each point on a two-dimensional grid of possible threshold values, which is why the literature often refers to the "grid search" method.

The bias and high variance of the profile likelihood threshold estimator are mentioned but not further pursued in the literature on TVECMs (see table 4 and figure 1 in Lo and Zivot 2001). The simulation results we present below confirm the existence of these weaknesses (see table 1 and figures 1 and 2). Greb et al. (2011) provide a detailed analysis of the problems associated with the profile likelihood approach to threshold estimation. In summary, there are two principal problems: i) the dependence on an arbitrary trimming parameter; and ii) the uncertainty inherent in the $\hat{\beta}_k$ which are estimated for each combination of possible threshold values. These problems can be pronounced in small samples.

In spatial arbitrage modeling, the first issue can be decisive. ψ places each observation into one of three regimes. In order to compute $\hat{\beta}_k$, it is essential that at least $d = \dim(\beta_{i,k})$ observations fall into the k-th regime. To ensure this, ψ_1 must be greater than or equal to $\gamma' p_{(d)}$, where $\gamma' p_{(1)}, \ldots, \gamma' p_{(n)}$ is the ordered sequence of error correction terms, and ψ_2 must be correspondingly less than $\gamma' p_{(n-d)}$. The trimming parameter restricts ψ accordingly. A variety of trimming parameters are suggested in the literature. Goodwin and Piggott (2001) specify that each regime in the TVECM that they estimate must include at least 25 observations. Balcombe, Bailey, and Brooks (2007) impose the restriction that each regime must include at least 20% of the observations in their sample, while Andrews (1993) proposes a minimum proportion of 15%. However, if markets are well-integrated, then arbitrage will lead to rapid correction of any price differences that exceed the thresholds, and the outer regimes will contain correspondingly few observations. Especially in small samples, this can lead to a situation in which the outer regimes actually contain fewer observations than imposed by the chosen trimming parameter. In this case, the resulting estimator cannot be consistent as the threshold parameter space Ψ (and, hence, the grid that is searched) excludes the true thresholds. Despite its potential impact on threshold estimation, the literature only offers several arbitrary suggestions for the trimming parameter.

The second problem naturally becomes more pronounced as the number of parameters in the model (i.e. the dimension of β_k) increases. Each additional lag included in a bivariate TVECM with three regimes adds 12 coefficients. Hence, the number of coefficients to be estimated can grow rapidly relative to the potentially few observations in the outer regimes. If there is also little difference in coefficients between regimes, pinpointing the location of the thresholds becomes increasingly difficult.

As an alternative to profile likelihood, Bayesian estimators have been employed in some price transmission studies (Balcombe, Bailey, and Brooks 2007; Balcombe and Rapsomanikis 2008). As explained in Greb et al. (2011), the performance of a

Bayesian estimator in generalized threshold regression models crucially depends on the selected priors. In the absence of any prior knowledge of potential parameter values, so-called noninformative priors are the natural choice. However, these can distort estimates. In particular, the posterior density associated with noninformative priors for the β_k inherits the dependence on a trimming parameter from the profile likelihood function. Indeed, Greb et al. (2011) show that the posterior density takes its largest values exactly for those threshold values that are arbitrarily included or excluded from the threshold parameter space Ψ when the trimming parameter is varied. Consequently, the trimming parameter strongly affects the threshold estimate. Nevertheless, Balcombe, Bailey, and Brooks (2007) and Balcombe and Rapsomanikis (2008) base their Bayesian estimators on noninformative priors. Chen (1998) suggests a Bayesian estimator based on a normal prior with known hyper-parameters for the β_k and a uniform prior for the threshold parameter. However, she designs the latter to assign zero probability to threshold values that do not leave a minimum number of observations in each regime, which is equivalent to assuming an arbitrary trimming parameter.

Regularized Bayesian estimator

Given the deficiencies of profile likelihood and Bayesian estimation with noninformative priors, we explore the properties of an alternative threshold estimator in the context of TVCEMs. This regularized Bayesian (rB) estimator was developed for univariate generalized threshold regression models with one threshold (Greb et al. 2011). The idea of the estimator is to penalize differences between regimes so as to keep these differences reasonably small when the data contain little information. The strength of this regularizing penalty is fundamental to the estimator. It is determined in a data-driven manner employing the so-called empirical Bayes paradigm. The estimator is developed in a Bayesian framework and the penalization is a result of the choice of priors. As an important consequence of the regularization, the posterior density is well-defined on the entire threshold parameter space Ψ. Hence, there is no need to choose a trimming parameter and no risk of excluding the true threshold from Ψ. In the setting of generalized threshold regression models, the rB estimator outperforms commonly used estimators, especially when the threshold leaves only few observations in one of the regimes or coefficients differ little between regimes.

Extension of the theory detailed in Greb et al. (2011) to the TVECM with two thresholds in equation (1) is straightforward. It involves reparametrizing the model in equation (3),

$$\begin{aligned}(6)\qquad \Delta p &= (I_2 \otimes X_1)\beta_1 + (I_2 \otimes X_2)\beta_2 + (I_2 \otimes X_3)\beta_3 + \varepsilon \\ &\qquad + (I_2 \otimes X_3)(\beta_3 - \beta_2) + \varepsilon \\ &= (I_2 \otimes X_1)(\beta_1 - \beta_2) + (I_2 \otimes X)\beta_2 + (I_2 \otimes X_3)(\beta_3 - \beta_2) + \varepsilon \\ &= (I_2 \otimes X_1)\delta_1 + (I_2 \otimes X)\beta_2 + (I_2 \otimes X_3)\delta_3 + \varepsilon,\end{aligned}$$

and specifying a noninformative constant prior for β_2 and normal priors for δ_i, $\delta_i \sim \mathcal{N}(0, \sigma^2_{\delta_i} I_{2d})$, $i = 1, 3$. The empirical Bayes strategy amounts to replacing Σ, $\sigma^2_{\delta_1}$, and $\sigma^2_{\delta_3}$ by their maximum likelihood estimates $\widetilde{\Sigma}$, $\tilde{\sigma}^2_{\delta_1}$, and $\tilde{\sigma}^2_{\delta_3}$. As illustrated in the appendix, this yields a log posterior density

$$(7)\qquad \mathrm{P}_{rB}(\psi|\Delta p, X) \propto -\frac{1}{2}\left\{ \log|\widetilde{V}||Z'\widetilde{V}^{-1}Z| + (\Delta p - Z\tilde{\beta}_2)'\widetilde{V}^{-1}(\Delta p - Z\tilde{\beta}_2) \right\}$$

with $\tilde{\beta}_2 = (Z'\widetilde{V}^{-1}Z)^{-1}Z'\widetilde{V}^{-1}\Delta p$ and $\widetilde{V} = \widetilde{\Sigma} + \tilde{\sigma}^2_{\delta_1} Z_1 Z_1' + \tilde{\sigma}^2_{\delta_3} Z_3 Z_3'$ for $Z = I_2 \otimes X$, $Z_1 = I_2 \otimes X_1$ and $Z_3 = I_2 \otimes X_3$. A comparison of the (logged) profile likelihood function $\ell_p(\psi)$ in equation (5) with $\mathrm{P}_{rB}(\psi|\Delta p, X)$ in equation (7) shows that unlike the former, the latter does not depend on $\hat{\beta}_k$, $k = 1, 2, 3$, which are not well-defined unless ψ leaves a minimum of d observations in each regime. Accordingly, $\mathrm{P}_{rB}(\psi|\Delta p, X)$ is defined on the entire threshold parameter space $\Psi = \{(\psi_1, \psi_2)|\min(\gamma' p_t) < \psi_1 < \psi_2 < \max(\gamma' p_t)\}$.

The regularized Bayesian threshold estimator $\hat{\psi}_{rB} = \left(\hat{\psi}_{1rB}, \hat{\psi}_{2rB}\right)$ is computed as the posterior median

$$(8) \qquad \int_{\min(\gamma' p_t)}^{\hat{\psi}_{irB}} \mathrm{P}_{rB}(\psi_i|\Delta p, X) d\psi_i = 0.5, \quad i = 1, 2$$

assuming a prior $\mathrm{P}_{rB}(\psi|X) \propto I(\psi \in \Psi)$ for ψ. Here, $\mathrm{P}_{rB}(\psi_i|\Delta p, X)$ denotes the i-th threshold's marginal posterior density. We choose the median of the posterior distribution because it is more robust than the mode and yields more reliable results than the mean when this density is skewed (which tends to be the case when the true threshold is close to the boundary of the threshold parameter space).

Computational issues

Any two threshold values that produce the same allocation of observations into regimes produce identical values of the profile likelihood function $\ell_p(\psi)$. Hence, $\ell_p(\psi)$ is a step function and not differentiable. The same holds for the posterior density $\mathrm{P}_{rB}(\psi|\Delta p, X)$. However, searching a grid that includes all of the observed error correction terms yields the exact maximum of $\ell_p(\psi)$ and also makes it possible to calculate the precise value of the integral of $\mathrm{P}_{rB}(\psi|\Delta p, X)$.

Obviously, a complete grid can be computationally intensive in large samples. Hence, in practice, profile likelihood functions are often evaluated on a coarser grid. For example, some authors (e.g. Goodwin and Piggott 2001) employ evenly spaced grids that divide the threshold parameter space Ψ into a chosen number of equal steps and that therefore do not necessarily include each of the observed error correction terms. In the absence of local maxima and large jumps between individual steps, such a simplified grid will provide a reasonable approximation of the maximum/integral. However, when the dimension of β_k is high or the thresholds leave few observations in one of the regimes, $\ell_p(\psi)$ and $\mathrm{P}_{rB}(\psi|\Delta p, X)$ tend to be jagged and display several local maxima. In such a case, even a fairly dense grid can produce a poor approximation of the true maximum and, consequently, poor estimates, if it does not include all function values. We demonstrate this effect of an inappropriate grid choice in one of the empirical applications below.

Computation of the rB estimator is greatly simplified by taking advantage of functions for mixed models available in statistical software packages. Again, we refer to Greb et al. (2011) for details. R code for calculating rB estimates (for the general TVECM in equation (1) and for restricted models such as the BAND-TVECM) is available from the authors.

Simulations

In a simulation study, we generate data using model (1) with the following parameters: thresholds are set to $\psi_1 = -4$ and $\psi_2 = 4$; adjustment coefficients $\rho_1 = \rho_3 = (-0.25, 0)'$ and $\rho_2 = (0, 0)'$; intercepts $\theta_1 = (-1, 0)'$, $\theta_2 = (0, 0)'$, $\theta_3 = (1, 0)'$; and $\Theta_{11} = \Theta_{31} = \left(\begin{smallmatrix} 0.2 & 0.2 \\ 0 & 0 \end{smallmatrix}\right)$, $\Theta_{21} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. The cointegrating vector $\gamma = (1, -1)'$ is assumed to be known; this implies an error correction term $\gamma' p_t = p_{1,t} - p_{2,t}$ that is simply equal to the difference between p_1 and p_2. Errors are normally distributed, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2 I_2)$ with $\sigma^2 = 1$. The length of the simulated series is $n = 200$. We have selected the parameters to take on values that are plausible in real data applications. In most simulations with these parameters about one half of the data belongs to the inner and one fourth to each of the outer regimes.

We estimate thresholds by applying the profile likelihood and rB estimators to a Monte Carlo sample of 300 replications of the data generating process defined above. We show profile likelihood estimates for three different trimming parameters. These are, first, the least restrictive trimming parameter possible ($d = 2M + 2$, which ensures that each regime contains at least exactly the minimum number of observations necessary to estimate all model parameters), second, 15%, and third, 20% of the sample size. Results are summarized in figures 1 and 2 together with table 1. The rB estimator clearly outperforms the profile likelihood estimator. We observe a considerable reduction in both bias and variance and, consequently, mean squared error. In contrast to the profile likelihood estimates, the rB estimates are not drawn towards zero. The histograms show that the distribution of the rB estimates is also less skewed. Further simulations (including restricted models) confirm these findings. Altogether, the results indicate that the rB estimator is not only superior for generalized threshold regression models, but also for TVECMs specifically.

Empirical Application 1: Goodwin and Piggott (2001) revisited

In the first empirical application, we revisit Goodwin and Piggott's (2001) seminal analysis of spatial price transmission with TVECMs. We apply the rB estimator to their dataset and compare the results with their profile likelihood estimates. Goodwin and Piggott (2001) explore daily corn and soybean prices at important North Carolina terminal markets (figures 3 and 4). These are Williamston, Candor, Cofield, and Kinston for corn, and Fayetteville, Raleigh, Greenville, and Kinston for soybeans. Observations range from 2 January 1992 until 4 March 1999. For each commodity, Goodwin and Piggott (2001) evaluate pairs consisting of a central market – Williamston for corn and Fayetteville for soybeans – and each of the other markets in turn. They estimate the TVECM in equation (1) with logarithmic prices by maximizing the (logged) profile likelihood function $\ell_p(\psi)$ under the assumption of Gaussian errors $\varepsilon_t \sim \mathcal{N}(0, \sigma^2 I_2)$ (or, equivalently, minimizing the sum of squared errors). In accordance with spatial equilibrium theory they assume that $\psi_1 \leq 0$ and $\psi_2 \geq 0$ and search for the maximum of $\ell_p(\psi)$ among those ψ that meet this condition. To obtain comparable results, we also incorporate this information in the rB estimator; we specify a prior on ψ which is zero for any ψ such that $\psi_1 > 0$ or $\psi_2 < 0$, and uniform otherwise. Goodwin and Piggott (2001) evaluate the estimating function at 100 equally spaced grid points for each threshold. In contrast, we compute the rB estimates exactly, that is, the posterior density is evaluated on a grid that is complete (i.e. that includes all observed values of the error correction term). Goodwin and Piggott (2001) assume a trimming parameter that ensures that each regime contains at least 25 observations. We also impose this restriction when replicating their results. As explained above, the rB estimator does not require a trimming parameter.

In table 2, along with the rB and Goodwin and Piggott's (2001) original profile likelihood estimates of the threshold parameters ψ_1 and ψ_2, we also present estimates of ρ_{k1} and ρ_{k2} as well as the total adjustment $\rho_{k1} - \rho_{k2}$ for each regime k ($k = 1, 2, 3$). To interpret these results, note that we employ the conventional specification of the bivariate TVECM in equation (1) in which the error correction term is normalized on the first of the two prices. Thus, assuming that the cointegrating vector $\gamma = (1, -1)'$ and ignoring the lagged difference terms $\sum_{m=1}^{M} \Theta_{km} \Delta p_{t-m}$ in (1), each regime k of the bivariate TVECM takes the following equation-by-equation form:

$$(9) \qquad \Delta p_{1,t} = \rho_{k1}\left(p_{1,t-1} - p_{2,t-1}\right)$$

$$\Delta p_{2,t} = \rho_{k2}\left(p_{1,t-1} - p_{2,t-1}\right)$$

For this specification, the stability condition $|1-(\rho_{k2}-\rho_{k1})|<1$, which is equivalent to $0<\rho_{k2}-\rho_{k1}<2$, ensures that deviations from the long-run equilibrium $p_{1,t}-p_{2,t}=0$ are corrected (Zivot and Wang 2003). This condition, which must hold in the outer regimes $k=1$ and $k=3$ of the TVECM in (1), allows for a wide range of error correction behavior. For example, the pair of adjustment parameters $(\rho_{k1},\rho_{k2})=(-3,-2)$ satisfies this condition. Given these parameters, if $p_{1,t}-p_{2,t}=\eta$ (i.e. p_1 is too large relative to p_2 by the amount η), $\rho_{k1}=-3$ will cause p_1 to fall by 3η in period $t+1$, and $\rho_{k2}=-2$ will cause p_2 to fall by 2η in the same period. Together, these adjustments will restore $p_{1,t}-p_{2,t}=0$.

However, when prices deviate from equilibrium in the context of spatial arbitrage, trade restores equilibrium by causing the higher price to fall and the lower price to rise. Hence, it is reasonable to expect that $\rho_{k1}\leq 0$ and $\rho_{k2}\geq 0$, which precludes combinations such as $(\rho_{k1},\rho_{k2})=(-3,-2)$.[1] Furthermore, combinations that satisfy $1<\rho_{k2}-\rho_{k1}<2$ (for example $\rho_{k1}=-1.3$ and $\rho_{k2}=0.5$) imply exponentially declining oscillations toward equilibrium, which is difficult to reconcile with rational spatial arbitrage. Hence, we also expect that the more restrictive condition $0<\rho_{k2}-\rho_{k1}<1$ will hold in regimes $k=1$ and $k=3$. A pair of adjustment parameters that satisfies these conditions is, for example, $(\rho_{k1},\rho_{k2})=(-0.15,0.1)$, according to which p_1 will fall (rise) in each period to correct 0.15 or 15% of any positive (negative) deviation from the equilibrium condition $p_{1,t}-p_{2,t}=0$, and p_2

[1] It is not necessary that both prices adjust to restore equilibrium. In other words, in regimes 1 and 3 one (but not both) of the adjustment parameters can equal zero. This can occur if, for example, one of the markets being analyzed is so much larger than the other that its price does not react to trade flows between the two.

will correct 0.1 or 10% by moving in the respective opposite direction. Together these price changes imply a total adjustment of $\rho_{k2} - \rho_{k1} = 0.25$ or 25% per period, and thus a smooth exponential error correction process with a half-life of roughly 2.4 periods.[2]

In table 2, we see that compared with the profile likelihood estimates, the rB estimates for both thresholds are always of greater magnitude. This is confirmed by the results reported in the last three columns of the same table, which show (in square brackets) for each pair of markets the number of observations assigned to each of the three regimes by the respective estimation method. Since the thresholds estimated by the regularized Bayesian method are farther from zero, this method assigns correspondingly less (more) observations to the outer (inner) regimes.

In the last three columns of table 2 we also illustrate the effect of using a complete rather than a uniform grid on the allocation of observations into regimes. For the profile likelihood results, the first number in square brackets is the number of observations allocated to the respective regime when Goodwin and Piggott's uniform grid is employed, and the second number is the corresponding number of observations when a complete grid is employed. If both grids lead to similar estimates of the thresholds ψ_1 and ψ_2, then they will also lead to similar allocations of observations into regimes. While this is the case for some market pairs, several cases (for example Kinston – Fayetteville) illustrate that a complete grid is necessary to ensure correct identification of the global maximum of the likelihood function.

What are the economic implications of these results? Several points can be made. First, the fact that the regularized Bayesian threshold estimates are further

[2] In regime 2 (the "band of inaction" between the two thresholds) deviations from the long-run price equilibrium do not trigger any response and both adjustment parameters ρ_{21} and ρ_{22} are expected to equal zero.

apart can be interpreted as evidence of greater market integration. It implies that more observations are in the inner "band of inaction", and correspondingly fewer are in the outer bands where spatial equilibrium is violated, triggering trade and price adjustments. However, if thresholds are estimates of the transaction costs of trade between two locations, then the rB estimates suggest that these costs are higher than indicated by the profile likelihood estimates (see O'Connell and Wei 2002). Hence, the rB threshold estimates suggest that the markets in question are more integrated in the sense that they display fewer violations of spatial equilibrium, but also that they are separated by higher transactions costs which must be overcome before arbitrage becomes profitable.

Second, market integration is reflected not only in how often violations of spatial equilibrium occur, but also in the speed with which such violations are corrected. According to the two-market spatial equilibrium theory discussed above, the outer regimes 1 and 3 should be characterized by more rapid error correction than the inner regime 2, within which prices can move independently and no error correction is expected. The rB estimates of the adjustment parameters fulfill this expectation in all of the six cases in table 2. The only slight exception in the case of Greenville – Fayetteville, in which total adjustment is as large in regime 1 as it is in regime 2 (0.048 in both cases). In comparison, the profile likelihood estimates are compatible with two-market spatial equilibrium theory in only two of the six cases in table 2 (Candor – Williamston and Cofield – Williamston). In the other four cases the profile likelihood estimates display a number of important inconsistencies. In the cases of Kinston – Williamston and Kinston – Fayetteville, for example, total adjustment in regime 2 is considerably stronger than in regimes 1 and 3. And in the cases of Raleigh – Fayetteville and Greenville – Fayetteville total adjustment is roughly twice as strong in regime 2 as it is in regime 3.

Turning to the magnitudes of the estimated adjustment parameters we see that the rB results are not only more consistent with spatial equilibrium theory, they also indicate more rapid correction of disequilibrium in regimes 1 and 3 and, thus, stronger market integration. With the exception of regime 1 in the case of Greenville – Fayetteville, total adjustment in the outer regimes is always stronger according to the rB estimates, and often considerably so, than it is according to the profile likelihood estimates. Specifically, using the rB estimator, the largest total adjustment (0.471) is found in regime 3 for the case of Raleigh – Fayetteville, and total adjustment in the outer regimes amounts to 0.3 or more in four of the six cases in table 2. While these adjustment effects might appear relatively small, they are much larger than the largest total adjustment estimated by profile likelihood (0.132 in regime 1 for Raleigh - Fayetteville). Furthermore, since the underlying price data are daily, a total adjustment of 0.3 corresponds to an adjustment half-life of just under two days, which is evidence of quite rapid reaction to arbitrage opportunities.

With two exceptions, all of the statistically significant estimated adjustment parameters in table 2 have the expected signs. One exception is found in regime 2 for the case of Kinston – Williamston, where the adjustment parameter corresponding to the first equation (for price changes in Kinston) is positive rather than negative. This result, which holds for both the rB and the profile likelihood estimates, is compensated in both cases by a slightly larger and correctly-signed adjustment parameter in the second equation (for price changes in Williamston), so that the total adjustment effect is positive and small. The same happens for the profile likelihood estimates in the case of Cofield – Williamston. Overall, relatively few estimated adjustment parameters are statistically significant, and many of the larger rB estimates of adjustment parameters in regimes 1 and 3 are not significant. This is presumably due to the small numbers of observations in these regimes in most cases.

One other aspect of the results in table 2 deserves mention. For two of the market pairs (Candor – Williamston and especially Greenville – Fayetteville) the rB estimates of the adjustment parameters are comparatively small and similar across regimes. In the case of Greenville – Fayetteville, for example, the total adjustments in regimes 1, 2 and 3 are 0.048, 0.048 and 0.078, respectively, compared with, for example 0.302, 0.060 and 0.298 in the case of Kinston – Fayetteville. These results might indicate that the two-threshold, three-regime model of price transmission is misspecified. Sephton (2003), who also revisits the Goodwin and Piggott (2001) data, finds that the pairs Raleigh-Fayetteville and Greenville-Fayetteville display little evidence of threshold effects. Our rB estimates of very similar or identical adjustment coefficients across regimes appear to corroborate Sephton's finding for Greenville – Fayetteville.

Empirical Application 2: Price transmission between German and Spanish pork prices

As a second empirical application, we analyze transmission between German and Spanish pork prices. The analysis is carried out using the data presented in figure 5, which are average weekly prices of grade E pig carcasses for Germany and Spain in Euro per 100 kg between May 21, 1989 and October 17, 2010 (1091 observations). We specify a TVECM with three regimes,

$$(10) \quad \Delta p_t = \begin{cases} \rho_1 \gamma' p_{t-1} + \theta_1 + \sum_{m=1}^{M} \Theta_{1m} \Delta p_{t-m} + \varepsilon_t \quad , \quad \gamma' p_{t-1} \leq \psi_1 \quad \text{(Regime 1)} \\ \rho_2 \gamma' p_{t-1} + \theta_2 + \sum_{m=1}^{M} \Theta_{2m} \Delta p_{t-m} + \varepsilon_t \quad , \quad \psi_1 < \gamma' p_{t-1} \leq \psi_2 \quad \text{(Regime 2)} \\ \rho_3 \gamma' p_{t-1} + \theta_3 + \sum_{m=1}^{M} \Theta_{3m} \Delta p_{t-m} + \varepsilon_t \quad , \quad \psi_2 < \gamma' p_{t-1} \quad \text{(Regime 3),} \end{cases}$$

with $\Delta p_t = \left(\Delta p_t^{Germany}, \Delta p_t^{Spain}\right)'$ and $M = 3$. We apply profile likelihood and the rB estimator with the error correction term $\gamma' p_{t-1}$ defined as the difference between the Spanish and the German prices, $\gamma' p_t = p_t^{Germany} - p_t^{Spain}$. Since the sample period spans more than twenty years and includes events such as the introduction of the Euro, the postulate of constant transaction costs (ψ_1 and ψ_2) over time is likely to be an oversimplification. It is beyond the scope of this article to model variable transaction costs, but this should be kept in mind when interpreting the results.

We plot the profile likelihood for the upper threshold (ψ_2) in figure 6. To generate this figure, the lower threshold (ψ_1) is fixed at its profile likelihood estimate. We see that the profile likelihood reaches its maximum at the boundary of the range defined by the smallest possible trimming parameter (i.e. the requirement that each regime contains at least one observation per parameter to be estimated). Hence, any more restrictive trimming parameter (such as requiring that each regime contains at least 2.5 or 5% of all observations) strongly influences the profile likelihood estimate (see figure 6), rendering it arbitrary and unreliable. Compared with an estimate $\hat{\psi}_2 = 26.1$ for the least restrictive trimming parameter, requiring 2.5% (5%) of the observations to fall into each regime produces the estimate $\hat{\psi}_2 = 21.8$ ($\hat{\psi}_2 = 14.0$).

The rB estimator does not require an arbitrary trimming parameter. It produces threshold estimates $(-37.8, 34.8)$ that are considerably larger in magnitude than the profile likelihood estimates $(-27.9, 26.1)$. As mentioned above with respect to the analysis of the Goodwin and Piggott (2001) data, larger estimated thresholds define a wider "band of inaction" in regime 2 that can be interpreted as evidence of poorer market integration. However, in the case of the German and

Spanish pork prices the profile likelihood thresholds estimates are smaller because they are restricted by the trimming parameter. Hence, they reflect biased estimation rather than lower transaction costs of trade and greater market integration.

Furthermore, the rB estimator produces estimates of the adjustment parameters that are more plausible than their profile likelihood counterparts (table 3). In regime 1, where the difference between the German and Spanish prices is less than the lower threshold value, the profile likelihood estimate of the adjustment parameter for the Spanish price is large and significant (-0.665), but has an implausible sign. Both magnitude and sign are implausible for the corresponding parameter estimate in regime 3 (-1.193), where the difference between the German and the Spanish prices exceeds the upper threshold. The corresponding estimated adjustment parameters for the German price in regimes 1 and 3 (-0.198 and -0.334) have the expected negative signs, but they are insignificant. Altogether, the total adjustments for regimes 1 and 3 are negative according to the profile likelihood method (see the third-to-last and last colums of table 3). Hence, the profile likelihood estimates suggest that there is no mechanism that returns German and Spanish prices to their long run equilibrium when shocks drive them apart.

In comparison, the rB estimates of the adjustment parameters make considerably more sense. All of the rB estimates that are significant have the expected sign, and together they indicate that when the difference between the German and the Spanish prices exceeds one of the thresholds, adjustments are triggered that return these prices to their long run equilibrium (total adjustment equals 0.318 in regime 1 and 0.347 in regime 3).

In summary, the empirical applications illustrate the advantages of the rB estimator in the context of spatial price transmission analysis. The rB estimator does not depend on a trimming parameter that arbitrarily influences the profile likelihood results in the application with Spanish and German pork prices. Furthermore, in both applications the rB estimates of the adjustment parameters are more consistent with spatial equilibrium theory and price transmission between the markets in question than the corresponding profile likelihood estimates.

Conclusions

We discuss the estimation of TVCEMs in spatial price transmission analysis. We point out shortcomings of the common (profile likelihood) estimation procedure and emphasize the relevance of these problems for applied price transmission studies. As an alternative, we suggest employing a regularized Bayesian estimator (Greb et al. 2011), and we demonstrate this estimator's superior performance in a simulation study. Revisiting the empirical analysis in Goodwin and Piggott's influential paper on TVECMs in price transmission analysis, we find that the regularized Bayesian estimates are free of several inconsistencies that characterise the profile likelihood estimates. A second application, with German and Spanish pork prices, confirms the advantages of the regularized Bayesian estimator in spatial price transmission modeling, producing results that are more consistent with the theory of spatial equilibrium than the corresponding profile likelihood results.

Future work could move beyond the pairwise consideration of markets to study multivariate sets of prices and the more complex multiple-threshold relationships that exist between them. Another extension would be to investigate time-varying thresholds, since especially for longer time-series the assumption of constant transaction costs is questionable.

Acknowledgments

We acknowledge the support of the German Research Foundation (Deutsche Forschungsgemeinschaft) as part of the Institutional Strategy of the University of Göttingen. We would like to thank Barry Goodwin and Nicholas Piggott for sharing their dataset with us.

Appendix

We aim to compute the posterior density $\mathrm{P}_{rB}\left(\psi|\Delta p, X\right)$ for the model

$$\Delta p = (I_2 \otimes X_1)\delta_1 + (I_2 \otimes X)\beta_2 + (I_2 \otimes X_3)\delta_3 + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

with a normal prior $\delta_1 \sim \mathcal{N}(0, \sigma_{\delta_1}^2 I_{2d})$, where $d = 2M + 2$ with M the number of lags included in the model; a uniform prior $\beta_2 \sim U\left(\mathbb{R}^{2d}\right)$; a normal prior $\delta_3 \sim \mathcal{N}(0, \sigma_{\delta_3}^2 I_{2d})$; and a uniform prior $\psi \sim U(\psi \in \Psi)$.

To this end, we first calculate $\mathrm{P}_{rB}\left(\Delta p|\psi, X\right)$, since

$$\mathrm{P}_{rB}\left(\psi|\Delta p, X\right) = \mathrm{P}_{rB}\left(\Delta p|\psi, X\right) \mathrm{P}_{rB}(\psi|X) \Big/ \mathrm{P}_{rB}(\Delta p|X) \propto \mathrm{P}_{rB}\left(\Delta p|\psi, X\right)$$

given a constant prior $\mathrm{P}_{rB}(\psi|X)$. Employing an empirical Bayes approach, it suffices to compute $\mathrm{P}_{rB}\left(\Delta p|\psi, X, \Sigma, \sigma_{\delta_1}^2, \sigma_{\delta_3}^2\right)$: parameters Σ, $\sigma_{\delta_1}^2$, and $\sigma_{\delta_3}^2$ are replaced by their maximum likelihood estimates $\widetilde{\Sigma}$, $\tilde{\sigma}_{\delta_1}^2$, and $\tilde{\sigma}_{\delta_3}^2$. Given our specification of priors,

$$\begin{aligned} \mathrm{P}_{rB}\left(\Delta p|\psi, X, \Sigma, \sigma_{\delta_1}^2, \sigma_{\delta_3}^2\right) &= \int \mathrm{P}_{rB}\left(\Delta p, \beta_2|\psi, X, \Sigma, \sigma_{\delta_1}^2, \sigma_{\delta_3}^2\right) d\beta_2 \\ &= \int \mathrm{P}_{rB}\left(\Delta p|\beta_2, \psi, X, \Sigma, \sigma_{\delta_1}^2, \sigma_{\delta_3}^2\right) \mathrm{P}_{rB}\left(\beta_2|\psi, X, \Sigma, \sigma_{\delta_1}^2, \sigma_{\delta_3}^2\right) d\beta_2 \\ &= \int \mathrm{P}_{rB}\left(\Delta p|\beta_2, \psi, X, \Sigma, \sigma_{\delta_1}^2, \sigma_{\delta_3}^2\right) d\beta_2 \end{aligned}$$

and

$$\Delta p|\beta_2, \psi, X, \Sigma, \sigma^2_{\delta_1}, \sigma^2_{\delta_3} \sim \mathcal{N}\left\{(I_2 \otimes X)\beta_2, \Sigma + \sigma^2_{\delta_1}(I_2 \otimes X_1)(I_2 \otimes X_1)' + \sigma^2_{\delta_3}(I_2 \otimes X_3)(I_2 \otimes X_3)'\right\}.$$

To simplify notation, define $Z = I_2 \otimes X$, $Z_1 = I_2 \otimes X_1$, $Z_3 = I_2 \otimes X_3$, and $V = \Sigma + \sigma^2_{\delta_1} Z_1 Z_1' + \sigma^2_{\delta_3} Z_3 Z_3'$ and write

$$\Delta p|\beta_2, \psi, X, \Sigma, \sigma^2_{\delta_1}, \sigma^2_{\delta_3} \sim \mathcal{N}\left(Z\beta_2, V\right).$$

Consequently,

$$\begin{aligned}
&\mathrm{P}_{rB}\left(\Delta p|\psi, X, \Sigma, \sigma^2_{\delta_1}, \sigma^2_{\delta_3}\right) \\
&= \int \left(\frac{1}{2\pi}\right)^{2n/2} \frac{1}{\sqrt{|V|}} \exp\left\{-\frac{1}{2}(\Delta p - Z\beta_2)' V^{-1} (\Delta p - Z\beta_2)\right\} d\beta_2 \\
&= \left(\frac{1}{2\pi}\right)^{2n/2} \frac{1}{\sqrt{|V|}} \exp\left\{-\frac{1}{2}(\Delta p - Z\tilde{\beta}_2)' V^{-1} (\Delta p - Z\tilde{\beta}_2)\right\} \cdot \\
&\qquad \int \exp\left\{-\frac{1}{2}(\beta_2 - \tilde{\beta}_2)' Z' V^{-1} Z (\beta_2 - \tilde{\beta}_2)\right\} d\beta_2 \\
&= \left(\frac{1}{2\pi}\right)^{2n/2} \frac{1}{\sqrt{|V|}} \exp\left\{-\frac{1}{2}(\Delta p - Z\tilde{\beta}_2)' V^{-1} (\Delta p - Z\tilde{\beta}_2)\right\} (2\pi)^{2d/2} \frac{1}{\sqrt{|Z'V^{-1}Z|}} \\
&= \left(\frac{1}{2\pi}\right)^{2(n-d)/2} \frac{1}{\sqrt{|V||Z'V^{-1}Z|}} \exp\left\{-\frac{1}{2}(\Delta p - Z\tilde{\beta}_2)' V^{-1} (\Delta p - Z\tilde{\beta}_2)\right\}
\end{aligned}$$

with $\tilde{\beta}_2 = (Z'V^{-1}Z)^{-1} Z'V^{-1}\Delta p$. Substituting $\widetilde{\Sigma}$, $\tilde{\sigma}^2_{\delta_1}$, and $\tilde{\sigma}^2_{\delta_3}$ for Σ, $\sigma^2_{\delta_1}$, and $\sigma^2_{\delta_3}$, respectively, that is, $\widetilde{V} = \widetilde{\Sigma} + \tilde{\sigma}^2_{\delta_1} Z_1 Z_1' + \tilde{\sigma}^2_{\delta_3} Z_3 Z_3'$ for V, yields a log posterior density

$$\mathrm{P}_{rB}(\psi|\Delta p, X) \propto \mathrm{P}_{rB}\left(\Delta p|\psi, X\right) \propto -\frac{1}{2}\left\{\log|\widetilde{V}||Z'\widetilde{V}^{-1}Z| + (\Delta p - Z\tilde{\beta}_2)'\widetilde{V}^{-1}(\Delta p - Z\tilde{\beta}_2)\right\}.$$

References

Andrews, D. W. K. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61(4): 821–56.

Balcombe, K., A. Bailey, and J. Brooks. 2007. Threshold effects in price transmission: the case of Brazilian wheat, maize, and soya prices. *American Journal of Agricultural Economics* 89(2): 308–23.

Balcombe, K., and G. Rapsomanikis. 2008. Bayesian estimation and selection of nonlinear vector error correction models: the case of the sugar-ethanol-oil nexus in Brazil. *American Journal of Agricultural Economics* 90(3): 658–68.

Balke, N. S., and T. B. Fomby. 1997. Threshold cointegration. *International Economic Review* 38(3): 627–45.

Barndorff-Nielsen, O. 1983. On a formula for the distribution of the maximum likelihood estimator. *Biometrika* 70(2): 343–365.

Chen, C. 1998. A Bayesian analysis of generalized threshold autoregressive models. *Statistics & probability letters* 40(1): 15–22.

Fackler, P. L., and B. K. Goodwin. 2001. Spatial price analysis. *Handbook of agricultural economics* 1: 971–1024.

Goodwin, B. K., and N. E. Piggott. 2001. Spatial market integration in the presence of threshold effects. *American Journal of Agricultural Economics* 83(2): 302–17.

Greb, F., T. Krivobokova, A. Munk, and S. von Cramon-Taubadel. 2011. *Regularized Bayesian estimation in generalized threshold regression models*. Courant Research Centre "Poverty, Equity and Growth in Developing Countries", Georg-August-Universität Göttingen. CRC-PEG Discussion Paper No.99.

Hansen, B. E., and B. Seo. 2002. Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics* 110(2): 293–18.

Ihle, R. 2010. *Models for Analyzing Nonlinearities in Price Transmission*: Niedersächsische Staats-und Universitätsbibliothek Göttingen.

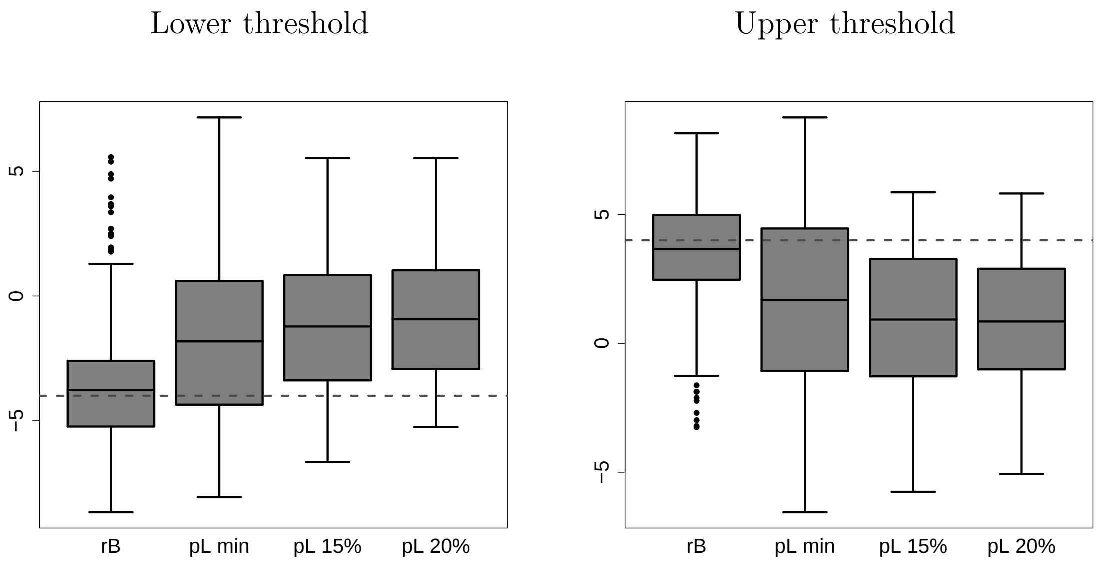
Lo, M. C., and E. Zivot. 2001. Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics* 5(4): 533–76.

O'Connell, P., and S. Wei. 2002. "The bigger they are, the harder they fall": Retail price differences across US cities. *Journal of International Economics* 56(1): 21–53.

Sephton, P. 2003. Spatial market arbitrage and threshold cointegration. *American Journal of Agricultural Economics* 85(4): 1041–1046.

Stephens, E. C., E. Mabaya, S. Cramon-Taubadel, and C. B. Barrett. 2011. Spatial Price Adjustment with and without Trade. *Oxford Bulletin of Economics and Statistics*. Published online July 5, 2011.

Zivot, E., and J. Wang. 2003. *Modeling financial time series with S-PLUS*. New York: Springer Verlag.



Note: The horizontal dashed gray line indicates the true threshold. The dark lines in the shaded boxes are the respective sample means. "rB" denotes the regularized Bayesian estimates. "pL min", "pL 15%", and "pL 20%" denote profile likelihood estimates with trimming parameters equal to the smallest possible value ($d = 2M + 2$), 15% of the sample size, and 20% of the sample size, respectively.

Figure 1. Simulation results – boxplots

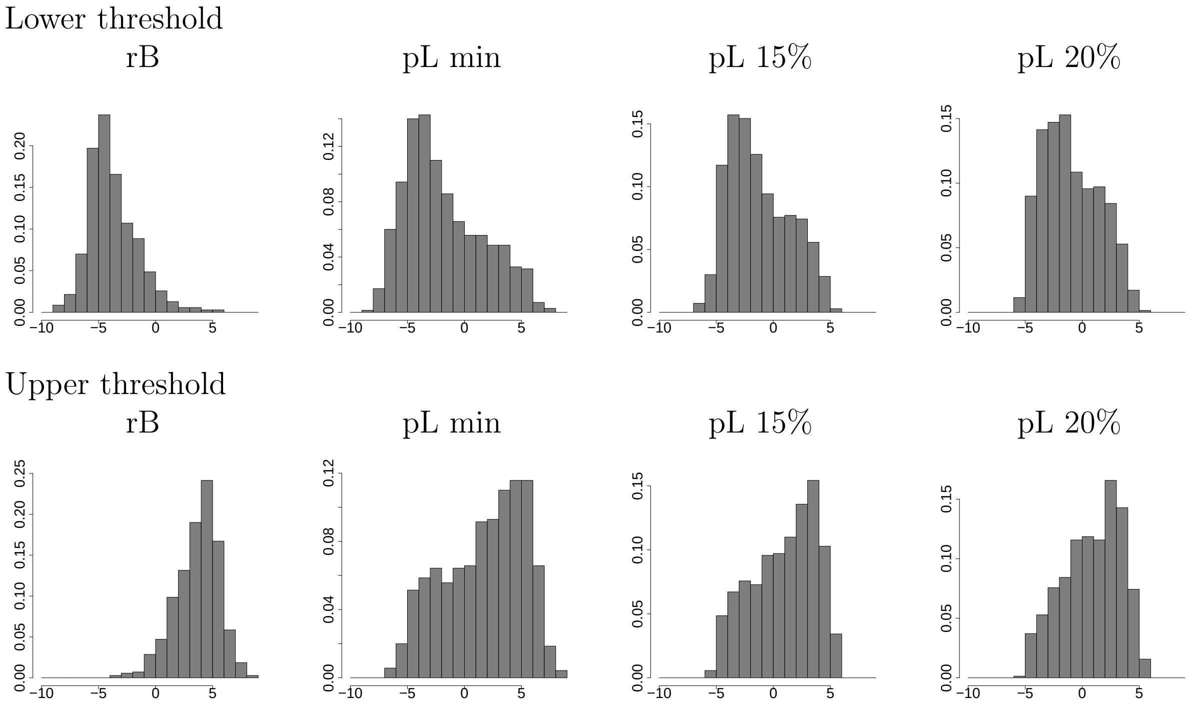


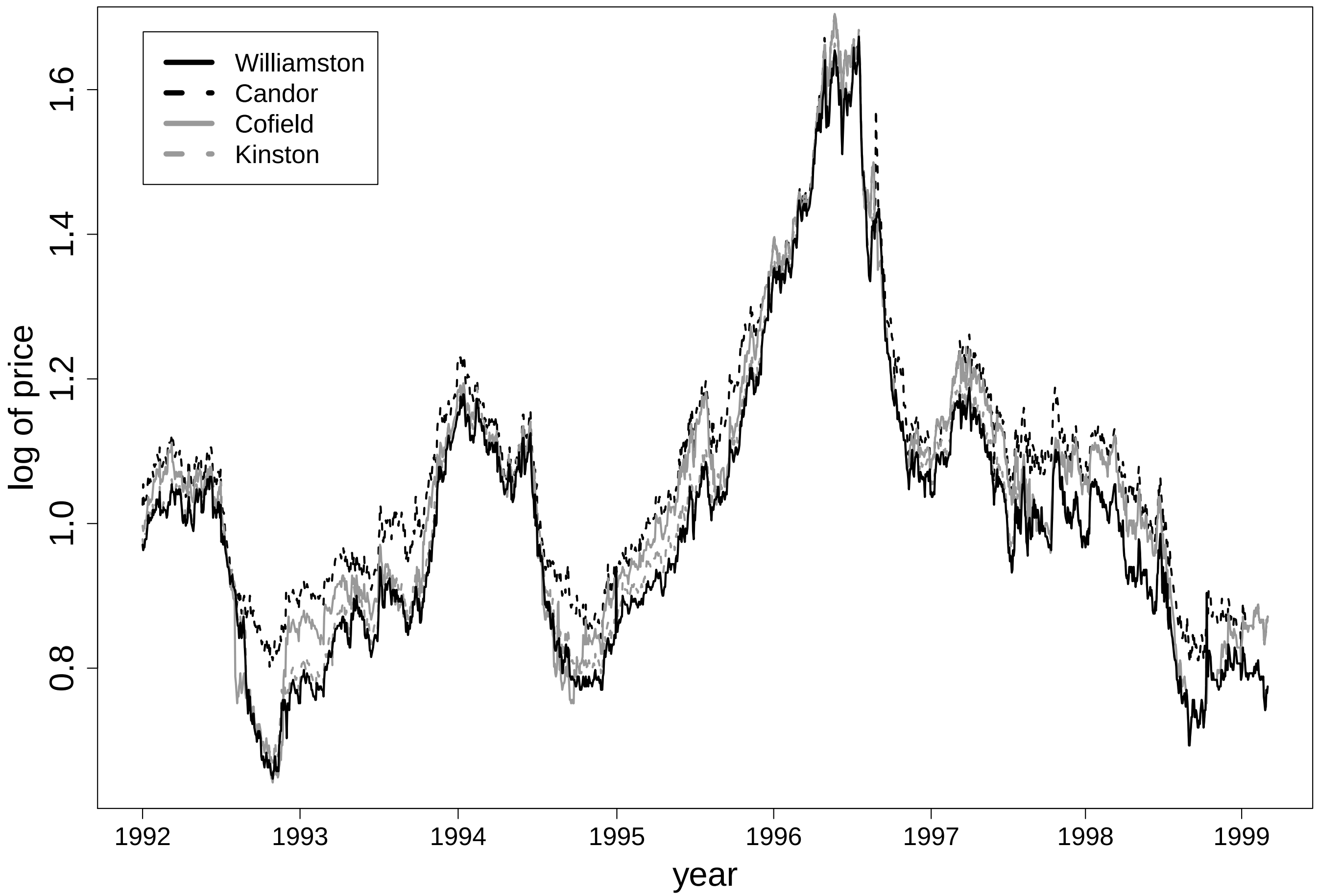
Figure 2. Simulation results – histograms

Note: "rB" denotes the regularized Bayesian estimates. "pL min", "pL 15%", and "pL 20%" denote profile likelihood estimates with trimming parameters equal to the smallest possible value ($d = 2M + 2$), 15% of the sample size, and 20% of the sample size, respectively.

Table 1. Simulation Results

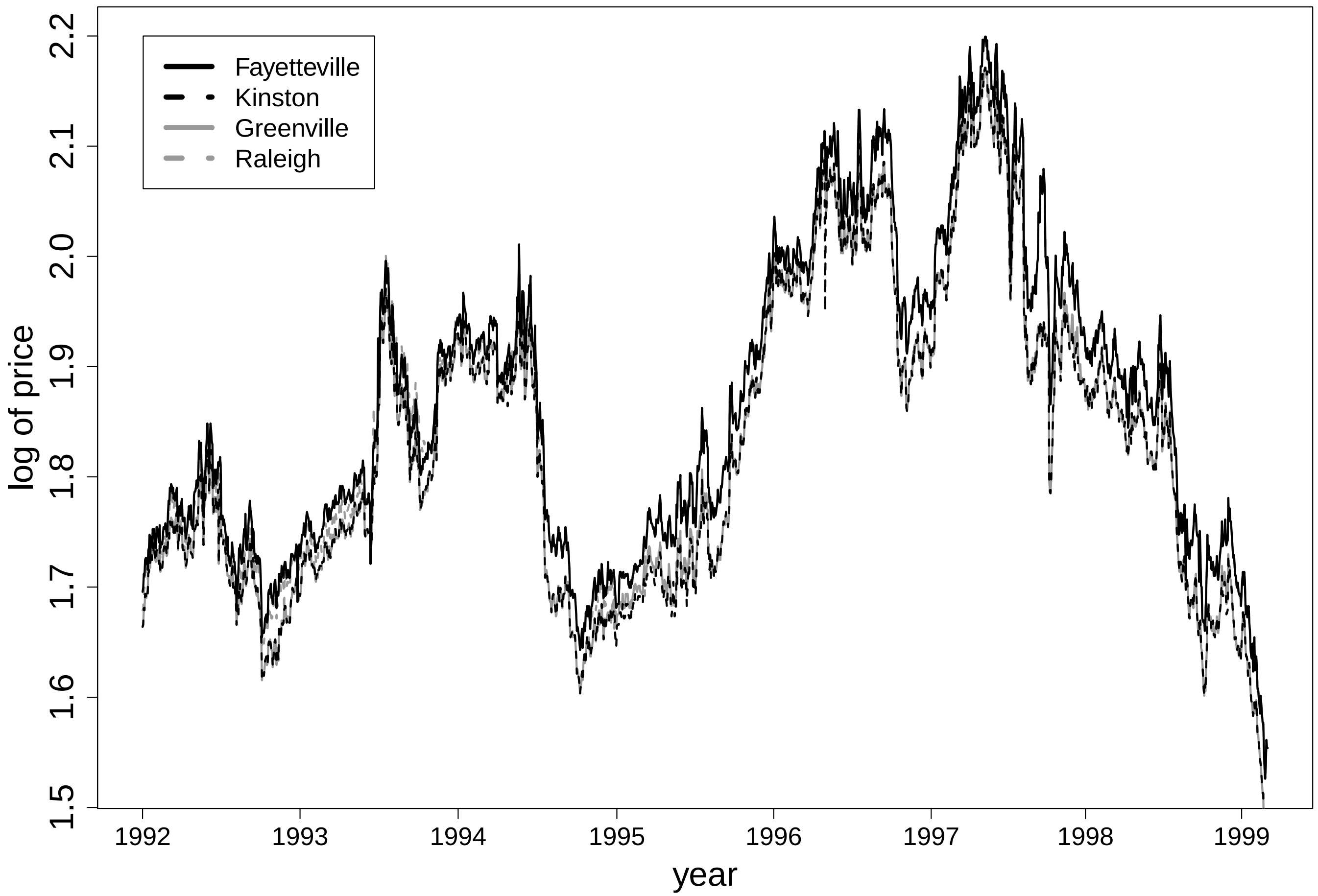
	Regularized Bayesian estimator		Profile likelihood estimator					
	lower threshold	upper threshold	lower threshold			upper threshold		
			min	15%	20 %	min	15%	20 %
true	-4	4	-4	-4	-4	4	4	4
mean	-3.76	3.66	-1.82	-1.22	-0.93	1.69	0.92	0.85
	(2.16)	(1.90)	(3.40)	(2.67)	(2.45)	(3.50)	(2.79)	(2.52)
MSE	4.71	3.73	16.31	14.86	15.40	17.57	17.21	16.24

Note: Standard errors are reported in parentheses below the mean. "min", "15%", and "20%" denote trimming parameters equal to the smallest possible value ($d = 2M + 2$), 15% of the sample size, and 20% of the sample size, respectively.



Source: Goodwin and Piggott (2001), who kindly made these data available.

Figure 3. Logged daily corn prices at four North Carolina terminal markets



Source: Goodwin and Piggott (2001), who kindly made these data available.

Figure 4. Logged daily soybean prices at four North Carolina terminal markets

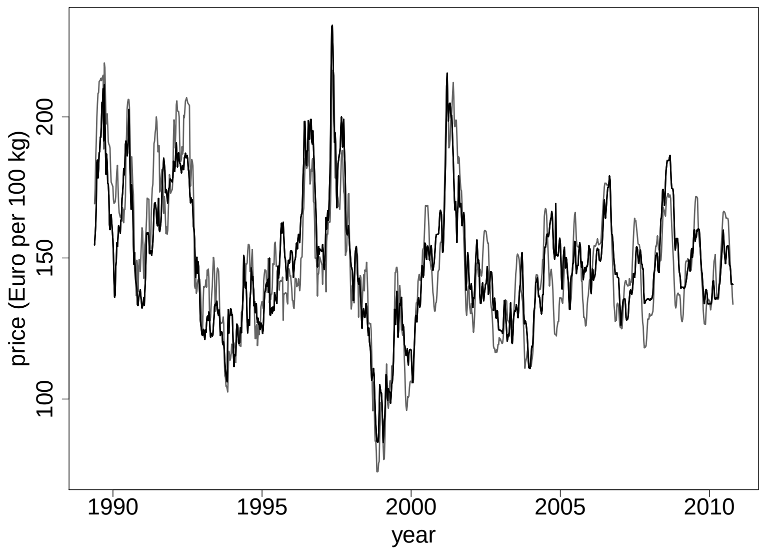
Table 2. Estimates for the Data in Figures 3 and 4 – TVECM with three Regimes

Est.	Dep.var.	ρ_1	$\sigma\,(\rho_1)$	ψ_1	ρ_2	$\sigma\,(\rho_2)$	ψ_2	ρ_3	$\sigma\,(\rho_3)$	Total(ρ_1) [#obs.]	Total(ρ_2) [#obs.]	Total(ρ_3) [#obs.]
Corn: Candor-Williamston												
pL	Δp^{CAN}	0.030	(0.020)	−0.0255	−0.003	(0.039)	0.0073	−0.008	(0.018)	0.047	0.009	0.057
	Δp^{WIL}	**0.077**	(0.020)		0.006	(0.039)		**0.049**	(0.018)	[295/299]	[762/676]	[716/798]
rB	Δp^{CAN}	0.064	(0.055)	−0.0799	0.003	(0.013)	0.0677	0.064	(0.055)	0.097	0.04	0.097
	Δp^{WIL}	**0.162**	(0.055)	(0.016)	**0.043**	(0.013)	(0.020)	**0.162**	(0.055)	[11]	[1760]	[2]
Corn: Cofield-Williamston												
pL	Δp^{COF}	**−0.056**	(0.021)	−0.0572	**0.027**	(0.012)	0.0594	0.017	(0.044)	0.118	0.007	0.076
	Δp^{WIL}	**0.062**	(0.020)		**0.034**	(0.012)		**0.094**	(0.043)	[69/69]	[1669/1698]	[35/6]
rB	Δp^{COF}	−0.144	(0.170)	−0.1908	0.011	(0.010)	0.0688	**−0.267**	(0.143)	0.363	0.033	0.305
	Δp^{WIL}	0.220	(0.170)	(0.032)	**0.043**	(0.010)	(0.019)	0.037	(0.142)	[1]	[1769]	[3]
Corn: Kinston-Williamston												
pL	Δp^{KIN}	0.064	(0.068)	−0.0125	**0.156**	(0.052)	0.0178	0.061	(0.107)	0.005	0.028	0.001
	Δp^{WIL}	0.070	(0.068)		**0.184**	(0.052)		0.062	(0.108)	[249/198]	[1469/1568]	[55/7]
rB	Δp^{KIN}	−0.179	(0.346)	−0.0293	**0.107**	(0.038)	0.0192	−0.180	(0.357)	0.456	0.023	0.456
	Δp^{WIL}	0.276	(0.347)	(0.009)	**0.130**	(0.038)	(0.009)	0.276	(0.357)	[6]	[1762]	[5]
Soybeans: Raleigh-Fayetteville												
pL	Δp^{RAL}	−0.126	(0.115)	−0.006	−0.098	(0.106)	0.0103	−0.039	(0.134)	0.132	0.009	0.004
	Δp^{FAY}	0.006	(0.116)		−0.089	(0.107)		−0.035	(0.136)	[166/499]	[1560/1227]	[47/47]

Est.	Dep.var.	ρ_1	$\sigma\,(\rho_1)$	ψ_1	ρ_2	$\sigma\,(\rho_2)$	ψ_2	ρ_3	$\sigma\,(\rho_3)$	Total(ρ_1) [#obs.]	Total(ρ_2) [#obs.]	Total(ρ_3) [#obs.]
rB	Δp^{RAL}	-0.200	(0.161)	-0.0353	-0.081	(0.063)	0.021	-0.219	(0.280)	0.371	0.093	0.471
	Δp^{FAY}	0.171	(0.161)	(0.009)	0.012	(0.063)	(0.004)	0.252	(0.280)	[5]	[1764]	[4]
Soybeans: Greenville-Fayetteville												
pL	Δp^{GRE}	-0.012	(0.028)	-0.0102	0.028	(0.039)	0.0216	0.055	(0.079)	0.058	0.04	0.022
	Δp^{FAY}	0.046	(0.028)		**0.068**	(0.039)		0.077	(0.080)	[411/438]	[1292/1030]	[70/305]
rB	Δp^{GRE}	0.012	(0.021)	-0.1011	0.012	(0.021)	0.0251	0.017	(0.097)	0.048	0.048	0.078
	Δp^{FAY}	**0.060**	(0.022)	(0.024)	**0.060**	(0.022)	(0.005)	0.095	(0.098)	[2]	[1760]	[11]
Soybeans: Kinston-Fayetteville												
pL	Δp^{KIN}	-0.012	(0.027)	-0.006	0.023	(0.183)	0.007	-0.026	(0.036)	0.061	0.125	0.084
	Δp^{FAY}	**0.050**	(0.026)		0.148	(0.180)		**0.058**	(0.035)	[544/6]	[508/1755]	[721/12]
rB	Δp^{KIN}	-0.231	(0.207)	-0.1201	-0.005	(0.021)	0.0265	-0.112	(0.326)	0.302	0.06	0.298
	Δp^{FAY}	0.071	(0.207)	(0.012)	**0.055**	(0.021)	(0.003)	0.186	(0.324)	[1]	[1765]	[7]

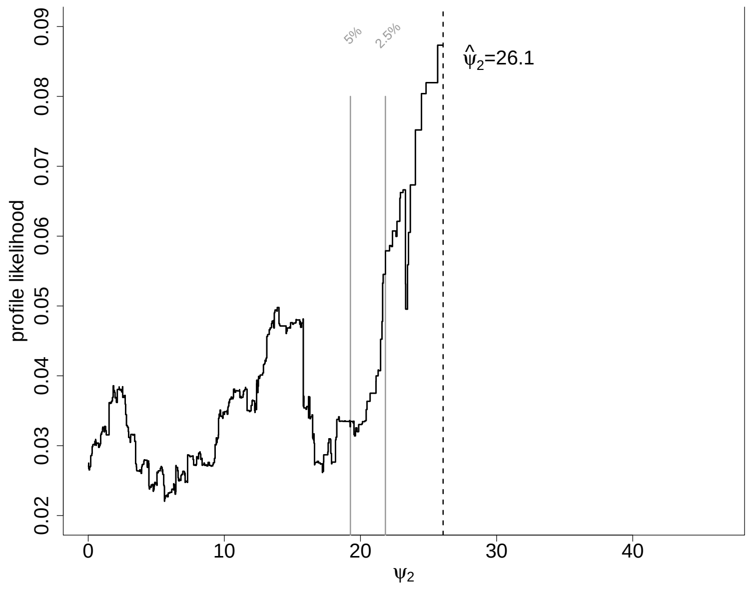
Notes:

- pL is the profile likelihood estimator; rB is the regularized Bayesian estimator.
- pL estimates are computed as in Goodwin and Piggott with a trimming parameter that ensures that each regime contains at least 25 observations.
- Standard errors of the estimated adjustment parameters (ρ_k) are provided in brackets. These must be interpreted with care because they are computed without accounting for the variability of the threshold estimate. Estimates that are significant at the 10% level are in **bold**. Standard errors for rB threshold estimates (in brackets below the estimate) are calculated in the customary Bayesian manner as their posterior standard deviation. To the best of our knowledge, how to compute standard errors for pL threshold estimates in TVECMs remains an open question.
- The number in square brackets below Total(ρ_k) is the estimated number of observations in regime k. For pL, the first number corresponds to Goodwin and Piggott's estimates, the second to pL estimates based on a complete grid.



Source: European Commission: http://ec.europa.eu/agriculture/markets/pig/porcs.pdf.

Figure 5. Weekly prices for grade E pig carcasses in Germany and Spain (Euro per 100 kg)



The dashed vertical line indicates the profile likelihood estimate for the upper threshold, $\hat{\psi}_2$, estimated using the least restrictive possible trimming parameter. Solid grey lines indicate how the threshold parameter space is restricted when 2.5% (5%) of the observations are required to fall into each regime. The lower threshold is fixed at its profile likelihood estimate, $\hat{\psi}_1 = -27.9$.

Figure 6. Profile likelihood function for the upper threshold, ψ_2, estimated with the pig price data in figure 5.

Table 3. Estimates for the Data in Figure 5 – TVECM with three Regimes

Est.	Dep.var.	ρ_1	$\sigma\,(\rho_1)$	ψ_1	ρ_2	$\sigma\,(\rho_2)$	ψ_2	ρ_3	$\sigma\,(\rho_3)$	Total(ρ_1) [#obs.]	Total(ρ_2) [#obs.]	Total(ρ_3) [#obs.]
pL	$\Delta p^{Germany}$	-0.198	(0.342)	-27.9	$\mathbf{-0.028}$	(0.012)	26.1	-0.334	(1.448)	-0.467	0.08	-0.859
	Δp^{Spain}	$\mathbf{-0.665}$	(0.365)		**0.052**	(0.012)		-1.193	(1.545)	[20]	[1059]	[8]
rB	$\Delta p^{Germany}$	$\mathbf{-0.288}$	(0.103)	-37.8	$\mathbf{-0.028}$	(0.011)	34.8	$\mathbf{-0.355}$	(0.115)	0.318	0.092	0.347
	Δp^{Spain}	0.030	(0.106)	(7.4)	**0.063**	(0.012)	(8.6)	-0.008	(0.117)	[1]	[1085]	[1]

Note: The notes below table 2 apply.