



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

# NCCC-134

APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

## **A Comparison of Threshold Cointegration and Markov-Switching Vector Error Correction Models in Price Transmission Analysis**

by

Rico Ihle and Stephan von Cramon-Taubadel

Suggested citation format:

Ihle, R., and S. v. Cramon-Taubadel. 2008. "A Comparison of Threshold Cointegration and Markov-Switching Vector Error Correction Models in Price Transmission Analysis." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [<http://www.farmdoc.uiuc.edu/nccc134>].

**A Comparison of Threshold Cointegration and Markov-Switching Vector Error Correction Models in Price Transmission Analysis**

*Rico Ihle*

and

*Stephan von Cramon-Taubadel\**

*Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis,  
Forecasting, and Market Risk Management  
St. Louis, Missouri, April 21-22, 2008*

*Copyright 2008 by Rico Ihle and Stephan von Cramon-Taubadel.  
All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.*

---

\* Rico Ihle (rihle@gwdg.de) is a PhD student in the Centre for Statistics ZfS and the Department of Agricultural Economics and Rural Development at the Georg-August-Universität Göttingen, Germany; Stephan von Cramon-Taubadel is Professor in the Department of Agricultural Economics and Rural Development at the Georg-August-Universität Göttingen, Germany.

## A Comparison of Threshold Cointegration and Markov-Switching Vector Error Correction Models in Price Transmission Analysis

*We compare two regime-dependent econometric models for price transmission analysis, namely the threshold vector error correction model and Markov-switching vector error correction model. We first provide a detailed characterization of each of the models which is followed by a comprehensive comparison. We find that the assumptions regarding the nature of their regime-switching mechanisms are fundamentally different so that each model is suitable for a certain type of nonlinear price transmission. Furthermore, we conduct a Monte Carlo experiment in order to study the performance of the estimation techniques of both models for simulated data. We find that both models are adequate for studying price transmission since their characteristics match the underlying economic theory and allow hence for an easy interpretation. Nevertheless, the results of the corresponding estimation techniques do not reproduce the true parameters and are not robust against nuisance parameters. The comparison is supplemented by a review of empirical studies in price transmission analysis in which mostly the threshold vector error correction model is applied.*

**Keywords:** price transmission; market integration; threshold vector error correction model; Markov-switching vector error correction model; comparison; nonlinear time series analysis

### 1 Introduction

Economists have devoted considerable attention to testing the Law of One Price (LOP) in a variety of settings, and agricultural economists in particular have generated an extensive literature on the empirical analysis of price transmission (PT) along spatial (prices for a homogeneous commodity at different locations - e.g. wheat in France and Germany) and vertical (prices for a commodity at different stages of processing - e.g. wheat-flour-bread) dimensions. Early studies focused on correlations or linear time series analysis involving prices, but in recent years attention has increasingly turned to the use of models that can capture the regime-dependent nature of relationships between prices. In a spatial context, the key insight, derived from Takayama and Judge (1971), is that prices will only co-move if spatial arbitrage conditions are binding (Baulch, 1994). If the difference between prices at two locations is greater than the cost of trade between these locations, then arbitrage will drive the price difference net of transaction costs to zero, and this equilibrating mechanism will lead to an observable relationship between the prices in question. If the difference between these prices is less than the transaction costs, however, arbitrage will not take place and in the simplest case the prices will move independently of one another. The result is a two-regime model of PT that extends to three regimes if the possibility of trade reversal is considered, and possibly more regimes if factors such as links to third markets or equilibrating mechanisms other than physical trade are accounted for.

The threshold vector error correction model has been used extensively in PT analysis (Goodwin and Piggott, 2001, etc.). Recently, Brümmer et al. (2008) propose the use of the

Markov-switching vector error correction model to study price transmission in a vertical context between wheat and flour in Ukraine. So far no systematic attempt has been made to contrast and compare these models as regards their theoretical underpinnings and their performance and interpretation in the context of PT. In this paper we carry out such a comparison in order to provide some indication regarding the common and differing features of both models. Both models allow for regime-switching; does this characteristic imply that they may be used interchangeably and lead to congruent results? We show that this is not the case and that each approach suits best particular analytical objectives in PT analysis. The comparison discusses the most important aspects for the empirical application of both models in PT analysis in detail in order to give some indication for the application of both models.

Section 2 outlines the relationship of both models to other time series models by introducing the class of nonlinear time series models in general and the subclass of threshold autoregressive models. These considerations are followed by a detailed characterization of the threshold and the Markov-switching vector error correction model, respectively, by focusing on the basic idea, the model structure, the estimation and the interpretation of each. Section 3 provides a conceptual comparison of the characteristics of both models outlined before. It is supplemented by a simulation study which assesses the performance of the estimation methods of each model. The last section summarizes and draws conclusions. Appendix A provides a literature review of applications of the threshold vector error correction model to PT analysis. Appendix B contains details on the simulation study.

## 2 Model Review

### 2.1 Classification of nonlinear time series models

Many model classes for nonlinear time series analysis were developed during the second half of the seventies and the eighties of the past century.<sup>1</sup> Tong (1978) introduced the class of so-called *threshold models*. Fairly general formulations of nonlinear models have been developed by Priestley (1980b) (the class of *state-dependent models*) and Tjøstheim (1986) (the class of *doubly stochastic models*) which encompass a wide range of classes of less general models, among others threshold models. Tong (1990) suggests a comprehensive classification of model classes for nonlinear time series analysis (Figure 1). Model classes characterized by a specific functional relationship which do not contain other subclasses are called *elementary* model classes. On the next higher level, groups of such elementary classes, which are called *first-generation* models, can be identified according to common properties. In turn, first generation models can be generalized in various ways. The resulting meta-classes such as the above-mentioned state-dependent and doubly stochastic models are called *second-generation* models, which are very general in their specification and each of which includes various first-generation models.<sup>2</sup>

Among the first-generation models, a wide variety of model classes has been developed. Classes such as bilinear (BL) models, threshold autoregressive (TAR) models<sup>3</sup> or

---

<sup>1</sup> For a narrative about the “Birth of the Threshold Time Series Model” see Tong (2007).

<sup>2</sup> However, Tong (1990), among others, questions their usefulness for practical analysis because of the high degree of generalization.

<sup>3</sup> TAR models are called *nonlinear mean reversion* (NMR) models in real exchange rate analysis; compare, for

autoregressive models with conditional heteroscedasticity (ARCH) are examples. For the purpose of this paper, the class of TAR models is most interesting. It subdivides into the three groups of piecewise polynomial, piecewise linear and smooth autoregressive models depending on the functional relationship  $f$  between the history  $\{X_p\}_{p \in \mathbb{Z}, p < t}$  of the time series<sup>4</sup>  $\{X_t\}_{t \in \mathbb{Z}}$  and its value  $X_t$  at time  $t$ :

$$X_t = f(\underbrace{X_{t-1}, \dots, X_{t-k}}_{\text{history of the time series}}, \epsilon_{t-1}, \dots, \epsilon_{t-k}) + \epsilon_t. \quad (1)$$

A general formulation of the threshold model might take the form

$$\mathbf{X}_t = \mathbf{A}^{(J_t)} \mathbf{X}_{t-1} + \mathbf{H}^{(J_t)} \epsilon_t + \mathbf{C}^{(J_t)} \quad (2)$$

where  $\mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-k+1})^\top$  and  $J_t$  denotes a random variable which takes one of the integer values  $\{1, 2, \dots, l\}$  at each time  $t$ .  $J_t$  is an indicator variable signaling the state (regime) in which the series  $\{X_t\}$  is at time  $t$ . For a particular state  $J_t = j$ , the  $(k \times k)$  non-random matrices  $\mathbf{A}^{(j)}$  and  $\mathbf{H}^{(j)}$  contain the autoregressive coefficients and the coefficients that reflect heteroscedasticity, respectively. The  $(k \times 1)$  vector  $\mathbf{C}^{(j)}$  comprises the constants of the relationship.  $\{\epsilon_t\}$  denotes a sequence of identically and independently distributed (*iid*)  $k$ -dimensional random vectors with zero mean and existing covariance matrix. Thus, for each state  $J_t = j$  the relationship is locally linear<sup>5</sup> with a particular set of coefficients and/or a constant.

The determination of  $J_t$  remains unspecified in (2). One might think of various ways in which the states of  $\{X_t\}$  are determined. This indicator variable is the key element of the nonlinear character of the equation; as Tong (1983) puts it, “ $J_t$  indicates the mode of the dynamic mechanism”. The realizations of  $J_t$  at all time points  $t$  form the series of the states (regimes)  $\{J_t\}$  which is referred to as the *regime-generating* process (RGP) of the time series. This generation mechanism of the regime process characterizes elementary model classes within the class of piecewise linear TAR models. The state of a threshold model can be generated by one of the following basic mechanisms:

$$J_t = f(X_{t-p}) \quad (3a)$$

$$J_t = f(Y_{t-q}) \quad (3b)$$

$$J_t = f(X_{t-p}, Y_{t-q}) \quad (3c)$$

where  $t, p \in \mathbb{N}^+$ ,  $q \in \mathbb{N}$  and  $t > p, t > q$ .

The first case refers to the endogenous determination of the regimes of  $\{X_t\}$  by some part of its history. Tong (1990) calls this the class of *self-exciting threshold autoregression* models (SETAR)<sup>6</sup> since the regimes of  $\{X_t\}$  are completely generated by the series itself. The

---

example, Norman (2007) and O’Connell and Wei (2002). However, we will stick to the former term throughout this paper.

<sup>4</sup> We use the abbreviated form  $\{X_t\}$  for denoting a time series in this paper.

<sup>5</sup> As Priestley (1980a) notes, the term *local* refers in this setting not to the proximity to a particular point in time but to a certain region of the state space of the series. Furthermore, *linear* refers to the constancy of parameters in such a region. *Local linearity* is thus the key property of TAR models, namely that their parameters are not constant over the whole range of observations, but take (constant) values depending on the current state/ the regime of the time series. Hence they are only constant within each state and called *state-dependent* or *regime-dependent* parameters.

<sup>6</sup> The abbreviation may be complemented by the number of regimes and the lag length as

second case denotes the exogenous determination by some other series  $\{Y_t\}$  lagged by  $q$  periods that is independent of  $\{X_t\}$ . One can think of a number of ways of exogenous determination. The most obvious generation mechanism is a second time series  $\{Y_t\}$  which is known. Tong (1990) refers to this case as an *open-loop threshold autoregressive system* (TARSO). If  $\{Y_t\}$  itself follows a threshold model and its regimes are exogenously determined by  $\{X_t\}$ , the model is called a *closed-loop threshold autoregressive system* (TARSC), i.e., each of the two series is determining the states of the other one. Another possibility, among others, is the determination of the states by a set of unknown (exogenous) variables which cannot be identified or measured for some reason so that only quantify conditional probabilities of staying in a state or switching to another can be quantified. Thus, the states of a series  $\{X_t\}$  might be generated by a Markov chain. The resulting model is called Markov-switching autoregressive (MSAR)<sup>7</sup> which can easily be transformed into the *Markov-switching vector error correction model* (MSVECM). The third type of regime determination can be thought of as a mixture of the two above-mentioned ones in which the states of  $\{X_t\}$  are determined by a combination of lagged values of the series itself and of some exogenous series  $\{Y_t\}$ . The case that the states of the second series  $\{Y_t\}$  are in turn determined by a combination some lag of itself and of  $\{X_t\}$  can be called *simultaneous TARSC*. If regressands of such a system are not expressed in levels but in differences, the resulting piecewise linear TAR model with mixed regime determination is called a *threshold vector error correction model* (TVECM). Hence, the TVECM and the MSVECM both belong to the class of piecewise linear TAR models.

## 2.2 Detailed Characterization

### *The Threshold Principle*

In a simple market setting it is often postulated that quantity demanded will equal zero above a certain price, or that quantity supplied will equal zero below a certain price. As a result, the functional relationship between quantity (supplied or demanded) and price will be subject to different regimes depending on whether the price is above or below certain values. Such values are called *thresholds*. A threshold introduces nonlinearities into the functional relationship and “specifies the operation modes of the system” (Tong, 1990). The relationship between two or more variables might be locally linear<sup>8</sup>, however, globally it exhibits nonlinear behavior because of the existence of one or more structural changes in the relationship.

Tong (1990) notes that “threshold is a generic concept” resulting from the general property of saturation<sup>9</sup>, i.e., the structural changes as found, for example, in the mentioned quantity-price relationship. Tong defines the *threshold principle* as “the local approximation over the states, i.e., the introduction of regimes via thresholds”. Such regime-dependent parameter stability of some time series is usually referred to as *threshold behavior*. Balke and Fomby (1997, and

---

SETAR( $l; k_1, k_2, \dots, k_l$ ) where  $l$  denotes the number of regimes and  $k_j, j = 1, 2, \dots, l$  the lag-length in the  $j^{th}$  state as, for example, in Tong (1990), or only by the number of regimes SETAR( $l$ ) as, for example, in Hansen (1999).

<sup>7</sup> A multidimensional version of this class are the Markov-switching vector autoregressive (MSVAR) models which are discussed in depth in Krolzig (1997).

<sup>8</sup> Compare footnote 5 on page 4 for a definition of local linearity.

<sup>9</sup> Tong (1990) and Tong and Lim (1980) provide a large number of examples in various disciplines of science.

references therein) list several examples in which threshold behavior is found in economics, e.g. prices, inventories, consumer durables or employment.

## THRESHOLD VECTOR ERROR CORRECTION MODELS

### *Basic Idea*

Although Whittle (1954) is first to suggest a statistical model based on the threshold idea, the class of threshold models is formally introduced by Tong in 1978. He and many other researchers subsequently extend this area of research. Bhansali suggests that, as early as 1980, “commodity price series [are] a possible class of economic time series where applications of these models may be useful”. In the second half of the 1980s, cointegration theory is developed to deal with the analysis of non-stationary time series.<sup>10</sup> In 1997, Balke and Fomby publish a paper on threshold cointegration in which they unite both developments. Their essential insight is the assumption that the correction of deviations from the long-run equilibrium, i.e., the equilibrium errors, might display threshold behavior. The TVECM has attracted much attention in, among others, PT analysis since the publication of Balke and Fomby (1997).<sup>11</sup>

The possible existence of nonlinear PT was first hypothesized by Heckscher (1916).<sup>12</sup> In the context of international trade, he proposed a *band of inaction* in which small deviations from the equilibrium price are not adjusted because transaction costs are higher than potential earnings due to the price differential. These transaction costs not only encompass transport costs, but for example also costs of searching, negotiating, insurance and risk premia. Heckscher termed the boundaries of this neutral band in which prices are supposed to move freely, *commodity points*. In other words, the of transmission of price signals between markets depends on whether deviations from the equilibrium price are inside the band of inaction or not, i.e., PT changes structurally depending on the magnitude of the deviations. Hence, PT is likely to follow threshold behavior. Such a regime-dependent nature of PT also results from the Enke-Samuleson-Takayama-Judge spatial equilibrium model formulated in Takayama and Judge (1971). The model implies that trade will only occur if the price spread of some homogeneous commodity between two spatially separated markets is at least as large as the transaction costs of trading between these two markets. Consequently, PT depends on the magnitude of the price spread, i.e., it shows regime-dependent behavior.

Figure 2 depicts the threshold behavior of PT.<sup>13</sup> It shows the quantity traded  $trade_t^{AB}$  from market  $A$  to market  $B$  as a function of the price differential  $p_t^B - p_t^A$  between  $B$  and  $A$ .  $\tau$  denotes the price differential above which trade takes place. Rational traders will only engage in trade if it is profitable, i.e., when they make a net profit. Thus,  $\tau$  can be interpreted as the commodity point for trade from  $A$  to  $B$  which is equivalent to the transaction costs involved in the trading process. Price differentials below  $\tau$  will not trigger trade flows and are

<sup>10</sup> Compare, for example, Engle and Granger (1987), Johansen (1995), Hendry and Juselius (2000) or Hendry and Juselius (2001).

<sup>11</sup> We provide a review of publications which study PT in commodity trade using mainly the TVECM in the econometric analysis in Appendix A, pp. 36.

<sup>12</sup> This idea is based on the LOP as it was formulated by Marshall (1890, p. 325) who also mentioned the role of transaction costs.

<sup>13</sup> A cointegration vector  $\beta = (1 \ -1)^\top$  is implicitly assumed here.



not adjusted. However, if the price differential is greater than  $\tau$ , trade, by shifting supply from A to B, will cause  $p_t^A$  to rise and  $p_t^B$  to fall. This mechanism reduces the price differential in a process that will continue until it returns to  $\tau$ .  $p_t^B - p_t^A - \tau = 0$  is therefore an equilibrium relationship. If both  $p_t^A$  and  $p_t^B$  are  $I(1)$ , it will be a cointegrating relationship, with an equilibrium error  $p_t^B - p_t^A - \tau$  that is corrected by trade whenever it exceeds zero<sup>14</sup>; values of the error that are less than zero are not corrected<sup>15</sup>. Hence, trade leads to the many times studied price adjustment process. Consequently, the magnitude of PT will differ depending on whether trade takes place or not, that is, PT shows regime-dependent behavior. Thus, threshold models are both theoretically and intuitively appropriate in general for the analysis of PT. Moreover, the regressands are usually expressed in first differences, i.e.,  $\Delta p_t^A = p_t^A - p_{t-1}^A$  and  $\Delta p_t^B = p_t^B - p_{t-1}^B$ . The regimes of each price series, directly corresponding to the regimes of PT, are determined by the error correction term, which is itself a function of both series. Thus, a simultaneous TARSC in the form of the TVECM is an appropriate model. Obstfeld and Taylor (1997) provide the first publication which explicitly refers to the hypothesis of Heckscher. O'Connell and Wei (1997) and Trenkler and Wolf (2003) derive this idea from economic theory. Several theoretical models in the area of real exchange rate analysis yield results in line with Heckscher's hypothesis; see, for example, Dumas (1992), Uppal (1993), Sercu et al. (1995), Coleman (1995, 2004).

### Model Structure

The TVECM may generally be formulated as follows<sup>16</sup>:

$$\Delta \mathbf{p}_t = \boldsymbol{\mu}^{(j)} + \boldsymbol{\alpha}^{(j)} \boldsymbol{\beta}^\top \mathbf{p}_{t-1} + \mathbf{C}^{(j)}(L) \Delta \mathbf{p}_t + \boldsymbol{\epsilon}_t \quad \text{if } \theta^{(j-1)} < \boldsymbol{\beta}^\top \mathbf{p}_{t-d} \leq \theta^{(j)} \quad (4)$$

$$= \boldsymbol{\mu}^{(j)} + \boldsymbol{\alpha}^{(j)} ect_{t-d} + \sum_{i=1}^{k-1} \boldsymbol{\Psi}_i^{(j)} \Delta \mathbf{p}_{t-i} + \boldsymbol{\epsilon}_t \quad \text{if } \theta^{(j-1)} < ect_{t-d} \leq \theta^{(j)} \quad (5)$$

$$= \boldsymbol{\mu}^{(J_t)} + \boldsymbol{\alpha}^{(J_t)} ect_{t-d} + \sum_{i=1}^{k-1} \boldsymbol{\Psi}_i^{(J_t)} \Delta \mathbf{p}_{t-i} + \boldsymbol{\epsilon}_t \quad (6)$$

where  $\mathbf{p}_t = (p_t^A \ p_t^B)^\top$  is the vector of prices in markets A and B,  $t = 1, \dots, T$  denotes the time index and  $j \in \{1, 2, \dots, l, l+1\}$  the index of the regimes.  $\boldsymbol{\mu}^{(j)}$  denotes the regime-dependent mean where the superscript  $(j)$  signals the regime-dependency of the parameter.  $ect_{t-d} = \boldsymbol{\beta}^\top \mathbf{p}_{t-d}$  denotes the deviation from the long-run equilibrium, i.e. the error correction term lagged by  $d$  periods.<sup>17</sup>  $\boldsymbol{\beta} = (\beta^A \ \beta^B)^\top$  denotes the cointegration vector of the prices  $\mathbf{p}_t$  and  $\boldsymbol{\alpha}^{(j)} = (\alpha^A \ \alpha^B)^\top$  is called the loading vector. It contains the regime-dependent parameters characterizing to what extent the price changes  $\Delta \mathbf{p}_t$  react on deviations from the long-run equilibrium lagged by  $d$  periods. These parameters are

<sup>14</sup> Hence, the price spread  $p_t^B - p_t^A$  is directly proportional to the equilibrium error. Since, for example, the price change in market B  $\Delta p_t^B = p_t^B - p_{t-1}^B$  is a measure for trade from A to B, the error correction mechanism as depicted, for example, in Meyer (2004) corresponds to Figure 2.

<sup>15</sup> However, negative values of the error are bounded from below by a second threshold which measures the transaction costs of trade in the opposite direction. This second threshold need not be of the same magnitude as the first, as, for example, the costs of moving up- as opposed to downriver or with and without backhauls might differ.

<sup>16</sup> For a derivation see, for example, Balke and Fomby (1997) or Lo and Zivot (2001).

<sup>17</sup> Here it becomes obvious, that the threshold variable  $ect_{t-d}$  is a linear combination of the price series  $\mathbf{p}_t$  and thus a function of those.

interpreted as the magnitudes of error-correction of both prices which are equivalent to the speed (the rate) of price adjustment to the long-run equilibrium and characterize the regime-dependent magnitudes of PT.  $\mathbf{C}^{(j)}(L)$  denote lag polynomials of order  $k$  and, alternatively, the  $\Psi_i^{(j)}$  are  $(2 \times 2)$  matrices containing the autoregressive coefficients of each price difference (the coefficients for short-run adjustment of deviations). The errors  $\epsilon_t$  are  $(2 \times 1)$  vectors of *iid* random variables with mean zero and finite covariance matrix  $\Sigma$ .

The values  $\theta^{(j)} \in \mathbb{R}$  are ordered so that  $\theta^{(0)} < \theta^{(1)} < \dots < \theta^{(l)} < \theta^{(l+1)}$  where  $\theta^{(0)} = -\infty$ ,  $\theta^{(l+1)} = \infty$ . They are called *threshold parameters* or in short *thresholds*.<sup>18</sup> We impose the assumption on the thresholds to be time-invariant since this specification is almost exclusively used in applied research.<sup>19</sup> The variable determining the relevant regime at time  $t$  is called *threshold variable*.<sup>20</sup> It is assumed to be stationary and to follow a continuous distribution.  $d \in \mathbb{N}^+$  is called the *delay parameter*. Alternatively, the model can be formulated using the indicator variable  $J_t$  introduced in (2). It takes the value  $j$  at time  $t$  if  $\theta^{(j-1)} < ect_{t-d} \leq \theta^{(j)}$ .

Obviously, the nonlinear TVECM is a generalization of the linear vector error correction model (VECM). Each threshold  $\theta^{(j)}$  is only meaningful if

$$0 < \mathbf{P}(\theta^{(j-1)} < ect_{t-d} \leq \theta^{(j)}) < 1. \quad (7)$$

That is, only if realizations of the threshold variable occur with a probability larger than zero, i.e., are observable in each regime, the respective threshold exists.<sup>21</sup> By introducing dummy variables for each regime, the model can more compactly be formulated in terms of a multivariate regression model similarly to Hansen and Seo (2002):

$$\Delta \mathbf{p}_t = \mathbf{A}^{(1)\top} \mathbf{X}_{t-1} \mathbf{d}_t^{(1)} + \dots + \mathbf{A}^{(l)\top} \mathbf{X}_{t-1} \mathbf{d}_t^{(l)} + \epsilon_t \quad (8)$$

$$= \sum_{j=1}^l \mathbf{A}^{(j)\top} \mathbf{X}_{t-1} \mathbf{d}_t^{(j)} + \epsilon_t \quad (9)$$

$$= \mathbf{A}^{(J_t)\top} \mathbf{X}_{t-1} + \epsilon_t \quad (10)$$

where  $\mathbf{A}^{(j)}$  denotes a  $((2k+2) \times 2)$  matrix of coefficients. The vector of the regressors of (5) with  $(2k+2)$  elements is contained in  $\mathbf{X}_{t-1} = (1 \quad \beta^\top \mathbf{p}_{t-1} \quad \Delta \mathbf{p}_{t-1} \dots \Delta \mathbf{p}_{t-k})^\top$ .

Furthermore,  $\mathbf{d}_t^{(j)} = \mathbb{1}_{(\theta^{(j-1)} < ect_{t-d} \leq \theta^{(j)})}$  denotes the dummy variable signaling the  $j$ 's regime of the series at time  $t$  where  $\mathbb{1}_{(\bullet)}$  is the indicator function. By expressing the regimes of the price series in terms of the indicator variable  $J_t$ , a special case of (2) is obtained.

In the analysis of PT, the thresholds are interpreted as the transaction costs for moving a

<sup>18</sup> The thresholds  $\theta^{(0)}$  and  $\theta^{(l+1)}$  are usually not referred to as thresholds in the proper sense of the term. They rather represent some kind of natural boundaries since the threshold variable of any meaningful model will take values between  $-\infty$  and  $\infty$ . Hence, they exist also for each linear model and are only introduced for the sake of the generality of (4) - (6). In general, if we speak of thresholds we refer only to the inner ones, i.e.,  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(l)}$ . Thus in general, a TVECM of  $s$  regimes has  $s - 1$  effective, i.e., inner thresholds and vice versa. Thus,  $l$  denotes the number of effective thresholds.

<sup>19</sup> For models relaxing this restriction see, e.g., Van Campenhout (2007) who models the threshold as a linear function of time and Park et al. (2007) who derive formulae for dynamic thresholds varying on a daily basis.

<sup>20</sup> In the case of the TVECM the threshold variable is always the deviation from the long-run equilibrium  $ect_{t-d}$ .

<sup>21</sup> If the realizations of the threshold variable are likely to occur only in one regime, no effective threshold exists and the TVECM in (5) simplifies to a linear VECM of the form  $\Delta \mathbf{p}_t = \boldsymbol{\mu} + \boldsymbol{\alpha} ect_{t-d} + \sum_{i=1}^{k-1} \Psi_i \Delta \mathbf{p}_{t-i} + \epsilon_t$ .

homogeneous commodity between any pair of markets which introduce the nonlinear behavior into the PT process. The error-correction mechanism is usually assumed to react immediately one period after some deviation from equilibrium, i.e., the delay parameter is usually assumed to equal one and the error correction term becomes  $ect_{t-1}$ . Furthermore, the number of regimes is restricted, often either set to two or three implying one or two thresholds respectively.<sup>22</sup> In line with the above-mentioned theoretical background, a TVECM(3) has much appeal since it accounts for trade into both directions between two spatially separated markets.<sup>23</sup>

Balke and Fomby (1997) and Lo and Zivot (2001) suggest certain restrictions on the model which might be particularly suitable for applied analysis. The two prices can, based on Heckscher's supposition, be expected not to be cointegrated inside the "band of inaction" spanned by the two transaction costs implying that the price differences  $\Delta \mathbf{p}_t$  move as random walks around zero. Consequently, no error correction takes place in regime  $j = 2$  between the two thresholds, i.e.,  $\alpha^{(2)} = \mathbf{0}$ , and the regime-dependent mean equals zero  $\mu^{(2)} = \mathbf{0}$ . Depending on the center of attraction of the error correction mechanism, special cases of the model can be distinguished. If the errors are corrected toward a band around the long-run equilibrium which is spanned by the regime-specific means  $\mu^{(1)}$  and  $\mu^{(3)}$  the model is called a *BAND-TVECM* as formulated in (11). However, if the errors are corrected toward the long-run equilibrium itself, implying  $\mu^{(1)} = \mu^{(3)} = \mathbf{0}$ , the model is called an *Equilibrium-TVECM* (EQ-TVECM). Moreover, the model is called *continuous* if  $\mu^{(1)} = -\alpha^{(1)}\theta^{(1)}$  and  $\mu^{(3)} = -\alpha^{(3)}\theta^{(2)}$ . If both (effective) thresholds are of the same magnitude, i.e., if  $-\theta^{(1)} = \theta^{(2)}$ , implying identical transaction costs in both directions of trade, the model is called *symmetric*.

$$\Delta \mathbf{p}_t = \begin{cases} \mu^{(1)} + \alpha^{(1)}ect_{t-1} + \sum_{i=1}^{k-1} \Psi_i^{(1)} \Delta \mathbf{p}_{t-i} + \epsilon_t & \text{if } \theta^{(0)} < ect_{t-1} \leq \theta^{(1)} \\ \mu^{(2)} + \alpha^{(2)}ect_{t-1} + \sum_{i=1}^{k-1} \Psi_i^{(2)} \Delta \mathbf{p}_{t-i} + \epsilon_t & \text{if } \theta^{(1)} < ect_{t-1} \leq \theta^{(2)} \\ \mu^{(3)} + \alpha^{(3)}ect_{t-1} + \sum_{i=1}^{k-1} \Psi_i^{(3)} \Delta \mathbf{p}_{t-i} + \epsilon_t & \text{if } \theta^{(2)} < ect_{t-1} \leq \theta^{(3)}. \end{cases} \quad (11)$$

Figure 3 depicts a realization of an EQ-TVECM characterized by three regimes. The regime  $J_t$  depends exclusively on the magnitude of the first lag of the error correction term  $ect_{t-d}$ . The price series  $p_t^A$  and  $p_t^B$  are plotted in the bottom panel. The variable causing regime switches is deviation from the long-run equilibrium, i.e., the error correction term  $ect_t$  which equals the difference between the prices at each time  $t$ . It is separately plotted in the middle panel. If it is either smaller or larger than the lower  $\theta^{(1)}$  or the upper threshold  $\theta^{(2)}$ ,  $J_t$  takes the values  $j = 1$  or  $j = 3$ , respectively, and error correction toward zero takes place. However, prices move independently inside the band spanned by the two thresholds since  $\alpha^{(2)} = \mathbf{0}$ . Whenever the threshold variable  $ect_t$  crosses one of the thresholds, the regime switches after a lag of  $d$  periods to the new regime as depicted in the middle and the upper panel of Figure 3.<sup>24</sup> The parameters of most interest in applied analysis are the thresholds  $\theta^{(j)}$ , the loading vectors  $\alpha^{(j)}$  and the cointegration vector  $\beta$ .

<sup>22</sup> In order to refer to the number of regimes, the name of the specified model is sometimes supplemented by this number, for example a TVECM with  $l$  thresholds has  $l + 1$  regimes and can be denoted by TVECM( $l + 1$ ) or TVECM $_{l+1}$ .

<sup>23</sup> A TVECM(2) where  $\theta^{(1)} = 0$  is suitable for the study of asymmetric PT, see, e.g., Chen et al. (2005).

<sup>24</sup> The rationale for such a lag is that markets need some time to react. Nevertheless, this time may depend on the product traded, the market infrastructure and the socio-economic environment of the market.

## Estimation

Several authors developed estimation techniques for threshold models in general or for the TVECM in particular. Tong (1978) suggests the Entropy Maximization Principle based on the Akaike Information Criterion (AIC) for the estimation of a general TAR model. Tsay (1989), Chan (1993) and Hansen (2000) propose approaches for threshold models with two regimes. Tsay (1998) shows that, asymptotically, the estimates of this sequential conditional multivariate least squares estimation are strongly consistent and that the estimated coefficients  $\mathbf{A}^{(j)}$  in equation (9) are independent of the thresholds  $\theta^{(j)}$  and the delay parameter  $d$  and normally distributed.

Balke and Fomby (1997) suggest conditional least squares estimation for TAR models applicable to any number of thresholds and delay parameters. Obstfeld and Taylor (1997, Appendix A) give a detailed description of their applied maximum likelihood estimation technique. Hansen (1999) presents an estimation technique for SETAR models with two or more regimes based on sequential conditional least squares estimation through concentration. Lo and Zivot (2001, Appendix A) suggest a combination of the methods of Hansen (1999) and Tsay (1998) for estimating one threshold and the delay parameter of a multivariate TVECM. Hansen and Seo (2002) propose a maximum likelihood estimation procedure for the TVECM for the bivariate case, i.e., with two regimes, which allows for the simultaneous estimation of the cointegration vector and the threshold, and provide a detailed description of the algorithm proposed.

Table 5 summarizes estimation approaches of selected publications in chronological order. It displays information on the underlying model such as the model class, the number of estimated thresholds  $l$  and potential restrictions on the delay parameter  $d$ . Moreover, it mentions whether the estimation follows the maximum likelihood or the least squares principle. The latter is referred to differently in the literature as *sequential*, *iterative* or *conditional* (multivariate) least squares.<sup>25</sup> This is complemented by information on the optimization method such as the considered optimization criterion, i.e., the objective function, its parameters and the type of the optimization. The RSS criterion, in contrast to the log-determinant of the variance-covariance matrix, ignores correlations across the regimes' equations. Nevertheless, Serra and Goodwin (2002) have shown that both criteria yield the same estimation results and might thus be considered to be equivalently suitable.

The functions of the presented criteria will usually not be smooth.<sup>26</sup> Hence, a grid search algorithm in form of SCLS is suitable for optimization. Its dimension depends on the number of parameters of the optimization criterion.<sup>27</sup> The challenge for estimation consists in the fact that the unknown parameters of the model depend on each other. The coefficients matrices  $\mathbf{A}^{(j)}$  in (9) additionally to the variance-covariance matrix  $\Sigma$  depend, among others, on the unknown thresholds  $\theta^{(j)}$ , however, the former are a precondition for estimating the latter.

The basic idea of the grid search is very pragmatic. In order to break the “vicious circle”, the parameters of the optimization criterion, i.e., among others the thresholds, are pretended to

---

<sup>25</sup> We refer to the method as sequential conditional least squares (SCLS) throughout the paper.

<sup>26</sup> For examples of the shape of such criterion functions see Hansen and Seo (2002, p. 298).

<sup>27</sup> The higher the dimension the higher the computational costs. Several authors have suggested alternatives, see, for example, Hansen and Seo (2002), Lo and Zivot (2001), Hansen (1999), Bai and Perron (1998), Bai (1997) or Dorsey and Mayer (1995).

be known and are set to some constants. Conditionally on the combination of the chosen optimization parameters, the remaining model parameters, i.e., the coefficients matrices  $\mathbf{A}^{(j)}$  and the variance-covariance matrix  $\Sigma$ , and the optimization criterion are computed. The computation is repeated for a number of combinations of possible values of the optimization parameters, and the criterion is evaluated.<sup>28</sup>

Candidate values of the optimization parameters are generated by an evenly spaced grid across the empirical support of the threshold variable and potentially a reasonable range of the criterion's other parameters. The combination which optimizes the criterion represents the final estimates of the optimization parameters. Conditionally on these, the final estimates of the remaining model parameters are obtained. In case of the maximum likelihood approach of Hansen and Seo (2002), this idea is called *concentrated* or *profile* likelihood.

For practical computation, the constraint formulated in (7) has to be accounted for in order to ensure a reasonable number of observations for the estimation of  $\mathbf{A}^{(j)}$ . It is modified in the following way to ensure a minimal proportion of observations in each regime

$$\pi_0 < \frac{T_j}{T} < 1 - l \cdot \pi_0 \quad (12)$$

where  $T_j$  denotes the number of observations in regime  $j$  and  $l$  the number of thresholds of the model. The trimming parameter  $\pi_0$  is usually set to 0.05 or 0.1.

### *Interpretation*

As sketched above, the TVECM specification, as long as the model is assumed or tested to have two effective thresholds, has a immediate economic motivation, namely the concept of commodity points suggested by Heckscher which corresponds to the transaction costs in a Enke-Samuelson-Takayama-Judge spatial equilibrium. Price adjustment as a consequence of trade only takes place if the price spread between two markets exceeds the “band of inaction”. The latter is delimited by the two estimated thresholds. Prices move independently within this band and are not cointegrated.

The thresholds are interpreted as transaction costs which render trade costly and thus inhibit it to a certain extent. Hence, the assumption that the thresholds are symmetric around the equilibrium error may not be appropriate in every case since economic theory does not give indications for this assumption and transaction costs are likely to be direction-specific (e.g. backhaul). The transaction costs encompass much more aspects than only the expenses for transportation. They are any costs with respect to temporal and financial expenses connected with the search for information, financing the trading process and legal duties (Shepherd, 1997). Barrett (2001) provides an extensive discussion of the components of such costs. A large part consists, of course, of the freight rates. Additionally, variable costs associated with insurance, financing or contracting are relevant. Exogenous costs such as underwriting fees or testing charges might apply. Furthermore, average duties on the product and immeasurable transaction costs such as opportunity and search costs or risk premia might also play a role.

Consequently, if the deviation  $ect_t$  from the long-run equilibrium, as depicted in the middle

---

<sup>28</sup> Detailed accounts of the algorithm are, for example, given in Serra and Goodwin (2003) and Park et al. (2007).

panel of Figure 3, is less than the transaction costs  $\theta^{(1)}$  of trade from market  $B$  to market  $A$  it is corrected upward toward zero, i.e., the long-run equilibrium, by trade into this direction. If it exceeds the transaction costs  $\theta^{(2)}$  of trade from  $A$  to  $B$  then it is corrected downward toward zero. Within the band spanned by the two transaction costs there is no incentive for trade in either direction. Prices move independently in this corridor, thus  $\alpha^{(2)} = 0$ , until the deviation from their long-run equilibrium exceeds either of the transaction costs once more, trade becomes profitable and the deviation decreases in the following periods, i.e., error correction takes place.

To our knowledge, only very few papers opt for a broader interpretation of the estimated thresholds than only as transaction costs. Trenkler and Wolf (2003) suggest an extension of the interpretation of these parameters which they call *cost parameter* to social, cultural and technical aspects. Moreover, they discuss the potential impact of nominal fixed transaction costs on market integration when aggregate price levels fluctuate. Obstfeld and Taylor (1997) note that the estimated thresholds, which they call in line with Heckscher's terminology commodity points, may reflect more aspects than only costs of transport and restrictions to trade. They refer to this additional component as *sunk costs of arbitrage*. O'Connell and Wei (2002) provide a comprehensive discussion of the importance of fixed and variable *market frictions* for deviations from the LOP. They hypothesize that also costs connected to the change of preferences and technology such as costs of labor migration or of entering and exiting a market might be relevant. Coleman (1995, 2004) addresses this argument as well.

## MARKOV-SWITCHING VECTOR ERROR CORRECTION MODELS

### *Basic Idea*

The underlying concept of the MSVECM emerged in the area of the econometrics of time-varying parameters. Goldfeld and Quandt (1973) develop a switching regression model characterized by parameter changes governed by a Markov chain. Hamilton (1989) extended this approach to the analysis of time series. A nonlinear VECM whose equilibrium errors follow a Markov process is suggested by several authors. Furthermore, the model is extended to the cointegration framework. Jackman (1995) proposes such a model for analyzing the determinants of presidential approval in the United States. Krolzig (1996, 1997) develops the MSVECM as a special case of the more general Markov-switching vector autoregression model whereas Hall et al. (1997) apply a MSVECM to the analysis of house prices in the United Kingdom.<sup>29</sup> Applications of the model are mainly found in business cycle and financial research, e.g. Krolzig and Toro (2001), Francis and Owyang (2003) or Psaradakis et al. (2004a), the latter suggesting further applications. Krolzig et al. (2002) analyze the British labor market. In PT analysis, the model is much less frequently applied than the TVECM; Brümmer et al. (2008) propose it to analyze vertical price transmission between wheat and flour in Ukraine.

As mentioned above, the MSVECM, as an elementary subclass of TAR models, is in general

---

<sup>29</sup> They suggest two specifications of the model with, first, constant transition probabilities and, second, transition probabilities as function of the equilibrium error. However, we focus in this paper only on the former case which represents the simplest form of the MSVECM with time-invariant transition probabilities. For extensions of this framework, see also, for example, Diebold et al. (1994), Hamilton and Raj (2002a) or Camacho (2005).

suitable for the analysis of PT due to the threshold behavior of trade as depicted in Figure 2. Hamilton (1989) characterizes the particular form of nonlinearities Markov-switching models are suitable for as “discrete shifts in regime-episodes across which the dynamic behavior of the series is markedly different”. Hamilton (1994, chap. 22.4) emphasizes the intuitively appealing characteristics of this class of models which are comprehensiveness and great flexibility. Hamilton and Raj (2002b) mention that “normal behavior of economies is occasionally disrupted by dramatic events that seem to produce quite different dynamics for the variables that economists study”. Moreover, Psaradakis et al. (2004a) note that the MSVECM is “best suited to situations where the change in regime is triggered by a sudden shock to the economy, situations which might not be adequately described by models with smooth transitions or threshold effects”.

The MSVECM can be characterized as a TAR model with exogenous determination of the states, i.e., the regimes are not a function of the analyzed price series themselves but of external determinants which, in contrast, do not have to be observed. Such determinants might act as general driving forces of trade, prices and a number of further economic variables. In the context of PT, regime-switching seems plausible which may not exclusively be determined by the equilibrium error, but rather by the “general state” of the trading process or even of the surrounding political or economic system. Price transmission behavior is likely to change temporarily due to external factors such as general characteristics of the political economic system. Raj (2002) mentions national policy changes, economic recessions, financial panics or wars in the context of business cycles. Further “sources of abrupt change” such as government actions in form of the introduction or the elimination of legal regulations are alluded to in Hamilton (1995). Chamley (1999) shows that an unique equilibrium may exist in a world characterized by imperfect information showing episodes of high and low economic activity which may, among others, result in differing equilibrium adjustment and short-run dynamics of PT. Furthermore he shows that switches between such regimes occur randomly.

One can think of traders’ temporary insecurity about the future due to elections or turmoils in politically unstable countries or due to exceptional positive or negative expectations about the near economic future such as forecasts of strong price rises.<sup>30</sup> Hence, traders face, for such limited periods, quite differing conditions ranging from increased uncertainty to the impossibility of trade and are likely to show, as a consequence, temporarily differing behavior. Such periods are hardly measurable. Agricultural scandals as they occasionally occur in Europe and the United States result in at least temporary changes of consumer demand might lead to transitionally different transmission of price signals. Further events such as temporal or new legal regulations, crop failures or transient demand changes are likely to change the “normal” trade dynamics in an abrupt manner.<sup>31</sup> Trade and hence price adjustment dynamics are furthermore subject to the asymmetries of the business cycle which first were hypothesized by Keynes (1936).

The importance of a further factor is stressed in the business cycle literature. Economic behavior and thus regime switches might be driven by extrinsic uncertainty in the sense of “random phenomena that do not affect tastes, endowments, or production possibilities” (Cass and Shell, 1983). This uncertainty is referred to in the literature as nonfundamentals, market

---

<sup>30</sup> For an example, see Agra Europe (2008, middle of page M/3): “Importers still panic buying...”.

<sup>31</sup> Martinez Peria (2002) aims at identifying speculative attacks on the European Monetary System during the first half of the 1990s by modeling a tranquil versus a speculative regime.

psychology, animal spirits, sunspots or self-fulfilling prophecies. Several publications show that these phenomena are apt to create business cycle fluctuations in the absence of shocks to fundamentals of the economy, see, for example, Azariadis (1981), Cass and Shell (1983), Howitt and McAfee (1992), Jeanne and Masson (2000) or Thomas (2004). Hamilton and Raj (2002b) see the cause of such potential impact in “agents’ believe such nonfundamentals affect aggregate economic activity”. Such phenomena are likely to affect trade processes as well causing random switches between regimes of PT.<sup>30</sup>

### Model Structure

In general, the MSVECM is formulated identically to the TVECM in (6):

$$\Delta \mathbf{p}_t = \boldsymbol{\mu}^{(J_t)} + \boldsymbol{\alpha}^{(J_t)} \boldsymbol{\beta}^\top \mathbf{p}_{t-1} + \sum_{i=1}^{k-1} \boldsymbol{\Psi}_i^{(J_t)} \Delta \mathbf{p}_{t-i} + \boldsymbol{\epsilon}_t \quad (13)$$

The number of regimes is denoted with  $M$  so that  $J_t = j \in \{1, 2, \dots, M\}$ . The model can, of course, compactly be written as in (10). Each regime-dependent variable takes a certain value depending on the value of the indicator variable  $J_t$  at time  $t$ , for example  $\boldsymbol{\alpha}^{(J_t)} = \boldsymbol{\alpha}^{(j)}$  if  $J_t = j$ , i.e.,

$$\boldsymbol{\alpha}^{(J_t)} = \begin{cases} \boldsymbol{\alpha}^{(1)} & \text{if } J_t = 1 \\ \vdots & \\ \boldsymbol{\alpha}^{(M)} & \text{if } J_t = M. \end{cases} \quad (14)$$

The regimes  $j$  of the MSVECM (13) are thought of as determined by a probabilistic process which has  $M$  states, i.e., they are assumed to be realizations of a *latent  $M$ -state Markov chain* with discrete state space in discrete time. The regime-dependent parameters are constant in each state but are allowed to change across states. Hence, each state of the underlying Markov chain directly corresponds to a regime of PT. Furthermore, the chain determines the regime switching.

The key element of the model is the  $(M \times M)$  *transition matrix*  $\boldsymbol{\Gamma}$  which contains the *transition probabilities*  $\gamma_{hj}$  for switching from state  $h$  to state  $j$

$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1M} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{M1} & \gamma_{M2} & \cdots & \gamma_{MM} \end{pmatrix} \quad (15)$$

where  $\gamma_{hj} = \Pr(J_{t+1} = j | J_t = h)$ . The Markov chain is assumed to be *homogeneous*, i.e., the transition probabilities are assumed to be time-invariant (compare footnote 29). Since switching from state  $h$  can only take place to one of the  $M$  states, the rows of  $\boldsymbol{\Gamma}$  sum up to unity by construction, i.e.,  $\boldsymbol{\Gamma} \mathbf{1}_M = \mathbf{1}_M$  where  $\mathbf{1}_M = (1 \ 1 \ \dots \ 1)^\top$  is a  $(1 \times M)$  vector, which is equivalent to  $\sum_{j=1}^M \gamma_{hj} = 1$ ,  $h = 1, \dots, M$ . The state process  $\{J_t\}$  determined by the transition probabilities  $\gamma_{hj}$  can thus be modeled quite flexibly. The larger, e.g., the probability on the diagonal of  $\boldsymbol{\Gamma}$  of some state is, the more persistent the behavior of this state will appear



and the less switches from this state to others will occur on average.

Several assumptions on the properties of the Markov chain have to be made in order to keep the model in a tractable complexity and to ensure desirable properties of the time series and the regimes. The RGP is assumed to satisfy the *Markov property*:

$$\Pr(J_{t+1}|J_t, J_{t-1}, \dots, \mathbf{p}_t, \mathbf{p}_{t-1}, \dots) = \Pr(J_{t+1}|J_t), \quad (16)$$

which is also referred to as a first-order or a memoryless process. This property states that the probability of switching to a new state in  $t + 1$  solely depends on the state of the preceding period  $t$  or as Chung (1960) puts it “the past should have no influence on the future except through the present”. Neither states before  $J_t$  nor any further variables such as the observed price series contain additional information regarding the regime switching. This assumption is not restrictive since each more complex model can be reparametrized into a first-order model, see, for example, Hamilton (1994, chap. 22.4) or MacDonald and Zucchini (1997, chap. 1.3). Moreover, the Markov chain has to be assumed to be ergodic and irreducible. The first condition is necessary to ensure a stationary unconditional probability distribution of the regimes.<sup>32</sup> The second one is needed to ensure the stationarity of the resulting time series. It requires that the ergodic probabilities of all states are larger than zero. Hence, it is assumed that any state can be reached from any state, that is that there are no absorbing states.

Figure 4 depicts the transition graph of a Markov chain of trade with  $M = 2$  states. It displays the possibilities for switching between two subsequent periods and the associated transition probabilities, i.e., it illustrates the information contained in the transition matrix  $\Gamma$ . In state  $j = 1$ , trade is not inhibited by, e.g., governmental measures, in state  $j = 2$  it is. The realization of a MSVECM in Figure 5 is generated according to (13) and corresponds to the Markov chain in Figure 4. If, say, the Markov chain is at  $t = 0$  in state  $J_0 = 2$ , as depicted in the upper panel of the figure, the loading parameters  $\alpha^{(J_t)}$  take the values  $\alpha^{(2)}$ , i.e., the correction of deviations from the long-run equilibrium in this period takes place with a high magnitude of PT of  $\pm 0.25$ . The switching to the state in the next time period  $t = 1$  solely depends on the previous state and the respective transition probabilities (Markov property). For  $J_0 = 2$  is the state  $J_1$  of the following period generated by a random switch based on the probabilities  $\gamma_{22} = 0.8$  and  $\gamma_{21} = 1 - \gamma_{22} = 0.2$ . Following this mechanism, the state in  $t = 1$  will be, say,  $J_1 = 2$ . PT in this period in turn is characterized by the adjustment speeds  $\alpha^{(2)}$ . These adjustment speeds will prevail until the Markov chain switches to state  $j = 1$  at some time  $t$  (the ninth time point in the figure).

### *Estimation*

In contrast to the estimation of the TVECM, one method is used for the estimation of the MSVECM as well as for general Markov-switching models in practice.<sup>33</sup> The particular challenge for estimation is similar to the TVECM. The researcher encounters uncertainty on

<sup>32</sup> The expected unconditional probabilities of the of being in any of the  $M$  states at arbitrary time are called the *ergodic* probabilities of the chain. Hence, the empiric frequencies of the regimes asymptotically equal the ergodic probabilities.

<sup>33</sup> Krolzig (1997, chap. 8) outlines with the multi-move Gibbs sampling a further estimation method which is based on Bayesian statistics. Mizrach and Watkins (2000) mention hill climbing. However, they recommend the EMA because of its superior properties.

two levels. First, the state process  $\{J_t\}$  depends on  $\mathbf{A}^{(j)}$  in (9). It has to be estimated since it is unknown. Second, the model parameters  $\mathbf{A}^{(j)}$  in turn depend on the unknown states  $J_t$  and are also to be estimated.<sup>34</sup>

Due to this two-fold uncertainty, the estimation consists of two steps which are the *expectation step* and the *maximization step*. These steps are iterated and the inference about the states and the estimates is updated until some convergence criterion is met. The procedure is called the *Expectation-Maximization algorithm* (EMA) (Figure 6). A particular filter is used in its first step which is the *Baum-Lindgren-Hamilton-Kim* (BLHK) filter.<sup>35</sup> The EMA was introduced by Dempster, Laird, and Rubin in 1977. Hamilton (1990) proposes the usage of the BLHK filter in connection with the EMA. Kim (1994) contributes an important improvement of the expectation step. Krolzig (1997, chap. 6) provides a detailed account of the method mentioning its major advantages which are computational simplicity and desirable convergence properties and discussing various extensions.

The algorithm is initialized by assuming starting values for the model parameters, the transition matrix and the probabilities of being in each of the  $M$  regimes at  $t = 1$ . The following expectation step draws inference about the unobserved regimes. First, the observations are filtered with the BHLK filter which yields the *filtered probabilities*. These are the probabilities that the observation at time  $t$  has been generated by each of the  $M$  regimes conditional on the data up to  $t$  and the estimated model parameters which are, in case of the first iteration, the initially assumed ones. Afterwards, the full sample *smoothed probabilities* are obtained on the basis of the filtered probabilities by a backward recursion. They represent the probabilities for each of the  $M$  regimes that it has occurred at time  $t$  conditional on the entire sample at hand. Equivalently, they might be interpreted as the probabilities that the observation at time  $t$  has been generated by regime  $j$  conditional on the entire sample.

The maximization step computes the update of the maximum likelihood estimates of all parameters which include the transition probabilities, the vector error-correction parameters and the probabilities of being in each of the  $M$  regimes at  $t = 1$ , that is, the initial state. The transition probabilities  $\gamma_{hj}$  are updated as the ratio between the summed probabilities of switches from  $h$  to  $j$  and of occurrences of regime  $h$  throughout the sample. Both quantities are calculated on the basis of the smoothed probabilities from the performed expectation step. The regime-dependent vector error correction parameters  $\mathbf{A}^{(j)}$  are calculated via generalized least squares estimation in which the observations are weighted by their smoothed probabilities. The second step finishes with the update of the probabilities of being in each of the  $M$  regimes at  $t = 1$  which are estimated by the smoothed probabilities for  $t = 1$ . The first iteration has thus been completed. The second iteration starts with utilizing the updated parameters from the previous one for the calculations in the expectation step where inference about the states is updated again. The second iteration is then completed by the update of the parameter estimates in the maximization step. The third iteration starts and so on until some reasonable convergence criterion is met.

This algorithm works for the estimation of Markov-Switching models in general. For the MSVECM in particular, Krolzig (1996) recommends a two-step estimation where first the

<sup>34</sup> Hamilton (1990) mentions three problems of interest to the researcher which are the inference about the unobserved regimes of the sample, the conditional forecast and the estimation of the model parameters including the transition matrix.

<sup>35</sup> For more details compare Krolzig (1997, chap. 5).

cointegration vector and the equilibrium errors are obtained. The equilibrium error may then be treated as an exogenous regressor in the model which becomes a general MSVAR model. The EMA can then be applied to the latter as described above.

### *Interpretation*

In the case of the MSVECM, the inference on the regimes is of probabilistic nature. The RGP is assumed to follow an latent Markov chain. Hence, the researcher cannot say with certainty which regime has occurred at some time  $t$ . The only measure allowing inference on this question are the smoothed probabilities which lie between zero and one.

Allowing for such non-deterministic statements regarding the occurrence of regimes turns out to be a reasonable and justified approach. Hamilton and Raj (2002b) note that there is a “growing consensus among economists that regime changes might be more appropriately modeled as arising from a probability process such as the Markov process”. Trade as well as business cycles or presidential approval are highly complex processes generated by unknown dynamics which are very likely to be of nonlinear character. Although the methodology finds evidence in the data at hand that some observations seem likely to follow a different regime, the researcher can, of course, not be completely sure about this evidence because the true RGP remains unknown. This fact is acknowledged by considering probabilistic statements regarding the incidence of the regimes. The regime with the highest smoothed probability for some time  $t$  is most likely to occur.

The interpretation of the regimes is far from being obvious a priori. It is much less straightforward as for the TVECM. The Markov-switching methodology is capable to identify distinct regimes among the observations of the sample. However, it relies exclusively on the sample by doing so. Hence, it is the researcher’s task to make sense of the identified regimes since no immediate interpretation based on economic theory is available as, e.g., in case of Heckscher’s supposition for the TVECM. The regimes have to be thoroughly analyzed and contrasted. Furthermore, an instructive endeavor might be to hypothesize the number and timing of regimes or at least potential determinants before performing the econometric analysis. By carefully analyzing potentially relevant events in the political and economic environment during the sample period, insights into the dynamics of the markets under study may be gained. The data analysis might then be used less as an exploratory but rather as a confirmatory tool. Jackman (1995) discusses the danger of the ex post “labelling of states”. He argues that a thorough interpretation of the estimates of each regime is necessary for characterizing the detected states. Alternatively, one might try to impose some structure on the Markov process or approach the issue from a Bayesian point of view by incorporating prior knowledge.<sup>36</sup>

The MSVECM allows, in a similar way as the TVECM does, not only for regimes characterized by different speeds of error adjustment but also for periods where no error correction takes place as, for example, in Psaradakis et al. (2004a). The latter case is particularly interesting. Such a regime is in contrast to the TVECM not bounded so that the longer the regime prevails the farther the prices, which are not cointegrated in this regime, may wander away from the equilibrium relationship. Such random walk behavior leads to high

---

<sup>36</sup> He provides in his article a comprehensive and detailed example of the first approach.

deviations from equilibrium which are not corrected despite of their magnitude. Thus, such a regime might be interpreted as being characterized by prohibitive transaction costs which do not allow for any trade although deviations from equilibrium might become huge. Such an extreme regime of PT might, for example, be caused by political intervention or other forms of prohibitive trade barriers which either lead to immense costs of trade or even do not allow for trade at all. Consequently, the MSVECM may be seen as being able to detect temporarily changing transaction costs where the change takes place in form of discrete shifts.

### 3 Comparison

#### 3.1 Conceptual Comparison

The application of TAR models to PT analysis is appropriate in general due to the regime-dependent behavior of PT as depicted in Figure 2. The application of the TVECM in particular possesses with Heckscher's supposition and the spatial arbitrage models of Takayama and Judge immediate economic justification; the MSVECM does have justification only to a limited extent mainly due to the lack of attention it has attracted yet in the field. Nevertheless, the application of the latter model in PT analysis seems intuitively very reasonable, particularly in cases where "discrete shifts in regime-episodes" (Hamilton, 1989) seem to be present in the data and the trade process was "occasionally disrupted by dramatic events" (Hamilton and Raj, 2002b) or "a sudden shock" (Psaradakis et al., 2004a).

Both models can be formulated in terms of (9) which represents a special case of the general threshold model specification in (2). Although both approaches model regime-dependent behavior of time series and belong to the group of piecewise linear TAR models (Figure 1), the philosophy regarding their underlying RGPs differs fundamentally. This leads to differing estimation methods and interpretation of results.

The regime process  $\{J_t\}$  is in case of the TVECM assumed to exclusively be generated by the first lag of some linear combination of the two price series under investigation, i.e.,  $J_t = f(p_{t-1}^A, p_{t-1}^B)$ . In case of the MSVECM, it is rather assumed to be a function of one or more exogenous variables  $y, z, \dots$  which might be thought of as the "general state" of the system  $J_t = f(y_t, y_{t-1}, \dots, y_{t-l}, z_t, z_{t-1}, \dots, z_{t-r}, \dots)$  where  $l, r \in \mathbb{N}^+$ . In contrast to the former model, the state process is allowed to be latent. Consequently, no observations on the regime generating variable(s) are required, they even may stay entirely unspecified. In this light, the assumption of the TVECM that the equilibrium error  $ect_t$  is the only variable determining the regimes seems restrictive. However, if the time series to be analyzed emerged in a stable economic and political environment in the absence of abrupt changes and other events which are likely to influence trade, the TVECM is the more appropriate model. It implies that there are at least two regimes in the data, in case of trade reversals even three regimes, and that the deviations  $ect_t$  from the long-run equilibrium is the only variable causing regime switching.

In the case of the existence of only one spatial equilibrium condition in the data or a highly unstable political and/or economic environment, in which trade as one aspect of the economy is embedded, regimes of PT are not likely to be (exclusively) determined by the equilibrium error. Regime shifts due to exogenous factors may superimpose the (weak)

regimes created by spatial equilibrium conditions and dominate the RGP. Consequently,  $ect_t$  does not represent the variable causing regime shifts. The MSVECM would in such a case be more appropriate. Consequently, the assumption that the equilibrium error represents the only variable causing nonlinear PT might in some settings not reflect reality.

These diverging suppositions regarding the RGP entail differing inference concerning the regime incidences. Whereas statements about the regime occurred at some time  $t$  can be made with certainty for the TVECM, they can only be of probabilistic nature in case of the MSVECM. Allowing for non-deterministic statements regarding the regime incidences turns out to be a reasonable and justified approach. Trade as well as business cycles or presidential approval are highly complex processes generated by unknown dynamics which are likely to be of nonlinear character. Although the methodology is capable to detect evidence in the data at hand that some observations are likely to follow a different regime, the researcher can, of course, not completely be sure about such findings because the true data generating process (DGP) remains unknown. This remaining uncertainty is acknowledged by considering probabilistic statements regarding the occurrence of the regimes.

In the case of the TVECM, the statements about the occurrence of regimes are deterministic in the sense that a certain regime  $j$  has or has not occurred at time  $t$  with certainty, i.e., the point estimates of  $ect_t$  can uniquely be assigned to the  $l$  regimes of the model which is implicitly formulated in (9). The binary variable  $d_t^{(j)}$  takes the value 1 if the regime  $j$  occurs at time  $t$  or zero otherwise. Clearly, such a deterministic all-or-nothing statement is more restrictive than the corresponding probabilistic statement of the MSVECM, however on the other hand, it allows for an easier interpretation. Whereas the Markov-switching approach acknowledges the uncertainty concerning the unknown true DGP, the TVECM approach does not. It instead suggests that always when the estimated threshold variable falls into a certain interval, the corresponding regime prevails with certainty. This implication is quite strong and may, of course, not be true for all observations. This assignment may have occurred occasionally by chance instead of being caused by the supposed underlying RGP.

Both models are capable to detect regimes characterized by different rates of error correction as well as regimes in which no adjustment behavior takes place. Although the TVECM does not explicitly model transaction costs, the threshold estimates are, at least in case of a TVECM(3) specification, usually interpreted as such. Moreover, they are often assumed to be constant during the sample period. The MSVECM does not model transaction costs either. Nevertheless, it might be understood as to allow the transaction costs to shift during the sample period since an identified regime without adjustment may potentially be caused by temporary prohibitive transaction costs.

The estimation of both models faces the same challenge. The parameters of the regime-dependent VECM are unknown and depend on the regime process  $\{J_t\}$ . This process itself is unknown because the quantities characterizing it, which are the thresholds and the transition probabilities, respectively, are unknown as well. Their estimates in turn depend on the unknown vector error correction parameters. The estimation methods of SCLS and EMA, though variants of the maximum likelihood principle, tackle this task in different ways. In the former case, a number of modifications are easily implemented to estimate the model in dependence of various optimization parameters (Table 5). The researcher determines candidate values of the optimization parameters which typically form a regular spaced grid. Conditionally on these, a optimization criterion is evaluated. The combination of parameters

optimizing the criterion is selected as the final estimates. The EMA, in contrast, iterates conditionally on one set of starting values until a convergence criterion is met. Inference on the unobserved regimes is recursively drawn for all observations conditional on the parameter estimates of the previous iteration. Parameter estimates, in turn, are obtained conditionally on the evidence on the regimes from the preceding step. In case of both methods, the number of regimes may either be justified theoretically, evaluated by econometric tests or determined by using a model selection criterion.<sup>37</sup>

Differences between both methods concerning the interpretation have already been addressed. Of course, regime frequencies and regime-dependent half-lives of the adjustment process may be calculated in both cases. Additionally, the expected duration of the regimes may be calculated for the MSVECM. The regime frequency is estimated by the proportion of observations generated by the regime. In the case of the MSVECM is it the proportion of observations which is likely to be generated by the regime whereat the meaning of the term “likely” has to be determined by the researcher. The regime with the highest smoothed probability among all regimes for some time  $t$  is considered to be most likely. The half-life of a adjustment process is the time which is required to correct half of the deviation from equilibrium of a given shock (Van Campenhout, 2007) and can easily be obtained.<sup>38</sup> However, the calculation of half-lives is more complicated for vector autoregressions of higher order as pointed out by Ben-Kaabia and Gil (2007). The expected duration  $\lambda^j$  of regime  $j$  can be calculated as  $\lambda^j = E[\lambda | J_t = j] = \frac{1}{1-\gamma_{jj}}$  as outlined in Krolzig (1997, section 11.3.4).  $\gamma_{jj}$  denotes the transition probability of staying in regime  $j$  as depicted in the transition matrix  $\Gamma$  (15).

Furthermore, it has been noted that the interpretation of the regimes of the TVECM is relatively straightforward. However, the results are occasionally interpreted in a narrower or broader sense. In case of the MSVECM, some effort has to be devoted to carefully analyzing the identified regimes. The parameters and further descriptive variables have to be interpreted and consulted in detail in order to receive insights regarding the distinguishing characteristics and the nature of the regimes.

### *Distinct Regime Generating Processes: Exogenous vs. Endogenous Switching*

As mentioned above, both models may be formulated in terms of the general specification (10). Nevertheless, the RGP differ fundamentally. A formulation in terms of a common notation permits insights from one perspective regarding their similarities and differences. In particular, the RGP of the TVECM can be reformulated by using the notation of the MSVECM. The key distinction in the philosophies underlying both approaches becomes apparent and can well be contrasted by using a unified notation. In the following, we restrict

<sup>37</sup> We do neither address the issue of testing for nonlinearity nor impulse response analysis in this paper since it is beyond its scope. However, we will briefly discuss the issue of model selection below.

<sup>38</sup> The half-life  $\kappa$  is the solution in  $z_{t+\kappa} = \frac{z_t}{2}$  based on a SETAR specification of the equilibrium error process as, for example, in Balke and Fomby (1997, equation (1)). As they have shown, the SETAR and the TVECM specification are equivalent, the former represents a reparametrization of the latter and vice versa. Hence, the half-life is calculated as  $\kappa = \frac{\ln(0.5)}{\ln(1+\rho^{(j)})}$  based on the SETAR specification. In the TVECM specification,  $\rho^{(j)}$  is not estimated. It thus has to be replaced by  $\rho^{(j)} = 1 + \beta^\top \alpha^{(j)} = 1 + \alpha^{A(j)} - \beta^B \alpha^{B(j)}$  where  $\alpha^{A(j)}$  denotes the magnitude of PT of the  $j$ 's regime of the price series of market  $A$  and  $\beta = (\beta^A \ \beta^B)^\top = (1 \ \beta^B)^\top$  the cointegration vector between both prices.

the comparison to the simplest specification of  $l + 1 = M = 2$  regimes for each of the models.

The transition matrix of a MSVECM with  $M = 2$  regimes has the following structure

$$\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & 1 - \gamma_{11} \\ 1 - \gamma_{22} & \gamma_{22} \end{pmatrix} \quad (17)$$

because  $\Gamma \mathbf{1}_M = \mathbf{1}_M$ . Thus, the corresponding Markov-chain is characterized by only two transition probabilities. It is assumed to be homogeneous, i.e., the transition probabilities are assumed to be constant, and irreducible, i.e., for the transition probabilities holds that  $0 < \gamma_{11} < 1$  and  $0 < \gamma_{22} < 1$ . The process is assumed to satisfy the Markov property. Alternatively, the transition matrix can be rewritten in terms of a set of conditional probabilities

$$\Pr(J_t = 1 | J_{t-1} = 1) = \gamma_{11} \quad (18)$$

$$\Pr(J_t = 1 | J_{t-1} = 2) = 1 - \gamma_{22} \quad (19)$$

$$\Pr(J_t = 2 | J_{t-1} = 1) = 1 - \gamma_{11} \quad (20)$$

$$\Pr(J_t = 2 | J_{t-1} = 2) = \gamma_{22}. \quad (21)$$

A TVECM of  $l + 1 = 2$  regimes possesses one (effective) threshold  $\theta^{(1)}$ . The RGP is characterized as implicitly formulated in (6)

$$J_t = \begin{cases} 1 & \text{if } ect_{t-1} \leq \theta^{(1)} \\ 2 & \text{if } ect_{t-1} > \theta^{(1)}. \end{cases} \quad (22)$$

As mentioned above,  $J_t$  takes the values  $j$  with certainty. Depending on the size of the threshold variable which is in this case the first lag of the error correction term  $ect_{t-1}$ , the regime  $j$  prevails with a probability of 100% at time  $t$ . Hence, it becomes evident that the RGP may be formulated in terms of conditional probabilities which can be summarized into a transition matrix. However, due to the mentioned certainty these probabilities take either of the values 0 or 1. They are conditional on the previous state  $J_{t-1}$ , however, they additionally depend on the threshold variable  $ect_{t-1}$  of the previous period. The corresponding transition probabilities are as follows

for  $ect_{t-1} \leq \theta^{(j)}$

for  $ect_{t-1} > \theta^{(j)}$

$$\Pr(J_t = 1 | J_{t-1} = 1, ect_{t-1}) = \omega_{11}^{(1)} = 1 \quad \Pr(J_t = 1 | J_{t-1} = 1, ect_{t-1}) = \omega_{11}^{(2)} = 0 \quad (23)$$

$$\Pr(J_t = 1 | J_{t-1} = 2, ect_{t-1}) = \omega_{21}^{(1)} = 1 \quad \Pr(J_t = 1 | J_{t-1} = 2, ect_{t-1}) = \omega_{21}^{(2)} = 0 \quad (24)$$

$$\Pr(J_t = 2 | J_{t-1} = 1, ect_{t-1}) = \omega_{12}^{(1)} = 0 \quad \Pr(J_t = 2 | J_{t-1} = 1, ect_{t-1}) = \omega_{12}^{(2)} = 1 \quad (25)$$

$$\Pr(J_t = 2 | J_{t-1} = 2, ect_{t-1}) = \omega_{22}^{(1)} = 0 \quad \Pr(J_t = 2 | J_{t-1} = 2, ect_{t-1}) = \omega_{22}^{(2)} = 1. \quad (26)$$

These probabilities can be summarized into the following transition matrix  $\Omega$  which depends

on the threshold variable  $ect_{t-1}$

$$\Omega = \Omega_t = \begin{cases} \Omega^{(1)} = \begin{pmatrix} \omega_{11}^{(1)} & \omega_{12}^{(1)} \\ \omega_{21}^{(1)} & \omega_{22}^{(1)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} & \text{if } ect_{t-1} \leq \theta^{(1)} \\ \Omega^{(2)} = \begin{pmatrix} \omega_{11}^{(2)} & \omega_{12}^{(2)} \\ \omega_{21}^{(2)} & \omega_{22}^{(2)} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} & \text{if } ect_{t-1} > \theta^{(1)}. \end{cases} \quad (27)$$

This transition matrix  $\Omega$  is equivalent to the usual specification of the RGP of a TVECM as, e.g., formulated in (22). It highlights the similarities and the differences of the TVECM in comparison to the transition matrix  $\Gamma$  of the MSVECM. In case that the error correction term is smaller than the threshold, it does not matter in which state the process was in the preceding period  $t - 1$  it takes the regime  $j = 1$  in time  $t$  with probability 1. This regime is either reached by staying in the regime 1 which is expressed by  $\omega_{11}^{(1)}$  or by switching from regime 2 to one 1 expressed by  $\omega_{21}^{(1)}$ . In case that the error correction term is larger than the threshold, the process will be in regime 2 at time  $t$  with probability 1.

It becomes apparent that the transition matrix  $\Omega$  is not constant over time because the respective transition probabilities take values depending on the magnitude of the error correction term. Thus, the matrix symbol is augmented by the time index  $t$  and denoted as  $\Omega_t$ . Consequently, this RGP is, in contrast to the MSVECM in (17), not homogeneous. Second, the transition probabilities are restricted to take either of the values 0 or 1. In this sense, the switching is purely deterministic. It either occurs or it does not, each of both with certainty.

Third, the process does not satisfy the Markov property because

$\Pr(J_t|J_{t-1}, ect_{t-1}) \neq \Pr(J_t|J_{t-1})$ . It has been mentioned that the size of the error correction term determines not only the regime but also the transition probabilities in  $\Omega_t$ . Hence, it contains additional information for the switching of the regimes. In contrast, the entire information relevant for switching is encompassed in the previous state in case of the MSVECM as denoted in (18) to (21). Moreover, the transition probabilities in  $\Omega_t$  do exclusively depend on the threshold variable because a state  $j$  is reached at time  $t$  from any preceding state with certainty, exclusively determined by the magnitude of the threshold variable. This view corresponds to the usual interpretation of the TVECM that the switching exclusively depends on the error correction term. Hence, it holds that

$\Pr(J_t|J_{t-1}, ect_{t-1}) = \Pr(J_t|ect_{t-1})$  and (23) to (26) and the transition matrix  $\Omega_t$  simplify to

$$\Omega'_t = \begin{pmatrix} \omega'_{11} & \omega'_{12} \\ \omega'_{21} & \omega'_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (28)$$

where  $\omega'_{1j} = \Pr(J_t = j|ect_{t-1} \leq \theta^{(1)})$  and  $\omega'_{2j} = \Pr(J_t = j|ect_{t-1} > \theta^{(1)})$ ,  $j = 1, 2$ . This formulation emphasizes the fact that the switching is exclusively governed by the observed threshold variable in a deterministic way whereas, in the case of the MSVECM, the switching is governed by the unobserved previous state in a probabilistic way as formulated in (17) to (21). Furthermore, the threshold variable is in case of the TVECM a linear combination of the two price series under study.

Consequently, the regime switching of the TVECM, in contrast to the MSVECM, is not exogenous. It has been shown that the probabilities  $\omega'_{hj}$ ,  $h, j = 1, 2$  for switching to regime  $j$  from time  $t - 1$  to time  $t$  depend exclusively on the error correction term  $ect_{t-1}$ . This



threshold variable itself is a function of the observed price series  $\{p_t^A\}$  and  $\{p_t^B\}$  of markets  $A$  and  $B$  because it represents a linear combination of the first lags of both series  $ect_{t-1} = \beta^\top \mathbf{p}_{t-1}$ . Hence, the transition probabilities are as well a function of the two price series  $\omega'_{hj} = f(ect_{t-1}) = f(p_{t-1}^A, p_{t-1}^B)$  and the switching is thus endogenous. The switching of the MSVECM is exogenous since the variable causing the switching remains unknown.

### *Model Selection*

The characterization of either of the considered approaches has been motivated in the previous sections by economic theory and heuristic evidence. To our knowledge, no econometric tests exist which explicitly test nonlinear model classes against each other. Mellows (1999) notes that this constitutes a common problem in nonlinear time series analysis. Although a number of tests have been developed to check for nonlinear behavior such as Hansen (1997), Hansen (1999) or Hansen and Seo (2002) for TAR or Hansen (1992) for Markov-switching models, the determination of the most appropriate model class for the data at hand remains an issue for future research. Mellows (1999, chap. 5) suggests a classification method based on an idea of Tong (1990) which uses parametric bootstrap and discriminant analysis. However, a simulation study reveals that a model class is more likely to be identified correctly by the approach the more pronounced the specific nonlinearities of the underlying process are. Under weak nonlinearities, one quarter to more than half of the models are wrongly classified as linear. However, diagnostic tests as, for example, developed in Hamilton (1996) help to assess the adequacy of the chosen model.

Considerations of testing one model against another might not be of immediate interest in the context of applied research in price transmission analysis since the application of a certain model class has to go along with an appropriate interpretation of results. Qualitative reasoning of the appropriateness of the chosen model accompanied by formal tests of nonlinearities in the time series and a thorough interpretation of estimation results might be a recommendable approach to tackle this issue.

Besides the questions which of the nonlinear models to choose, the question whether nonlinear are superior to linear time series models has to be addressed. It has been discussed above that both models considered here possess much appeal from an economic perspective. Clements and Krolzig (1998) assess the performance of two nonlinear time series models in comparison with linear AR models in business cycle analysis. They find that although both, Markov-switching autoregressive and SETAR models, are well capable to model the particular features of the data, their performance in forecasting is not as superior relatively to AR models. This question is not discussed for the models considered in this paper. However, more complex models are in general more capable to capture the distinctive features of the data. In contrast, their forecasting performance does not necessarily increase due to their complexity. This general fact is supposed also to hold for the TVECM and the MSVECM.

## 3.2 A Simulation Study

Additionally to the conceptual comparison of the TVECM and the MSVECM, we are interested in assessing the performance of the estimation methods under ideal circumstances.

In particular, we apply the sequential conditional least squares (SCLS) and the Expectation-Maximization-Algorithm (EMA) to simulated data which is generated by a TVECM and a MSVECM respectively.<sup>39</sup> Problems of the SCLS estimation have rarely been addressed in applied research. Lo and Zivot (2001) study the performance of the method for data generated according to a TAR and a TVECM for symmetric thresholds of 3,5 and 10 and time series of 100, 250 and 500 observations respectively. They find “considerable uncertainty in the estimates [of the thresholds] for moderate sample sizes” in case of the unrestricted model. However, the estimates of restricted TAR and TVECM models have a much smaller bias and also a smaller variance. Furthermore, Trenkler and Wolf (2003) note that the estimates of the unrestricted TVECM are very unreliable.

We extend with this simulation study the works of Clements and Krolzig (1998), Lo and Zivot (2001) and Psaradakis et al. (2004b). The first article assesses the forecast performance of Markov-switching and TAR models relative to linear autoregressive models in business cycle analysis via a simulation study. Lo and Zivot study the performance of tests for threshold cointegration, threshold nonlinearity and specification tests and evaluate the estimation of the TVECM by an extensive simulation study. Psaradakis et al. (2004b) examine tests for cointegration, parameter instability, neglected nonlinearity, Markov-switching and a model selection procedure based on data which follows Markov-switching error correction.

We follow Lo and Zivot (2001) by generating data according the simple cointegration model as outlined in Balke and Fomby (1997).<sup>40</sup> Lo and Zivot generate data with varying lengths of the time series and magnitudes of the true thresholds which they assume to be symmetric. However, we adopt a more comprehensive setting by generating time series of length 150, 500 and 1500, respectively. We vary the thresholds  $\theta^{(j)}$ , the transition probabilities  $\gamma_{hj}$  and the error correction parameters  $\alpha^{(j)}$ . The data is generated according to a TVECM and a MSVECM respectively. Each model is assumed to have three regimes one of which is not showing error correction. In the following we briefly present some results of the simulation study for  $T = 500$ . Further results may be obtained from the authors.

### *Case I: SCLS and TVECM Data*

A very low share, namely only around 56% of all observations are correctly identified by the method. Hence, SCLS seems not to be able to identify the true regimes of TVECM-data as generated by ((B.2)) to a satisfying extent. The reason lies in the very high and varying MSE as depicted for  $\hat{\theta}^{(1)}$  in Figure 7. The MSE is of considerable magnitude and strongly depends on the true thresholds  $\theta^{(1)}, \theta^{(2)}$  as well as on the nuisance parameters  $\rho^{(1)}, \rho^{(3)}$  which govern the autoregressive process of the  $ect_t$  as formulated in (B.2). The bias is the smallest for  $|\theta^{(j)}| = \sigma_{\nu_t}^2$ . Hence, not only the magnitude of the thresholds themselves but rather the ratio between the absolute values of the thresholds and the variance of the innovations  $\nu_t$  in (B.2) seems also to have influence on the MSE. Both, the bias and the variance tend to increase with decreasing  $\alpha^{(j)}$  and with increasing  $T$ . This result is plausible since both, a small  $\alpha^{(j)}$  and a large  $T$ , lead to more realizations of the data away from the true thresholds and thus

<sup>39</sup> The Bayesian approaches of Luoma et al. (2004) and Balcombe et al. (2007) for estimating the TVECM are not considered here. We focus on SCLS and EMA since they are the predominant estimation methods in applied research in PT.

<sup>40</sup> For more details see Appendix B.

tend to increase bias and variance.

#### *Case II: SCLS and MSVECM Data*

Table 7 shows the classification of true vs. estimated regime incidences as percentages of all observations of the dataset. By summing up each of the columns the ergodic probabilities of the MSVECM-data may be obtained. The percentages on the diagonal denote the correctly identified regimes which sum up to only 40%. SCLS performs here even worse as in case I. Hence, SCLS does not seem to be suitable to detect the true regimes in the data which is simulated by a MSVECM-DGP.

#### *Case III: EMA and TVECM Data*

Table 8 shows the classification of true vs. estimated regime incidences as estimated by the EMA on TVECM-data. A share of only about 30% was correctly identified indicating that EMA does not perform well with data that follow a TVECM.

#### *Case IV: EMA and MSVECM Data*

Figure 8 provides some evidence on the performance of the EMA on data that are generated by a MSVECM. The share of correctly identified regimes lies at about 42%. Similarly to case I, the MSE of, say, the estimate of  $\pi^{(1)}$ , i.e., the ergodic probability of the first regime, is heavily influenced by the true ergodic probabilities as well as by the nuisance parameters  $\rho^{(j)}$ . Nevertheless, a strange pattern appeared in the estimation results. The bias for the first two regimes is very high in tendency and decreases as the true ergodic probabilities approach  $\frac{1}{3}$  and increases with decreasing  $\alpha^{(j)}$ . The variance is small and increases slightly with increasing  $\alpha^{(j)}$ . Both, the bias as well as the variance, decrease with the increasing number of observations per time series.

#### *Summary*

This Monte Carlo analysis leads to rather pessimistic results regarding the performance of the assessed estimation methods since only one third to one half of the regimes have correctly been identified. Lo and Zivot (2001) demonstrated that the estimates of the thresholds are biased. We extended this evidence and demonstrated that the estimates of the thresholds seem also heavily influenced by the true error-correction parameters. The EMA did yield satisfying results neither. However, these results cannot be generalized since they depend on the assumptions made beforehand, particularly the specification of the DGPs. In order to derive general statements further experiments have to be conducted.

## 4 Conclusion

This paper compares two time series models which are relevant for PT analysis and allow for nonlinear adjustment of deviations from the long-run equilibrium. Both model classes, the TVECM as well as the MSVECM, are characterized by parameters which may take different values depending on the regime of the data. They are constant within one regime but may differ across regimes. Such regime-dependent models allow the study of trade processes from a dynamic point of view in which the transmission of price signals between markets changes temporarily. Such sophisticated models may thus enhance the understanding of the interaction of markets. Although both models seem at first glance very similar due to their common property of regime switching, their underlying statistical concepts differ fundamentally. Consequently, each model is suited for a particular type of nonlinearity.

The TVECM is characterized by endogenous switching. The variable causing regime switches is assumed to be fully determined by the prices under study. The restriction of the switching mechanism to a particular relationship facilitates the interpretation of the model which matches economic theory very well. Such a constraint implies two aspects. First, if the explicit information contained in the threshold variable is correct, the model will yield more reliable results than more general ones. However, if the opposite is the case, then it will be farer away from the thruth than general models.

The MSVECM is more general with respect to the switching mechanism since it allows for exogenous switching independent of the price series analyzed. Furthermore, the determinants causing switching may even remain completely unspecified. Its key element is a latent Markov chain modelling the transition of regimes between subsequent points in time. The higher flexibility of the model comes at the cost of limited straightforward interpretability. Making sense of the identified regimes requires more effort than in the case of the former model.

The two models reflect different aspects of the complex economic reality, spatial equilibrium conditions on the one hand and unobserved states of the system on the other. If the price data to be analyzed were predominantly not subject to external impacts such as changing political, economic or natural interferences, it can be assumed that markets and trade processes were the main forces generating the data. A TVECM would be the more appropriate model in such a case since it explicitly draws on the economic information contained in the prices. Nevertheless, a TVECM requieres at least two regimes in the data to be estimable. It requires, for example, changing spatial equilibrium conditions. If, however, trade took mainly place into one direction or external interferences dominated the markets during the time period studied then a MSVECM might be more suitable. Most often, the reality will lie inbetween these two extreme cases. In such a case, the most appropriate model depends on the dominating impact. The two models can be expected to yield differing results since each emphasizes a certain aspect of PT.

Although both models, from an economic perspective, seem capable to lead to interesting insights into PT, the simulation study confirmed and extended evidence that the empiric application and the applicability constitutes a drawback of these models. However, improving the estimation of such models is not the only area for future research. Most often, the quantitative components of the theoretically postulated and econometrically estimated thresholds, i.e., the determinants causing nonlinear price adjustment, receive little attention.

Acquiring empirical evidence regarding their structure might help to develop adapted models. As a consequence, the interpretation of the estimated thresholds only in terms of transaction costs might turn out to be too narrow in some cases. The magnitude of the time lag  $d$  of the TVECM is likely to depend on the product traded, infrastructure etc. so that by relaxing the assumption that prices react to deviations from equilibrium within a time lag of only one period, insights into the dynamic processes of trade might be extended.

## References

- Agüero, J.M. “Asymmetric Price Adjustments and Behavior Under Risk: Evidence from Peruvian Agricultural Markets.” Working paper, University of California, Riverside, USA, 2007.
- Agra Europe. “Grain prices hit records again.” *Agra Europe*, No. 2301(2008):M/2–M/3.
- Alemu, Z.G. and G.R. Biacuana. “Measuring Market Integration in Mozambican Maize Markets: A Threshold Vector Error Correction Approach.” Contributed paper prepared for presentation at the *International Association of Agricultural Economists Conference*, Gold Coast, Australia, August 12-18 2006.
- Azariadis, C. “Self-Fulfilling Prophecies.” *Journal of Economic Theory*, 25(1981): 380–396.
- Bai, J. “Estimating multiple breaks one at a time.” *Econometric Theory*, 13(1997): 315–352.
- Bai, J. and P. Perron. “Estimating and Testing Linear Models With Multiple Structural Changes.” *Econometrica*, 66(1998): 47–78.
- Bakucs, L.Z. and I. Fertő. “Spatial Integration on the Hungarian Milk Market.” Paper prepared for presentation at the *Joint IAAE - 104th EAAE Seminar Agricultural Economics and Transition*, Corvinus University, Budapest, Hungary, September 6-8 2007.
- 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): 308–323.
- Balke, N.S. and T.B. Fomby. “Threshold Cointegration.” *International Economic Review*, 38(1997): 627–645.
- Barrett, C.B. “Market Analysis Methods: Are Our Enriched Toolkits Well-Suited to Enlivened Markets?” *American Journal of Agricultural Economics*, 78(1996): 825–829.
- Barrett, C.B. “Measuring Integration and Efficiency in International Agricultural Markets.” *Review of Agricultural Economics*, 23(2001):19–32.
- Baulch, R.J. “Spatial Price Equilibrium and Food Market Integration.” Dissertation, Food Research Institute, Stanford University, USA, 1994.
- Ben-Kaabia, M. and J.M. Gil. “Asymmetric Price Transmission in the Spanish Lamb Sector.” *European Review of Agricultural Economics*, 34(2007): 53–80.
- Ben-Kaabia, M., J.M. Gil, and M. Ameer. “Vertical Integration and Non-linear Price Adjustments: The Spanish Poultry Sector.” *Agribusiness*, 21(2005): 253–271.
- Ben-Kaabia, M., J.M. Gil, and L. Boshnjaku. “Price Transmission Asymmetries in the Spanish Lamb Sector.” Paper prepared for presentation at the *Xth EAAE Congress ‘Exploring Diversity in the European Agri-Food System’*, Zaragoza, Spain, 28-31 August 2002.

- Bhansali, R.J. Discussion of Tong and Lim (1980), 1980, p. 270.
- Brümmer, B., S. von Cramon-Taubadel, and S. Zorya. “A Markov-Switching Vector Error Correction Model of Vertical Price Transmission between Wheat and Flour in Ukraine.” Under revision for the *European Review of Agricultural Economics*, 2008.
- Camacho, M. “Markov-Switching Stochastic Trends and Economic Fluctuations.” *Journal of Economic Dynamics and Control*, 29(2005): 135–158.
- Cass, D. and K. Shell. “Do Sunspots Matter?” *Journal of Political Economy*, 91(1983): 193–227.
- Chan, K.S.. “Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model.” *The Annals of Statistics*, 21(1993): 520–533.
- Chamley, C. “Coordinating Regime Switches.” *The Quarterly Journal of Economics*, 114(1999): 869–905.
- Chen, L.-H., M. Finney, and K.S. Lai. “A Threshold Cointegration Analysis of Asymmetric Price Transmission from Crude Oil to Gasoline Prices.” *Economics Letters*, 89(2005): 233—239.
- Chung, K.L. *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, Berlin, Germany, 1960.
- Clements, M.P. and H.-M. Krolzig. “A Comparison of the Forecast Performance of Markov-Switching and Threshold Autoregressive Models for US GDP.” *Econometrics Journal*, 1(1998): C47–C75.
- Coleman, A.M.G. “Arbitrage, Storage and the ‘Law of One Price’: New Theory for the Time Series Analysis of an Old Problem.” Working paper, Princeton University, USA, 1995.
- Coleman, A.M.G. “Storage, Slow Transport, and the Law of One Price: Evidence from the Nineteenth Century U.S. Corn Market.” Discussion Paper No. 502, University of Michigan, USA, 2004.
- Dercon, S. and B. van Campenhout. “Dynamic Price Adjustment in Spatially Separated Food Markets with Transactions Costs.” Working paper, Katholieke Universiteit Leuven, Belgium, 1998.
- Dempster, A.P., N.M. Laird, and D.B. Rubin. “Maximum Likelihood from Incomplete Data via the EM Algorithm (with Discussion).” *Journal of the Royal Statistical Society*, B 39(1977): 1–38.
- Diebold, F.X., J-H. Lee, and G. Weinbach. “Regime Switching with Time-Varying Transition Probabilities”, in C.P. Hargreaves (ed), *Nonstationary time series analysis and cointegration*, Oxford University Press, Oxford, UK, 1994, chapter 10.
- Doornik, J.A. *Object-Oriented Matrix Programming Using Ox*, Timberlake Consultants Press, London, and <http://www.doornik.com/index.html>, Oxford, UK, 2002.

- Dorsey, R.E. and W.J. Mayer. "Genetic Algorithms for Estimation Problems with Multiple Optima, no Differentiability, and Other Irregular Features." *Journal of Business and Economic Statistics*, 13(1995): 53–66.
- Dumas, B. "Dynamics Equilibrium and the Real Exchange Rate in a Spatially Separated World." *Review of Financial Studies*, 5(1992): 153–180.
- Ejrnaes, M. and K.G. Persson. "Market Integration and Transport Costs in France 1825–1903: A Threshold Error Correction Approach to the Law of One Price." *Explorations in Economic History*, 37(2000): 149—173.
- Enders, W. *Applied Econometric Time Series*. John Wiley & Sons, 2nd ed., Hoboken, NJ, USA, 2004.
- Engle, R.F. and C.W.J. Granger. "Co-Integration and Error Correction: Representation, Estimation and Testing." *Econometrica*, 55(1987): 251-76.
- Enke, S. "Equilibrium Among Spatially Separated Markets: Solution by Electrical Analogue." *Econometrica*, 19(1951): 40–47.
- Escobal, J. "The Role of Public Infrastructure in Market Development in Rural Peru." Dissertation, Wageningen University, The Netherlands, 2005.
- P.L. Fackler and B.K. Goodwin. "Spatial Price Analysis" in B. Gardner and G. Rausser (eds), *Handbook of Agricultural Economics*, Vol. 1, Elsevier, Amsterdam, The Netherlands, 2001, pp. 971-1024.
- Fanizza, D.G. "Multiple Steady States and Coordination Failures in Search Equilibrium: New Approaches to the Business Cycle." Dissertation, Northwestern University, USA, 1990.
- Federico, G. "Market Integration and Market Efficiency: The Case of 19th Century Italy." *Explorations in Economic History*, 44(2007): 293—316.
- Francis, N. and M. Owyang. "Asymmetric Common Trends: An Application of Monetary Policy in a Markov-Switching VECM." Federal Reserve Bank of St. Louis Working Paper 2003-001B, St. Louis, USA, 2003.
- Goldfeld, S.M. and R.E. Quandt. "A Markov Model for Switching Regressions." *Journal of Econometrics*, 1(1973): 3–16.
- Goodwin, B.K. and T.J. Grennes. "Tsarist Russia and the World Wheat Market." *Explorations in Economic History*, 35(1998): 405—430.
- Goodwin, B.K. and D.C. Harper. "Price Transmission, Threshold Behavior, and Asymmetric Adjustment in the U.S. Pork Sector." *Journal of Agricultural and Applied Economics*, 32(2000): 543—553.
- Goodwin, B.K. and M.T. Holt. "Price Transmission and Asymmetric Adjustment in the U.S. Beef Sector." *American Journal of Agricultural Economics*, 81(1999): 630–637.



- Goodwin, B.K. and N. Piggott. "Spatial Market Integration in the Presence of Threshold Effects." *American Journal of Agricultural Economics*, 83(2001): 302–317.
- Goodwin, B.K., T.J. Grennes, and L.A. Craig. "Mechanical Refrigeration and the Integration of Perishable Commodity Markets." *Explorations in Economic History*, 39(2002): 154—182.
- Hall, S., Z. Psaradakis, and M. Sola. "Switching Error-Correction Models of House Prices in the United Kingdom." *Economic Modelling*, 14(1997): 517–527.
- Hamilton, J.D. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica*, 57(1989): 357–384.
- Hamilton, J.D. "Analysis of Time Series Subject to Regime Changes." *Journal of Econometrics*, 45: 39–70.
- Hamilton, J.D. *Time Series Analysis*. Princeton University Press, Princeton, USA, 1994.
- Hamilton, J.D. "Rational Expectations and the Economic Consequences of Changes in Regime", in K.D. Hoover (ed), *Macroeconometrics: Developments, Tensions, and Prospects*, Kluwer Academic Publishers, Boston, USA, 1995, chapter 9.
- Hamilton, J.D. "Specification Testing in Markov-Switching Time Series Models." *Journal of Econometrics*, 70(1996): 127–157.
- Hamilton, J.D. and B. Raj (2002a)(eds). *Advances in Markov-Switching Models. Applications in Business Cycle Research and Finance*. Physica-Verlag, Heidelberg, Germany, 2002.
- Hamilton, J.D. and B. Raj (2002b). "New Directions in Business Cycle Research and Financial Analysis", in Hamilton, J.D. and B. Raj (eds), *Advances in Markov-Switching Models. Applications in Business Cycle Research and Finance*, Physica-Verlag, Heidelberg, Germany, 2002.
- Hansen, B.E. "The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP." *Journal of Applied Econometrics*, 7(1992): S61–82; Erratum (1996) 11, 195–198.
- Hansen, B.E. "Inference in TAR Models." *Studies in Nonlinear Dynamics and Econometrics*, 2(1997): 1–14.
- Hansen, B.E. "Testing for Linearity." *Journal of Economic Surveys*, 13(1999): 551–576.
- Hansen, B.E. "Sample Splitting and Threshold Estimation." *Econometrica*, 68(2000): 575–603.
- Hansen, B.E. and B. Seo. "Testing for Two-Regime Threshold Cointegration in Vector Error-Correction Models." *Journal of Econometrics*, 110(2002): 293–318.
- Heckscher, E.F. "Växelkurens Grundval vid Pappersmyntot." *Economisk Tidskrift*, 18(1916): 309–312.

- Hendry, D.F. and K. Juselius. "Explaining Cointegration Analysis: Part I." *The Energy Journal*, 21(2000): 1–42.
- Hendry, D.F. and K. Juselius. "Explaining Cointegration Analysis: Part II." *The Energy Journal*, 22(2001): 75–120.
- Howitt, P. and R.P. McAfee. "Animal spirits." *The American Economic Review*, 82(1992): 493–507.
- Jackman, S. "Re-Thinking Equilibrium Presidential Approval - Markov-Switching Error Correction." Paper presented at the *12th Annual Political Methodology Summer Conference*, Indiana University, Bloomington, USA, 1995.
- Jacks, D.S. "Intra- and International Commodity Market Integration in the Atlantic Economy, 1800–1913." *Explorations in Economic History*, 42(2005): 381–413.
- Jacks, D.S. "What Drove 19th Century Commodity Market Integration?" *Explorations in Economic History*, 43(2006): 383–412.
- Jeanne, O. and P. Masson. "Currency Crises, Sunspots and Markov-Switching Regimes." *Journal of International Economics*, 50(2000): 327–350.
- Johansen, S. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford, UK, 1995.
- Keynes, J.M. *The General Theory of Employment, Interest, and Money*. Macmillan, London, UK, 1936.
- Kim, C.-J. "Dynamic Linear Models with Markov-Switching." *Journal of Econometrics*, 60(1994): 1–22.
- Krolzig, H.-M. "Statistical Analysis of Cointegrated VAR Processes with Markovian Regime Shifts." SFB 373 Discussion Paper 25/1996, Humboldt-Universität zu Berlin, Berlin, Germany, 1996.
- Krolzig, H.-M. *Markov-Switching Vector Autoregressions. Modelling, Statistical Inference, and Applications to Business Cycle Analysis*. Springer, Berlin, Germany, 1997.
- Krolzig, H.-M. *MSVAR - An Ox Package Designed for the Econometric Modelling of Univariate and Multiple Time Series Subject to Shifts in Regime*, Version 1.31k. URL <http://www.economics.ox.ac.uk/research/hendry/krolzig/msvar.html>, 2004.
- Krolzig, H.-M. and J. Toro. "A New Approach to the Analysis of Business Cycle Transitions in a Model of Output and Employment." Department of economics discussion paper series No. 59, University of Oxford, Oxford, 2001.
- Krolzig, H.-M., M. Marcellino, and G. Mizon. "A Markov-Switching Vector Equilibrium Correction Model of the UK Labor Market." *Empirical Economics*, 27(2002): 233–254.
- Lo, M.C. and E. Zivot. "Threshold Cointegration and Nonlinear Adjustment to the Law of One Price." *Macroeconomic Dynamics*, 5(2001): 533–576.

- Luoma, A., J. Luoto, and M. Taipale. "Threshold Cointegration and Asymmetric Price Transmission in Finnish Beef and Pork Markets." Pellervo Economic Research Institute Working Papers No. 70, Helsinki, Finland, 2004.
- Lutz, C., W.E. Kuiper, and A. van Tilburg. "Maize Market Liberalisation in Benin: A Case of Hysteresis." *Journal of African Economies*, 16(2006): 102–133.
- MacDonald, I.L. and W. Zucchini. *Hidden Markov and Other Models for Discrete-valued Time Series*. Chapman and Hall, London, UK, 1997.
- Marshall A.. *Principles of Economics*. Macmillan Company, 8th ed., New York, USA, 1890.
- Martinez Peria, M.S. "A Regime-Switching Approach to the Study of Speculative Attacks: A Focus on EMS Crises", in J.D. Hamilton and B. Raj (eds), *Advances in Markov-Switching Models. Applications in Business Cycle Research and Finance*, Physica-Verlag, Heidelberg, Germany, 2002.
- McNew, K.P. and P.L. Fackler. "Testing Market Equilibrium: Is Cointegration Informative?" *Journal of Agricultural and Resource Economics*, 22(1997): 191–207.
- Mellows, M. *Testen und Auswählen von nichtlinearen Zeitreihenmodellen mit dem Bootstrap-Verfahren*. Peter Lang Europäischer Verlag der Wissenschaften, Frankfurt am Main, Germany, 1999.
- Meyer, J. "Measuring Market Integration in the Presence of Transaction Costs — a Threshold Vector Error Correction Approach." *Agricultural Economics*, 31(2004): 327–334.
- Mizrach, B. and J. Watkins. "A Markov Switching Cookbook", in P. Rothman (ed), *Nonlinear Time Series Analysis of Economic and Financial Data*, Kluwer Academic Publishers, Boston, USA, 2000, chapter 2.
- Moschini, G. and K.D. Meilke. "Modelling the Pattern of Structural Change in U.S. Meat Demand". *American Journal of Agricultural Economics*, 71(1989): 253–261.
- Noack, T. *Probleme der SETAR-Modellierung in der Zeitreihenanalyse*. Logos Verlag, Berlin, Germany, 2003.
- Norman, S. "How Well does Nonlinear Mean Reversion Solve the PPP Puzzle?" Working Paper University of Washington, Tacoma, USA, 2007.
- O'Connel, P.G.J. and S.-J. Wei. "The Bigger They Are, the Harder They Fall: How Price Differences Across U.S. Cities Are Arbitraged." Working paper 6089, National Bureau of Economic Research, Cambridge, USA, 1997.
- O'Connel, P.G.J. and S.-J. Wei. "'The Bigger They Are, the Harder They Fall': Retail Price Differences Across U.S. Cities". *Journal of International Economics*, 56(2002): 21–53.
- Obstfeld, M. and A.M. Taylor. "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited." *Journal of the Japanese and International Economies*, 11(1997): 441–479.

- Park, H., J.W. Mjelde, and D.A. Bessler. "Time-Varying Threshold Cointegration and the Law of One Price." *Applied Economics*, 39(2007): 1091–1105.
- Pede, V.O. and A.M. McKenzie. "Integration in Benin Maize Market: An Application of Threshold Cointegration Analysis." Selected Paper prepared for presentation at the *American Agricultural Economics Association Annual Meeting*, Providence, Rhode Island, USA, July 24-27 2005.
- Prakash, G. and A.M. Taylor. "Measuring Market Integration: A Model of Arbitrage with an Econometric Application to the Gold Standard, 1879–1913." Working Paper No. 6073, National Bureau of Economic Research, USA, 1997.
- Priestley, M.B. (1980a). Discussion of Tong and Lim (1980), 1980, p. 273.
- Priestley, M.B. (1980b). "State-Dependent Models: A General Approach to Non-linear Time Series Analysis." *Journal of Time Series Analysis*, 1(1980): 57–71.
- Priestley, M.B. *Nonlinear and Non-Stationary Time Series Analysis*. Academic Press, London, UK, 1988.
- Psaradakis, Z., M. Sola, and F. Spagnolo. "On Markov Error-Correction Models." Working Paper, Birkbeck College, London, UK, 2001.
- Psaradakis, Z., Sola, M. and F. Spangolo (2004a). "On Markov-Switching Models, with an Application to Stock Prices and Dividends." *Journal of Applied Econometrics*, 19(2004): 69–88.
- Psaradakis, Z., Sola, M. and F. Spangolo (2004b). Appendix of "On Markov-Switching Models, with an Application to Stock Prices and Dividends". *Journal of Applied Econometrics Data Archive*, URL <http://qed.econ.queensu.ca/jae/datasets/psaradakis002/pss-appendix.pdf>, 2004.
- R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>, 2007.
- Raj, B. "Asymmetry of Business Cycles: The Markov-Switching Approach", in A. Ullah, A.T.K. Wan and A. Chaturvedi (eds), *Handbook of Applied Econometrics and Statistical Inference*, Dekker, New York, USA, 2002.
- Samuelson, P. "Spatial Price Equilibrium and Linear Programming." *American Economic Review*, 42(1952): 283–303.
- Sephton, P.S. "Spatial Market Arbitrage and Threshold Cointegration." *American Journal of Agricultural Economics*, 85(2003): 435–450.
- Sercu, P., R. Uppal, and C. Van Hulle. "The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity." *Journal of Finance*, 10(1995): 1309—1319.

- Serra, T. and B.K. Goodwin. "Specification Selection Issues in Multivariate Threshold and Switching Models.", in *Proceedings of the AAEA-WAEA Annual Meeting*, 2002.
- Serra, T. and B.K. Goodwin. "Price transmission and asymmetric adjustment in the Spanish dairy sector." *Applied Economics*, 35(2003): 1889–1899.
- Serra, T. and B.K. Goodwin. "Regional Integration of Nineteenth Century U.S. Egg Markets." *Journal of Agricultural Economics*, 55(2004): 59–74.
- Serra, T., J.M. Gil, and B.K. Goodwin. "Local Polynomial Fitting and Spatial Price Relationships: Price Transmission in EU Pork Markets." *European Review of Agricultural Economics*, 33(2006): 415–436.
- Serra, T., B.K. Goodwin, J.M. Gil, and A. Mancuso. "Non-parametric Modelling of Spatial Price Relationships." *Journal of Agricultural Economics*, 57(2006): 501–521.
- Shepherd, A.W. *Market Information Services. Theory and Practice*. FAO Agricultural Services Bulletin 125, Rome, Italy, 1997.
- Siklos, P.L. and C.W.J. Granger. "Regime-Sensitive Cointegration with an Application to Interest-rate Parity." *Macroeconomic Dynamics*, 1(1997): 640–657.
- Takayama, T. and G. Judge. *Spatial and Temporal Price Allocation Models*. North-Holland, Amsterdam, The Netherlands, 1971.
- Thomas, J.K. "Do Sunspots Produce Business Cycles?" Working paper, University of Minnesota, Minneapolis, USA, 2004.
- Tjøstheim, D. "Some Doubly Stochastic Time Series Models." *Journal of Time Series Analysis*, 7: 225–273.
- Tong, H. "On a Threshold Model", in C.H. Chen (ed), *Pattern recognition and signal processing*, Sijthoff and Noordhoff, Amsterdam, The Netherlands, 1978.
- Tong, H. *Threshold Models and Non-linear Time Series Analysis*. Springer-Verlag, New York, USA, 1983.
- Tong, H. *Non-linear Time Series. A Dynamical System Approach*. Clarendon Press, Oxford, UK, 1990.
- Tong, H. "Birth of the Threshold Time Series Model." *Statistica Sinica*, 17(2007): 8–14.
- Tong, H. and K.S. Lim. "Threshold Autoregression, Limit Cycles and Cyclical Data (with Discussion)." *Journal of the Royal Statistical Society*, B42(1980): 245–292.
- Trenkler, C. and N. Wolf. "Economic Integration in Interwar Poland - A Threshold Cointegration Analysis of the Law of One Price for Poland (1924–1937)." European University Institute Working Paper ECO, No. 2003/5, San Domenico, Italy, 2003.
- Trenkler, C. and N. Wolf. "Economic Integration Across Borders: The Polish Interwar Economy 1921–1937." *European Review of Economic History*, 9(2005): 199 – 231.

- Tsay, R.S. "Testing and Modeling Threshold Autoregressive Processes." *Journal of the American Statistical Association*, 84(1989): 231–240.
- Tsay, R.S. "Testing and Modeling Multivariate Threshold Models." *Journal of the American Statistical Association*, 93(1998): 1188–1202.
- Uchezuba, D.I. "Measuring Market Integration for Apples on the South African Fresh Produce Market: A Threshold Error Correction Model." Master thesis, University of the Free State Bloemfontein, South Africa, 2005.
- Uppal, R. "A General Equilibrium Model of International Portfolio Choice." *Journal of Finance*, 48(1993): 529–553.
- Van Campenhout, B. "Modelling Trends in Food Market Integration: Method and an Application to Tanzanian Maize Markets." *Food Policy*, 32(2007): 112–127.
- Whittle, P. "The Statistical Analysis of a Seiche Record." *Journal of Marine Research*, 13(1954): 76–100.

## Appendix A: Review of Applications of the TVECM

Table 4 provides a review of studies of PT in commodity markets which are in most cases applications of the TVECM. We focus in the review on the data, models and estimation approaches used in the studies. It covers 35 publications of which 26 are journal articles. Tables 1 to 3 summarize the reviewed publications according to their publication type, research field and publication year. Table 4 lists the publications according to the initial letter

Table 1: Publication per Type

Type	Journal article	Conference paper	Working paper	Dissertation	MSc thesis
Number	26	4	3	1	1

of the first author's name. The first two columns give information on the type  $Ty$  of the article, where  $j$  denotes "journal article",  $c$  "conference paper",  $w$  "working paper",  $d$  "dissertation" and  $m$  "MSc thesis", and the field of research  $Fi$ , where  $ae$  denotes "agricultural economics",  $eh$  "economic history",  $ee$  "energy economics" and  $e$  "economics" in general. The following four columns provide information on the analyzed data. They outline the product(s) and the geographic region studied, their frequency  $Fr$ , where  $q$  denotes "quarterly",  $m$  "monthly",  $b$  "bi-weekly",  $w$  "weekly",  $d$  "daily" and  $n.m.$  "not mentioned", and the length of the time series  $T$  in number of periods. The next five columns of Table 4 summarize features of the

Table 2: Publications per Research Field

Type	Agricultural economics	Economic history	Economics	Energy economics
Number	23	7	3	2

estimated model(s). The column  $Model$  classifies the used model(s) according to the discussion of equation (11), p.9,  $Reg$  denotes the number of regimes in the model;  $Cont$  indicates whether the model, in case it has three regimes and is of the Band-type, is continuous.  $Sym$  indicates, in case the model has three regimes, whether the model is symmetric and  $Adj$  states whether the model allows for a nonzero adjustment coefficient in the middle regime, i.e., whether  $\alpha^{(2)} \neq 0$ . The last two columns provide information about

Table 3: Publications per Year

Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Number	1	2	1	2	2	3	2	3	7	5	7

the applied estimation method.  $Est$  outlines the (reference of) the estimation method applied where  $HS$  denotes the method of Hansen and Seo (2002),  $B$  denotes "Bayesian estimation",  $LZ$ ,  $LS$ ,  $PT$ ,  $OT$ ,  $BF$  the method of Lo and Zivot (2001), "multivariate least squares", Prakash and Taylor (1997), Obstfeld and Taylor (1997) and Balke and Fomby (1997) respectively.  $Par$  denotes the parameters used in the optimization as defined in (6).

Publication	Ty	Fi	Product	Region	Fr	T	Model	Reg	Cont	Sym	Adj	Est	Par
Agüero (2007)	w	ae	Tomatoes, potatoes, rice	Peru	d	1850	TVECM	2	-	-	-	HS	$\beta, \theta^{(1)}$
Alemu and Biacuana (2006)	c	ae	Maize	Mozambique	m	144	EQ-TVAR	3	-	no	yes	HS	$\theta^{(1)}, \theta^{(2)}$
Bakucs and Fertö (2007)	c	ae	Milk	Hungary	w	105	TVECM	2	-	-	-	HS	$\beta, \theta^{(1)}$
Balcombe et al. (2007)	j	ae	Wheat, soya, maize	Argentina, Brazil, USA	m	157	EQ-TVECM	3	-	yes	yes	B	-
Ben-Kaabia and Gil (2007)	j	ae	Lamb	Spain	w	365	Band-TVECM	3	no	yes	yes	B	-
Ben-Kaabia et al. (2005)	j	e	Poultry	Spain	m	264	Band-TVECM	3	no	no	no	LZ	$\theta^{(1)}, \theta^{(2)}$
Ben-Kaabia et al. (2002)	c	ae	Lamb	Spain	w	365	TVECM	2	-	-	-	HS	$\beta, \theta^{(1)}$
Chen et al. (2005)	j	ee	Crude oil, gasoline	USA	w	535	ECM	2	-	-	-	LS	-
Dercon and van Campenhout (1998)	w	ae	Rice	Philippines	m	162	Band-TAR	3	yes	yes	yes	n.m.	n.m.
Ejmas and Persson (2000)	j	eh	Wheat	France	b	1100	Band-TVECM	3	yes	no	no	PT	$\theta^{(1)}, \theta^{(2)}$
Escobar (2005)	d	ae	Potatoes	Peru	d	1540	Band-TAR	3	yes	yes	yes	n.m.	n.m.
Federico (2007)	j	eh	Wheat	Italy	w	260	Band-TAR	3	yes	yes	no	LZ	n.m.
Goodwin and Grennes (1998)	j	eh	Wheat	Russia	m	263	Band-TAR	3	yes	yes	no	OT	$\theta^{(1)}$
Goodwin and Harper (2000)	j	eh	Butter	USA	m	373	EQ-TVECM	3	-	no	yes	BF	$\theta^{(1)}, \theta^{(2)}$
Goodwin and Holt (1999)	j	ae	Pork	USA	w	626	EQ-TVECM	3	-	no	n.m.	BF	$\theta^{(1)}, \theta^{(2)}$
Goodwin and Piggott (2001)	j	ae	Beef	USA	w	897	EQ-TVECM	3	-	no	n.m.	BF	$\theta^{(1)}, \theta^{(2)}$
Goodwin et al. (2002)	j	ae	Corn, soybean	USA	d	2645	EQ-TVECM	3	-	no	n.m.	BF	$\theta^{(1)}, \theta^{(2)}$
Jacks (2005)	j	eh	Wheat	Atlantic economy	m	1300	Band-TVECM	3	yes	no	no	n.m.	$\theta^{(1)}, \theta^{(2)}$

Table 4: Studies of Price Transmission in Commodity Markets



Publication	Ty	Fi	Product	Region	Fr	T	Model	Reg	Cont	Sym	Adj	Est	Par
Jacks (2006)	j	eh	Wheat	Atlantic economy	m	1300	Band-TVECM	3	yes	no	no	n.m.	$\theta^{(1)}, \theta^{(2)}$
Lo and Zivot (2001)	j	e	CPI	USA	m	115	Band-TVECM	3	yes	yes	no	LZ	$\theta^{(1)}, \theta^{(2)}$
Luoma et al. (2004)	w	ae	Beef, pork	Finland	m	269	TVECM	2/3	no/yes	-/no	-/yes	B	-
Lutz et al. (2006)	j	ae	Maize	Benin	n.m.	422	Band-TVECM	3	yes	yes	no	LS	-
Meyer (2004)	j	ae	Pork	Germany, Netherlands	w	600	Band-TVECM	3	no	yes	yes	HS	$\theta^{(1)}$
Obstfeld and Taylor (1997)	j	e	CPI	World	m	192	Band-TVECM	3	yes	yes	no	OT	$\theta^{(1)}$
O'Connell and Wei (2002)	j	e	CPI	USA	q	42-68	TAR	2	-	-	-	n.m.	$\theta^{(1)}, d$
Park et al. (2007)	j	ee	Natural gas	Canada, USA	d	1290	Band-TVECM	3	no	no	yes	LZ	$\theta^{(1)}, \theta^{(2)}, d$
Pede and McKenzie (2005)	c	ae	Maize	Benin	w	162	TVECM	2	-	-	-	HS	$\beta, \theta^{(1)}$
Sephton (2003)	j	ae	Corn, soybean	USA	d	2645	TVECM	2	-	-	-	HS	$\beta, \theta^{(1)}$
Serra and Goodwin (2003)	j	ae	Dairy	Spain	m	78	EQ-TVECM	3	-	no	yes	LS	$\theta^{(1)}, \theta^{(2)}$
Serra and Goodwin (2004)	j	ae	Eggs	USA	m	372	TVECM	2/3	-/n.m.	-/n.m.	-/n.m.	HS	$\theta^{(1)} / \theta^{(1)}, \theta^{(2)}$
Serra, Gil et al. (2006)	j	ae	Pork	Denmark, France, Germany, Spain	w	574	EQ-TAR	3	-	no	yes	BF	$\theta^{(1)}, \theta^{(2)}$
Serra, Goodwin et al. (2006)	j	ae	Eggs	USA	m	360	EQ-TAR	3	-	no	yes	n.m.	$\theta^{(1)}, \theta^{(2)}$
Trenkler and Wolf (2005)	j	eh	Wheat flour	Poland	m	190	Band-TAR	3	yes	yes	no	LZ	n.m.
Uchezuba (2005)	m	ae	Apples	South Africa	m	148	Band-TVECM	3	no	no	yes	LZ	$\theta^{(1)}, \theta^{(2)}$
Van Campenhout (2007)	j	ae	Maize	Tanzania	w	626	EQ-TAR	3	-	yes	no	n.m.	$\theta^{(1)}, \theta^{(2)}$

Table 4: Studies of Price Transmission in Commodity Markets

## Appendix B: Simulation Study

We perform a simulation study in order to assess the performance of the estimation methods almost exclusively used in applied research which are sequential conditional least squares (SCLS) in case of the TVECM and the Expectation-Maximization-Algorithm (EMA) in case of the MSVECM. We generate 1000 replications of two prices  $p_t^A$  and  $p_t^B$  per dataset which follow a certain TVECM-specification and MSVECM-specification respectively. Each DGP is assumed to consist of  $l + 1 = M = 3$  regimes in one of which no error correction takes place, i.e. where the two prices are not cointegrated.<sup>41</sup> Since the equilibrium errors are corrected toward the long-run relationship itself and not toward a band around it in case of the MSVECM, we assume the error correction process of the TVECM to be of equilibrium-type in order to ensure comparability of the adjustment processes, i.e., the equilibrium errors of the threshold model are also assumed to be corrected toward the long-run relationship itself which means that we assume  $\mu^{(j)} = 0$  in (11), p. 9. Furthermore, we assume the short-run dynamics  $\Psi_i^{(j)} = 0$  for the sake of simplicity. Hence, the DGP corresponds to the simple nonlinear VECM as outlined in Balke and Fomby (1997, p. 629):

$$\begin{aligned} \Delta \mathbf{p}_t &= \boldsymbol{\alpha}^{(J_t)} ect_{t-1} + \boldsymbol{\epsilon}_t & (B.1) \\ &= \begin{pmatrix} \Delta p_t^A \\ \Delta p_t^B \end{pmatrix} = \begin{pmatrix} \alpha^{A(J_t)} \\ \alpha^{B(J_t)} \end{pmatrix} ect_{t-1} + \begin{pmatrix} \epsilon_t^A \\ \epsilon_t^B \end{pmatrix}. \end{aligned}$$

Each dataset is generated containing  $t = 1, \dots, T$ ;  $T = 150, 500, 1500$  time points respectively.<sup>42</sup> The lengths  $T$  are based on the time series used in the studies of the literature review in Table 4. The first two values of  $T$  denote a short and a long time series as typically used in empirical research. They roughly correspond to the first (162 measurements) and the third quartile (564 measurements) of the lengths of the time series of all studies except the five publications which use daily data. Very long time series of  $T = 1500$  observations will typically rarely be available in PT analysis. They correspond roughly to the mean length (1560 measurements) of the time series of daily observations used in Escobal (2005), Park et al. (2007) and Agüero (2007); the datasets of Goodwin et al. (2002) and Sephton (2003) with 2645 daily observations are not regarded since they are exceptionally long.

In particular, we assume the parameter of the cointegration relationship and the common trend to  $\beta = -1$  and  $\phi = 1$  respectively. We generate the equilibrium error process  $\{ect_t\}$  and the common stochastic trend  $\{B_t\}$  according to

$$ect_t = \rho^{(J_t)} \cdot ect_{t-1} + \nu_t \quad \text{where} \quad ect_t = p_t^B + \beta p_t^A \quad \text{and} \quad \nu_t \stackrel{iid}{\sim} N(0, 1) \quad (B.2)$$

$$B_t = B_{t-1} + \eta_t \quad \text{where} \quad B_t = p_t^B + \phi p_t^A \quad \text{and} \quad \eta_t \stackrel{iid}{\sim} N(0, 1) \quad (B.3)$$

The corresponding error correction parameters  $\boldsymbol{\alpha}^{(j)}$  and the price series  $p_t^A, p_t^B$  in (B.1) may

<sup>41</sup> In case of the TVECM, it is the middle regime that does not show error correction, i.e.,  $\boldsymbol{\alpha}^{(2)} = 0$ ; for the MSVECM it is the first, i.e.,  $\boldsymbol{\alpha}^{(1)} = 0$ .

<sup>42</sup> We actually generate  $T + 200$  observations of each dataset and throw away the first 200 in order to reduce the influence of the starting value.

then be calculated based on the generated values of  $ect_t$  and  $B_t$  in via the identities

$$p_t^A = \frac{B_t - ect_t}{\phi - \beta} \quad p_t^B = \frac{\phi \cdot ect_t - \beta B_t}{\phi - \beta}$$

$$\alpha^{A(j)} = \frac{1 - \rho^{(j)}}{\phi - \beta} \quad \alpha^{B(j)} = -\frac{\phi(1 - \rho^{(j)})}{\phi - \beta}.$$

We generate datasets for the four cases in Table 6, p. 43. In cases I and IV, we are interested in the performance of the estimation methods depending on the true parameters introducing nonlinearity, i.e., the thresholds and the transition matrix.<sup>43</sup> Hence, we focus on the estimation of the thresholds and the ergodic probabilities in dependence on a set of varying true thresholds and ergodic probabilities incorporating the error correction parameter as a nuisance parameter in order to study if and to what extent they influence the estimation results.<sup>44</sup> The parameters varied in the DGPs of the TVECM and the MSVECM are the thresholds  $\theta^{(1)}, \theta^{(2)}$  and the error correction parameters  $\rho^{(1)}, \rho^{(3)}$  of the outer regimes and the ergodic probabilities  $\pi^{(2)}, \pi^{(3)}$  and the error correction parameters  $\rho^{(1)}, \rho^{(3)}$  respectively.<sup>45</sup>

As criterion measuring the performance of the estimation of the thresholds via SCLS and the ergodic probabilities via EMA we choose the mean squared error

$MSE_{\hat{\lambda}} = E(\hat{\lambda} - \lambda)^2 = [\text{Bias}(\hat{\lambda})]^2 + \sigma^2(\hat{\lambda})$ . Hence, the MSE of SCLS is estimated for

$\hat{\lambda} \in \{\hat{\theta}^{(1)}, \hat{\theta}^{(2)}\}$  and the MSE of the EMA for  $\hat{\lambda} \in \{\hat{\pi}^{(1)}, \hat{\pi}^{(2)}, \hat{\pi}^{(3)}\}$ . We approximate the functional relationship between the MSE and the varying true parameters of the DGPs by

$\widehat{MSE}_{\hat{\lambda}} = [\widehat{\text{Bias}}(\hat{\lambda})]^2 + \widehat{\sigma^2}(\hat{\lambda})$ . The evaluation of  $MSE_{\hat{\lambda}}^{SCLS} = f(\theta^{(1)}, \theta^{(2)}, \rho^{(1)}, \rho^{(3)})$  and

$MSE_{\hat{\lambda}}^{EMA} = f(\pi^{(2)}, \pi^{(3)}, \rho^{(1)}, \rho^{(3)})$  for grids of true parameters is first computationally very demanding and can hardly graphically be illustrated due to its five dimensions. We thus

reduce the number of dimensions to three by evaluating the MSE function by holding its third and fourth parameters constant  $MSE_{\hat{\lambda}}^A = f(\theta^{(1)}, \theta^{(2)} | \rho^{(1)}, \rho^{(3)})$ ,

$MSE_{\hat{\lambda}}^C = f(\pi^{(2)}, \pi^{(3)} | \rho^{(1)}, \rho^{(3)})$  and evaluating it by holding its first and second parameters constant ( $MSE_{\hat{\lambda}}^B = f(\rho^{(1)}, \rho^{(3)} | \theta^{(1)}, \theta^{(2)})$ ,  $MSE_{\hat{\lambda}}^D = f(\rho^{(1)}, \rho^{(3)} | \pi^{(2)}, \pi^{(3)})$ ) respectively.

Despite we restrict the simulation on these two subspaces, detailed insights into the behavior of the estimation methods can be expected.

<sup>43</sup> Since the transition matrix of the MSVECM-DGP considered contains nine elements, we use the three corresponding ergodic probabilities  $\pi = (\pi^{(1)} \pi^{(2)} \pi^{(3)})^\top$  instead. These are, technically spoken, the normalized eigenvector of the matrix associated with its unit eigenvalue (cf. Hamilton, 1994, pp. 681). The transition matrix and the ergodic probabilities are connected via  $\Gamma = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$  (Hamilton, 1994, p. 730) where  $\mathbf{T}$  is the  $M \times M$  matrix of the eigenvectors and  $\mathbf{\Lambda}$  the  $M \times M$  diagonal matrix of the eigenvalues of  $\Gamma$ . Hence, by holding all elements of  $\mathbf{\Lambda}$  constant and only plugging in varying  $\pi$  in the column of  $\mathbf{T}$  associated with the unit eigenvalue, the corresponding transition matrices can be obtained. The ergodic probabilities corresponding to an estimated transition matrix  $\hat{\Gamma}$  can be obtained, in turn, by the eigenvalue decomposition of the latter and normalization of the

respective eigenvector. We obtained  $\mathbf{\Lambda}$  and  $\mathbf{T}$  from the eigenvalue decomposition of  $\Gamma = \begin{pmatrix} 0.95 & 0.03 & 0.02 \\ 0.15 & 0.8 & 0.05 \\ 0.2 & 0.1 & 0.7 \end{pmatrix}$

since it implies ergodic probabilities of  $\pi = (0.769 \ 0.154 \ 0.077)^\top$ , i.e., it assigns quite distinct unconditional probabilities to the three regimes.

<sup>44</sup> Since we simulate according to (B.2) and (B.3), we vary the autoregressive parameter  $\rho^{(j)}$  in (B.2) instead of  $\alpha^{(j)}$ .

<sup>45</sup> Note that  $\alpha^{(2)} = 0$  implying that  $\rho^{(2)} = 1$  in both cases and  $\pi^{(1)} = 1 - (\pi^{(2)} + \pi^{(3)})$ . As mentioned before, the ergodic probabilities denote the expected unconditional probabilities  $\pi^{(j)}$  of the  $j = 1, \dots, M$ ;  $M = 3$  regimes. The frequencies of the regimes generated by a MSVECM-DGP should hence equal them asymptotically.

In cases II and III it is not possible to estimate the bias by simulation since the respective method estimates thresholds and the transition matrix although the data was generated by a MSVECM and a TVECM respectively. We therefore try to assess for each observation whether its regime was correctly classified by the estimation method. However, this kind of inference has two identification problems in case of the EMA. First, the method only allows probabilistic statements regarding the regime incidences. The approach of considering the regime with the highest smoothed probability at some time  $t$  as the one that most likely occurred may, of course, randomly lead to incorrect identification. Second, Markov-switching models in general suffer from the problem regime identification regarding the transition matrix. This means that for repeated estimations the numbering of the regimes may change so that the first regime of the first estimation needs not to be identical to the first one of the second estimation. One can try to identify the regimes by the magnitudes of its estimated parameters, however, a considerable amount of uncertainty remains since the estimates vary randomly so that one can indeed obtain a wrong reordering.

### Case I

First, we generate TVECM datasets of the 16 combinations of  $\theta^{(1)} = -0.5, -1, -2, -3$  and  $\theta^{(2)} = 0.5, 1, 2, 3$  so that

$$\text{MSE}_{\hat{\lambda}}^A = f(M^A | \rho^{(1)} = 0.9, \rho^{(3)} = 0.9) \quad (\text{B.4})$$

where

$$M^A = \begin{pmatrix} -0.5, 0.5 & -1, 0.5 & -2, 0.5 & -3, 0.5 \\ -0.5, 1 & -1, 1 & -2, 1 & -3, 1 \\ -0.5, 2 & -1, 2 & -2, 2 & -3, 2 \\ -0.5, 3 & -1, 3 & -2, 3 & -3, 3 \end{pmatrix} \quad (\text{B.5})$$

and the MSE is evaluated for each combination, i.e., for each element of  $M^A$  individually.

Second, we generate TVECM datasets of the 16 combinations of  $\rho^{(1)} = 0.98, 0.9, 0.8, 0.5$  and  $\rho^{(3)} = 0.98, 0.9, 0.8, 0.5$  so that

$$\text{MSE}_{\hat{\lambda}}^B = f(M^B | \theta^{(1)} = -1, \theta^{(2)} = 1) \quad (\text{B.6})$$

where

$$M^B = \begin{pmatrix} 0.98, 0.98 & 0.9, 0.98 & 0.8, 0.98 & 0.5, 0.98 \\ 0.98, 0.9 & 0.9, 0.9 & 0.8, 0.9 & 0.5, 0.9 \\ 0.98, 0.8 & 0.9, 0.8 & 0.8, 0.8 & 0.5, 0.8 \\ 0.98, 0.5 & 0.9, 0.5 & 0.8, 0.5 & 0.5, 0.5 \end{pmatrix}. \quad (\text{B.7})$$

For  $T = 1500$ , we only generate 9 combinations of  $\theta^{(1)} = -0.5, -1, -2$ ,  $\theta^{(2)} = 0.5, 1, 2$  and  $\rho^{(1)} = 0.98, 0.9, 0.5$ ,  $\rho^{(3)} = 0.98, 0.9, 0.5$ , respectively, due to the high computational cost involved for such long time series.

*Case II*

We generate one MSVECM dataset with  $\Gamma = \begin{pmatrix} 0.8 & 0.15 & 0.05 \\ 0.03 & 0.95 & 0.02 \\ 0.1 & 0.2 & 0.7 \end{pmatrix}$  and  $\rho^{(1)} = 0.9$ ,  
 $\rho^{(3)} = 0.5$ .

*Case III*

We generate one TVECM dataset with  $\theta^{(1)} = -1$  and  $\theta^{(2)} = 1$  and  $\rho^{(1)} = 0.9$ ,  $\rho^{(3)} = 0.5$ .

*Case IV*

First, we generate MSVECM datasets of the 16 combinations of  $\pi^{(2)} = \frac{1}{12}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}$  and  
 $\pi^{(3)} = \frac{1}{12}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}$  so that

$$\text{MSE}_{\lambda}^C = f(M^C | \rho^{(2)} = 0.9, \rho^{(3)} = 0.5). \quad (\text{B.8})$$

where

$$M^C = \begin{pmatrix} 0.08, 0.08 & 0.17, 0.08 & 0.25, 0.08 & 0.33, 0.08 \\ 0.08, 0.17 & 0.17, 0.17 & 0.25, 0.17 & 0.33, 0.17 \\ 0.08, 0.25 & 0.17, 0.25 & 0.25, 0.25 & 0.33, 0.25 \\ 0.08, 0.33 & 0.17, 0.33 & 0.25, 0.33 & 0.33, 0.33 \end{pmatrix}. \quad (\text{B.9})$$

Second, we generated MSVECM datasets of the 16 combinations of  $\rho^{(2)} = 0.98, 0.9, 0.8, 0.5$   
and  $\rho^{(3)} = 0.98, 0.9, 0.8, 0.5$  so that

$$\text{MSE}_{\lambda}^D = f(M^D | \pi^{(2)} = \frac{1}{6}, \pi^{(3)} = \frac{1}{12}) \quad (\text{B.10})$$

where

$$M^D = \begin{pmatrix} 0.98, 0.98 & 0.9, 0.98 & 0.8, 0.98 & 0.5, 0.98 \\ 0.98, 0.9 & 0.9, 0.9 & 0.8, 0.9 & 0.5, 0.9 \\ 0.98, 0.8 & 0.9, 0.8 & 0.8, 0.8 & 0.5, 0.8 \\ 0.98, 0.5 & 0.9, 0.5 & 0.8, 0.5 & 0.5, 0.5 \end{pmatrix}. \quad (\text{B.11})$$

For  $T = 1500$ , we only generate 9 combinations of  $\pi^{(2)} = \frac{1}{12}, \frac{1}{6}, \frac{1}{3}, \pi^{(3)} = \frac{1}{12}, \frac{1}{6}, \frac{1}{3}$  and  
 $\rho^{(2)} = 0.98, 0.9, 0.5, \rho^{(3)} = 0.98, 0.9, 0.5$ , respectively, due to the high computational cost  
involved for such long time series. We do not present all results in this paper. Further results  
may be obtained from the authors.

Table 5: Estimation Approaches of Selected Publications<sup>46</sup>

Publication	Model	$l$	$d$	Principle	Criterion	Par's	Optim.
Balke and Fomby (1997)	SETAR	any	est.	LS	RSS	$\theta, d$	Min
Obstfeld and Taylor (1997)	SETAR	1 <sup>47</sup>	1	ML	LR <sup>48</sup>	$\theta$	Max
Hansen (1999)	SETAR	0-2	est.	LS	RSS	$\theta, d$	Min
Lo and Zivot (2001)	TVECM	1	est.	LS	RSS	$\theta, d$	Min
Hansen and Seo (2002)	TVECM	1	1	ML	$\log  \Sigma $	$\theta, \beta$	Min

Table 6: Design of the Simulation Study

Estimation	DGP	
	TVECM	MSVECM
SCLS	I	II
EMA	III	IV

Table 7: True vs. Estimated Regime Incidences (Case II)

	R1 true	R2 true	R3 true
R1 est	0.339	0.076	0.041
R2 est	0.216	0.039	0.018
R3 est	0.214	0.039	0.019

Table 8: True vs. Estimated Regime Incidences (Case III)

	R1 true	R2 true	R3 true
R1 est	0.088	0.105	0.028
R2 est	0.133	0.152	0.044
R3 est	0.186	0.199	0.065

<sup>46</sup> Obstfeld and Taylor estimate a symmetric TVECM(3), i.e. in absolute terms, only one symmetric threshold is to be estimated.

<sup>47</sup>  $LR$  denotes the likelihood ratio between a SETAR(3,1,1,1) and an AR(1). The latter model might also be called a SETAR (1,1), for the notation of SETAR models see footnote 6, p. 4.

<sup>48</sup>  $l$  denotes the number of thresholds (compare footnote 18, page 8),  $d$  the delay parameter,  $ML$  maximum likelihood,  $LS$  least squares,  $RSS = \text{trace}(\Sigma)$  residual sum of squares,  $\log |\Sigma|$  the log determinant of the variance-covariance matrix of the residuals;  $\theta$  means threshold, *est.* estimated and  $\beta$  the cointegration vector.

Figure 1: Classification of Nonlinear Models after Tong (1990)

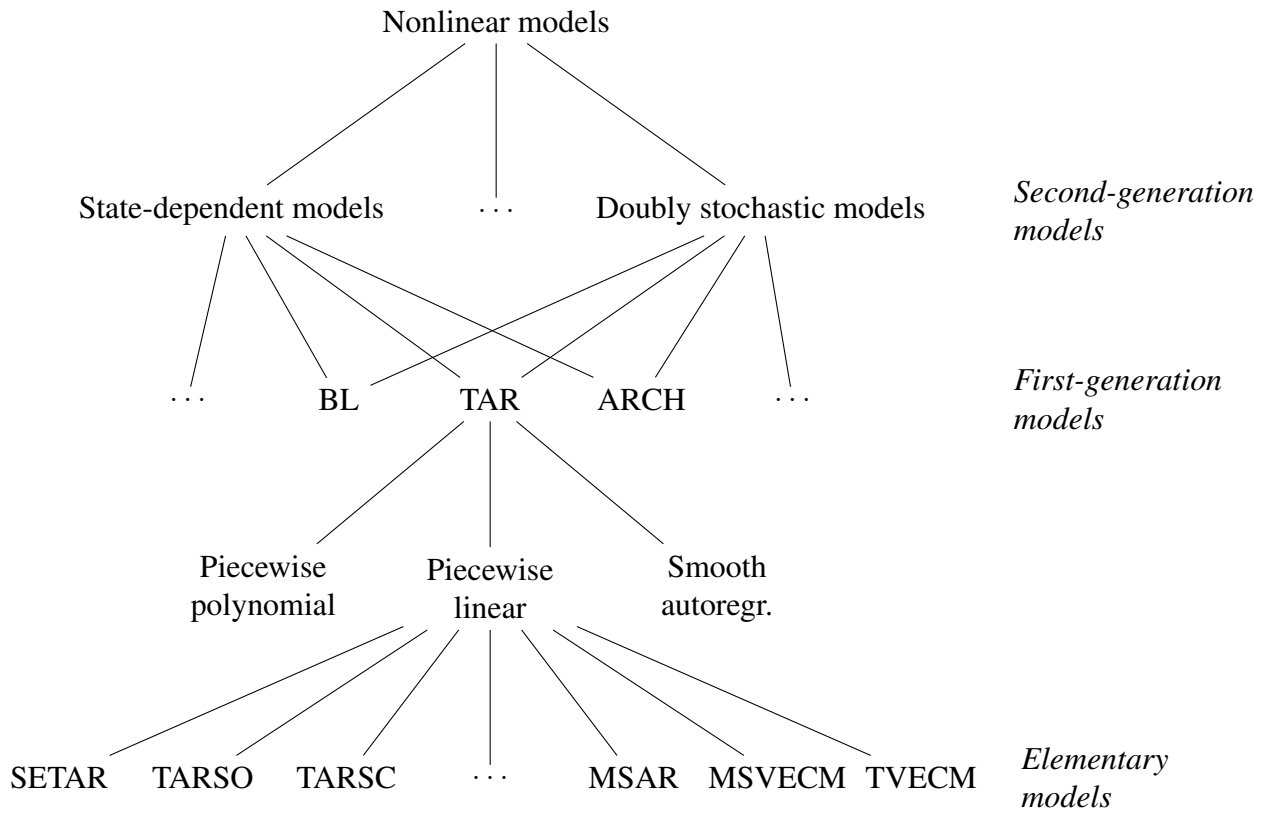


Figure 2: Transactions Costs and Regime-Dependent PT

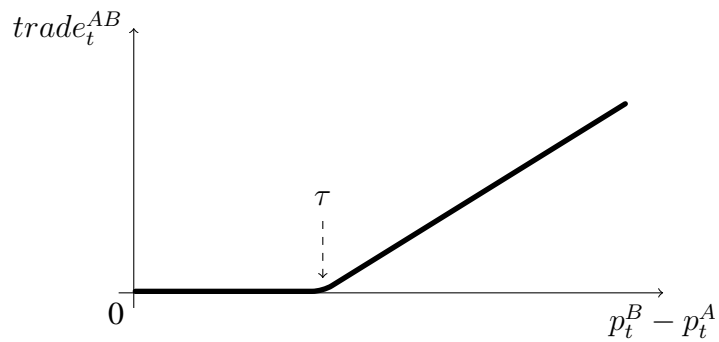
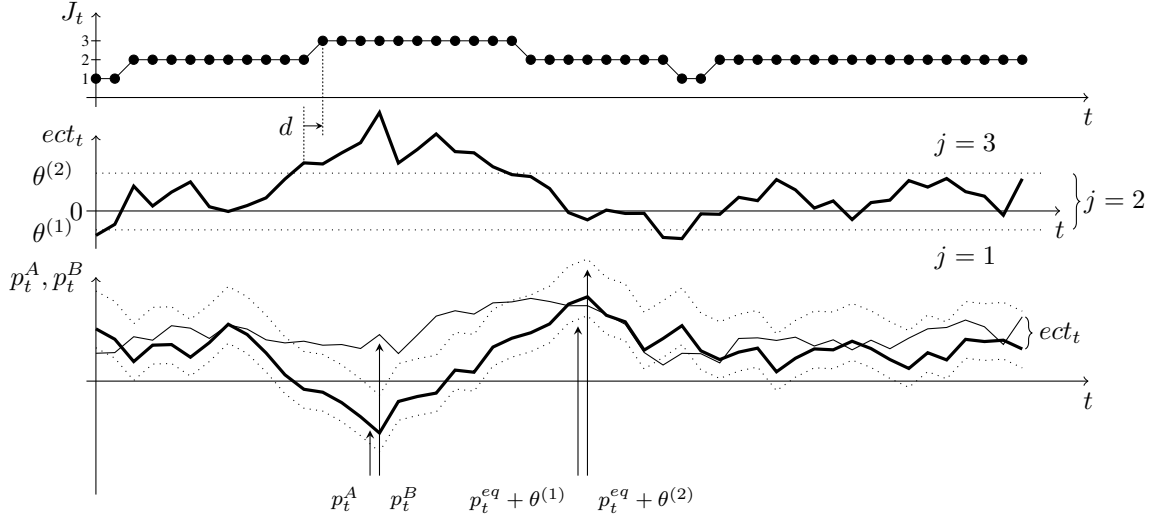


Figure 3: Realization of a TVECM(3)



Parameters:  $\mu^{(j)} = \mathbf{0}$ ,  $\alpha^{(1)} = (0.05 \ -0.05)^\top$ ,  $\alpha^{(3)} = (0.05 \ -0.05)^\top$ ,  $\theta^{(1)} = -1$ ,  $\theta^{(2)} = 2$ ,  $d = 1$ ,  $\Psi_i^{(j)} = 0$  and  $\beta = (1 \ -1)^\top$  implying a long-run equilibrium price  $p_t^{eq} = p_t^A$ .

Figure 4: Transition Graph of a Two-State Markov Chain

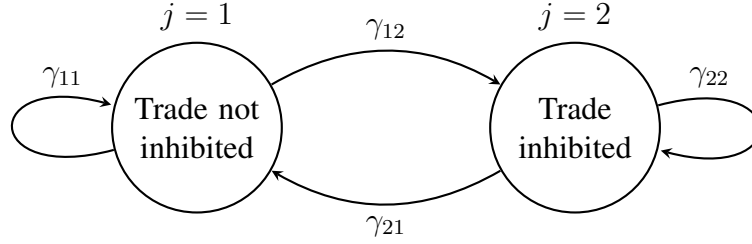
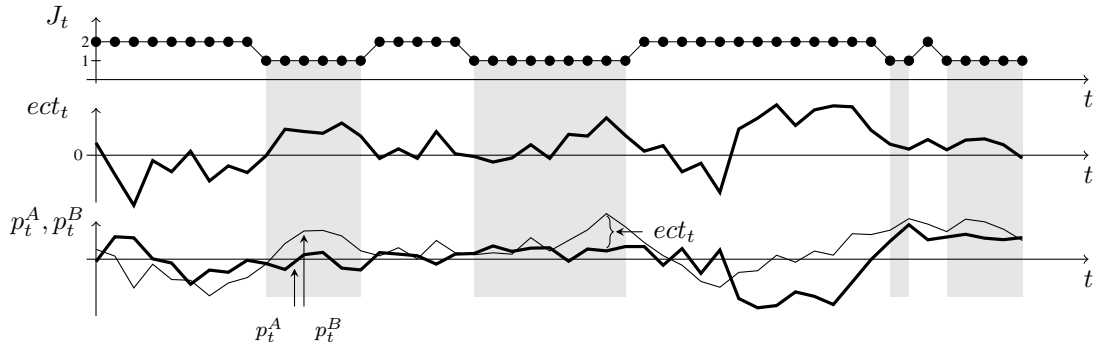


Figure 5: Realization of a MSVECM(2)



Parameters:  $M = 2$ ,  $\mu^{(J_t)} = \mathbf{0}$ ,  $j = 1$ :  $\alpha^{(1)} = (0.05 \ -0.05)^\top$ ,  $j = 2$ :  $\alpha^{(2)} = (0.25 \ -0.25)^\top$ ,  $\gamma_{11} = \gamma_{22} = 0.8$ ,  $\Psi_i^{(J_t)} = 0$  and  $\beta = (1 \ -1)^\top$  implying a long-run equilibrium price  $p_t^{eq} = p_t^A$ .



Figure 6: The Expectation-Maximization Algorithm

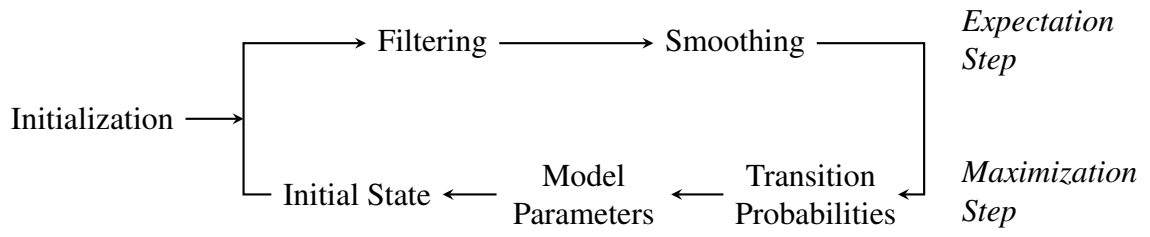


Figure 7:  $\widehat{\text{MSE}}_{\hat{\theta}^{(1)}} = f(\theta^{(1)}, \theta^{(2)})$  vs.  $\widehat{\text{MSE}}_{\hat{\theta}^{(1)}} = f(\rho^{(1)}, \rho^{(3)})$

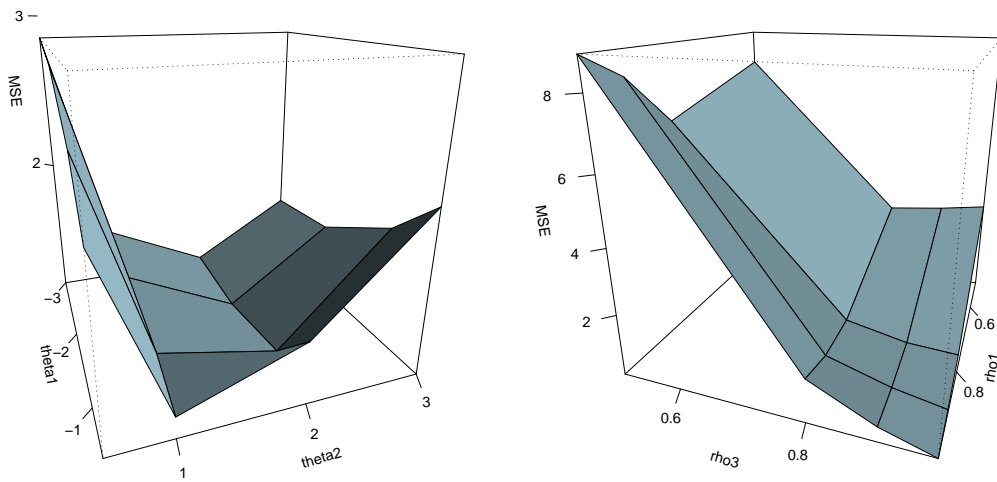


Figure 8:  $\widehat{\text{MSE}}_{\hat{\pi}^{(1)}} = f(\pi^{(2)}, \pi^{(3)})$  vs.  $\widehat{\text{MSE}}_{\hat{\pi}^{(1)}} = f(\rho^{(1)}, \rho^{(3)})$

