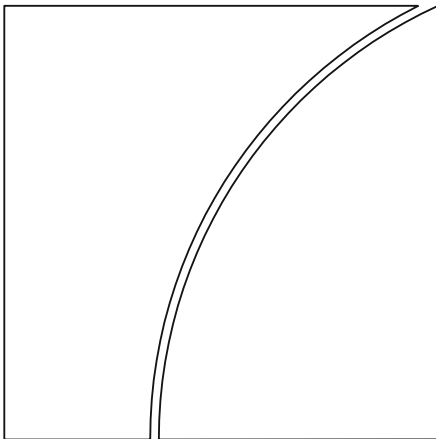# BIS Working Papers
No 689

Estimating unknown arbitrage costs: evidence from a three-regime threshold vector error correction model

by Kristyna Ters and Jörg Urban

Monetary and Economic Department

January 2018

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

Estimating unknown arbitrage costs: Evidence from a 3-regime threshold vector error correction model[*]

Kristyna Ters[†]   Jörg Urban[‡]

Abstract

We present a methodology for estimating a 3-regime threshold vector error correction model (TVECM) with an unknown cointegrating vector based on a new dynamic grid evaluation. This model is particularly suited to estimating deviations from parity conditions such as unknown arbitrage costs in markets with a persistent non-zero basis between two similar financial market instruments traded in the spot and the derivative markets. Our proposed 3-regime TVECM can estimate the area where arbitrageurs have no incentives for trading. Only when the basis exceeds a critical threshold, where the potential gain from the basis trade exceeds the overall transaction costs, do we expect arbitrageurs to step in and carry out the respective trade. This leads to non-linear adjustment dynamics and regimes with different characteristics. The overall transaction costs for the basis trades can be inferred from the estimated no-arbitrage regime. Our methodology allows us to quantify overall transaction costs for an arbitrage trade in markets where trading costs are opaque or unknown, as in credit risk or index arbitrage trading. The key contributions of this paper are the further development of the 2-threshold VECM, together with the numerical evaluation of the model through numerous simulations to prove its robustness. We present two short applications of the model in arbitrage trades in the palladium market and index trading for the S&P 500.

[†] University of Basel, Basel, Switzerland. email: kristyna.ters@unibas.ch

[‡] University of Basel, Basel, Switzerland. email: joerg.urban@unibas.ch (corresponding author)

1   Introduction

We present a methodology for estimating three-regime vector error correction models with two thresholds and an unknown cointegrating vector. Our proposed methodology is particularly suited to modelling arbitrage in markets with frictions and allows us to estimate unknown transaction costs for such trades.

The theoretical no-arbitrage condition between two similar financial market instruments traded in the spot and derivative markets, is a cornerstone for the empirical research on economic parity relationships. The no-arbitrage condition requires that the pricing in the spot market must be equal to the derivative market. If not, any pricing discrepancy would present investors with an arbitrage opportunity which will disappear rapidly, as arbitrageurs will exploit any mispricing. This mispricing is measured by the so-called basis. For the arbitrage condition to hold, markets must be perfect and frictionless. In practice, however, frictions and imperfections often make such arbitrage trades difficult and costly to varying degrees. These imperfections include, amongst others, limited and time-varying liquidity across market segments, unavailability of instruments with identical maturity and payout structures, and the fact that some arbitrage trades require large amounts of capital to be tied up for extended periods of time.

The basis trading strategy, in which an arbitrageur believes that two similar financial market instruments are mispriced relative to each other, aims to take opposing long and short positions in these two securities in order to make a gain on the convergence of their values. In the case of a positive basis, arbitrageurs will bet on a weakening basis (short basis position) and in the case of a negative basis, arbitrageurs bet on a strengthening basis (long basis position). There exists no universal definition of the basis, and different definitions are more commonly used in different markets. In credit risk markets the basis is defined as derivatives minus spot price (Gyntelberg et al.; 2013) or more specifically as the CDS spread minus the spread on a par risky fixed-rate bond over the risk-free rate. Lien and Yang (2008) define the basis as the difference between spot and future prices in their application in commodity markets. Fama and French (1987) and McMillan (2005), on the other hand, define the basis as future minus spot prices. Our proposed methodology is independent of the chosen specification.

A substantial part of the transaction costs of an arbitrage transaction is unknown when the trade is initiated, making the arbitrage trade risky. For index trades, for example, Sutcliffe (2006) states that this risk can occur because the bid-ask spread and brokers' commission, when unwinding the spot position at delivery, vary with the value of the index basket and that there may be a transaction tax which varies in proportion to the index (Sutcliffe; 2006). Adams and van Deventer (1993) suggest, that in the case of unknown arbitrage costs, traders should depart from the usual one-to-one ratio for the size of the

spot and futures positions. In order to eliminate the transaction cost risk, while buying shares and selling futures, the arbitrageur should buy, for every one futures contract sold, $1/(1-p)$ index baskets. p is the proportion of the value of the index basket that must be paid in transaction costs at delivery. When selling shares and buying futures, the arbitrageurs should sell $1/(1+p)$ index baskets for every futures contract bought. Adams and van Deventer (1993) state that this will remove the transaction cost risk from the arbitrage trade. However, they do not propose a methodology that can estimate the unknown transaction costs of such trades.

As a result of existing transaction costs on arbitrage trades, the difference between the prices in the spot and derivatives market for two similar financial market instruments, the basis, is typically not zero. Moreover, the basis can become sizeable and persistent in times of market stress. Finally, when entering into a basis trade, the arbitrageur is exposed to the risk that the trade will move in the wrong direction.

A persistent non-zero basis is therefore likely to reflect the unwillingness of arbitrageurs to try to exploit it, unless the pricing mismatch is greater than the cost of undertaking the arbitrage trade. Empirically, we would therefore expect to see such arbitrage forces intensifying as the magnitude of the basis exceeds some level that reflects the costs that traders face in the market. This suggests that the adjustment process towards the long-run equilibrium is non-linear, in that it differs depending on the level of the basis.

Some research specifically aims to estimate the effect of transaction costs on arbitrage such as Stevens (2015) in the market for crude oil. Stevens (2015) finds that transaction costs increase the persistence of the basis in the market for crude oil. He explains the non-zero basis by the absence of arbitrage. Forbes et al. (1999) investigate index futures arbitrage for the S&P500 stock index and the nearest-to-delivery futures contract and find significant transaction costs that prevent arbitrage. Forbes et al. (1999) also find clear indications for arbitrage trading when the basis breaks out beyond a threshold. However, both Forbes et al. (1999) and Stevens (2015) employ a univariate structural change test to the cointegrating residual based on Tsay (1989). This approach would, however, only be valid when the cointegrating vector is known a priori. Neither of the above-mentioned papers provide a solution for this problem. Forbes et al. (1999) also state in their conclusion, that the problem of an unknown cointegrating relationship in multiple threshold error correction models has not yet been resolved.

Hansen and Seo (2002) provide a methodology for estimating 2-regime threshold vector error correction (TVECM) models with an unknown cointegrating relationship. However, they do not provide a solution for the case beyond two regimes. An extension to a 3-regime TVECM with two thresholds is important from an economic viewpoint, as it allows us to test for positive and negative deviations from theoretical parity relationships. For example it allows to test for the existence of transaction costs in positive and negative

2

basis trading strategies. Also, the model setup as proposed by Hansen and Seo (2002) is not ideal for economic and financial market problems, where often a significant deviation from the theoretical parity relationship exists, because they did not account for a persistent deviation in their long-term equilibrium condition.

We contribute to the existing literature by developing an estimation procedure for threshold error correction models with three regimes (two thresholds) and an unknown cointegrating vector which is especially suited to modelling arbitrage in markets with frictions and unknown transaction costs. The estimation of an unknown cointegrating vector is particularly important for distorted parity relationships such as in financial markets and economic applications that exhibit a significant non-zero deviation from the theoretical parity relationship. Our proposed model allows for non-linear adjustment of prices, in derivative and spot markets, towards the long-run equilibrium. Consequently, we can estimate the region where arbitrageurs step into the market as the trading opportunity is 'sufficiently profitable' for both positive and negative basis trades. The overall transaction costs can be inferred from the no-arbitrage regime.

The rest of the paper is structured as follows: Section 2 discusses the setup and estimation of the TVECM. Section 3 provides some empirical applications for the illustration of our methodology and Section 4 concludes.

2   3-regime threshold vector error correction model

Threshold cointegration was introduced by Balke and Fomby (1997) as a feasible mean to combine regime switches and cointegration. In the case of a persistent non-zero basis between the spot and the derivative market we expect to see that if the basis is lower than the cost of undertaking an arbitrage trade, arbitrageurs have no incentives to carry out the trade. Only when the deviation from the long-term equilibrium exceeds a critical threshold, such that the expected profit exceeds the costs, will economic agents act to move the basis back towards its long-term equilibrium. As a result, adjustments to the long-term equilibrium are likely to be regime-dependent, with a relatively weak adjustment mechanism in the regime where arbitrageurs have no incentive for trading as the overall transaction costs exceed the expected profit from the arbitrage trade. By extending the linear VECM approach to a threshold vector error correction model (TVECM) we can model these non-linearities in the adjustment dynamics.

The TVECM approach extends the VECM by allowing the behaviour of price quotes for spots S_t and derivatives D_t for a specific reference entity or underlying to depend on the state of the system. One can formulate a general TVECM with l regimes, ie with $l-1$

thresholds as follows[1]:

$$\Delta y_t = \sum_{j=1}^{l} \left[\lambda^j (S - \beta_1 D - \beta_0)_{t-1} + \Gamma^j(L)\Delta y_t \right] d_t(\beta_0, \, \beta_1, \, \theta^{j-1}, \, \theta^j) + \varepsilon_t, \qquad (1)$$

where $y_t = (S_t \ D_t)^\mathsf{T}$. All thresholds θ^j are ordered and

$$d_t(\beta_0, \, \beta_1, \, \theta^{j-1}, \, \theta^j) = I(\theta^{j-1} \le \mathrm{ec}_{t-1}(\beta_0, \beta_1) < \theta^j), \qquad (2)$$

with the error correction term $\mathrm{ec}_{t-1}(\beta_0, \beta_1) = (S - \beta_1 D - \beta_0)_{t-1}$. The indicator function $I(\theta^{j-1} \le \mathrm{ec}_{t-1}(\beta_0, \beta_1) < \theta^j)$ is 1 if the error correction term $\mathrm{ec}_{t-1}(\beta_0, \beta_1)$ is in the interval $[\theta^{j-1}, \, \theta^j)$ and otherwise 0. Further, by definition the threshold θ^0 is $-\infty$ and θ^l is ∞ in Equation (1). We focus on the bi-variate case where $\Delta y_t = (\Delta S_t \ \Delta D_t)^\mathsf{T}$ is a 2-dimensional I(0) time series. The vector of price quotes $y_t = (S_t \ D_t)^\mathsf{T}$ are cointegrated with unknown β_0 and β_1. The error correction term $\mathrm{ec}_{t-1}(\beta_0, \beta_1)$ is stationary. $\varepsilon_t = (\varepsilon_t^S \ \varepsilon_t^D)^\mathsf{T}$ is a vector of i.i.d. shocks and $j \in \{1, 2, ..., l\}$ is the index denoting the l different regimes.

Equation (1) constitutes a vector autoregressive model in first-order differences with $\Gamma^j(L) = \sum_{k=1}^{m} \alpha^{j,k} L^k$ and L as lag operator, m as the number of VAR lags and an additional error correction term $\lambda^j \mathrm{ec}_{t-1}(\beta_0, \beta_1)$. $\lambda^j = (\lambda_1^j \ \lambda_2^j)^\mathsf{T}$ and the lagged VAR terms are regime-dependent conditioned on the state of the error correction term $\mathrm{ec}_t(\beta_0, \beta_1)$. The speed of adjustment parameters characterize to what extent the price changes in $\Delta y_t = (\Delta S_t \ \Delta D_t)^\mathsf{T}$ react to deviations from the long-term equilibrium.

The Schwarz (Bayesian) information criterion (SIC) is used to determine the VAR order. Lütkepohl (2006) states that in large samples for multivariate models when $T \to \infty$ only the SIC criterion is seen to be strongly consistent for any K-variate system.

The error correction term represents the long-term equilibrium of the two time series which has to be an AR(1) process by construction (Johansen; 1988). The VAR-term represents the short-run dynamics coming from market imperfections (Baillie et al.; 2002).

Contrary to the 1-threshold TVECM proposed by Hansen and Seo (2002) we introduce an intercept β_0 in the error correction term which denotes the deviation from the theoretical parity relationship (long-term equilibrium). This is motivated by our no-arbitrage discussion in Section 1 as the local constant β_0 represents the persistent non-zero basis. In frictionless markets (without deviations from the theoretical parity relationship), the error correction term in Equation (1) is equal to the observed basis $(S - D)_t$, with $\beta_0 = 0$ and $\beta_1 = 1$.

The transaction costs for a basis trade prevent a complete adjustment towards a zero basis. As such, in markets with frictions there may be a neutral band between the derivative and the spot market in which the error correction term in Equation (1) may fluctuate

---

[1] for a derivation of the TVECM see for example Balke and Fomby (1997)

without incentives for market participants to switch funds between the spot and derivatives market. Outside of that neutral band there might however be strong incentives for market participants to switch funds, which results in an adjustment towards the long-term equilibrium. We expect to find the speed of adjustment parameters to indicate that arbitrageurs engaging in $S_t - D_t$ basis trades as soon as the basis exceeds a threshold. In a market with a positive basis ($S_t > D_t$), arbitrageurs bet on a declining basis and will go short in the spot market and go long in the derivative market. In case of a negative basis ($S_t < D_t$), arbitrageurs bet on an increasing basis while carrying out the reverse trade.

According to arbitrage theory we would in general expect to find a 3-regime TVECM with two thresholds when the basis fluctuates between positive and negative figures with sizeable and persistent deviations from zero. The lower regime is defined as $ec_{t-1}(\beta_0, \beta_1) < \theta^1$, the middle regime as $\theta^1 \leq ec_{t-1}(\beta_0, \beta_1) < \theta^2$, and the upper regime is defined as $\theta^2 \leq ec_{t-1}(\beta_0, \beta_1)$. The middle regime is the neutral band where no arbitrage trading occurs while the outer regimes (lower and upper) are the arbitrage regimes. There may also be certain markets or time periods with only a persistent positive basis. In that case we expect to find at most two regimes ($l = 2$) with only one threshold θ^1. The lower regime (neutral regime) is defined as $ec_{t-1}(\beta_0, \beta_1) < \theta^1$, and the upper regime as $\theta^1 \leq ec_{t-1}(\beta_0, \beta_1)$. The regimes are reversed in case of a persistent negative basis market.

Therefore, we will discuss three classes of nested models: model class $\mathcal{H}_1$ is a 1-regime VECM, with no statistical significant threshold, ie where markets are efficient enough to not allow the basis to deviate too far from zero. In markets described by model class $\mathcal{H}_1$ we have no market frictions such as transaction costs, hence arbitrageurs will step in as soon as the basis deviates from zero. Model class $\mathcal{H}_2$ is a 2-regime TVECM and model class $\mathcal{H}_3$ is a 3-regime TVECM.

The threshold θ^j is computed relative to the estimated basis $S - \beta_1 D - \beta_0$. Therefore, transaction costs, which are defined relative to the observed/real basis $S - D$, are the sum $\theta^j + \beta_0$. In the case of $\theta^j + \beta_0 < 0$, we have transaction costs for an arbitrage trade on basis strengthening and in the case of $\theta^j + \beta_0 > 0$ we have transaction costs for an arbitrage trade on basis weakening. In the case of $\theta^j + \beta_0 < 0$, the transaction costs have to be interpreted in absolute terms.

We will discuss the three model classes: i) $\mathcal{H}_1$: VECM, ii) $\mathcal{H}_2$: 2-regime TVECM with an unknown cointegrating vector and iii) $\mathcal{H}_3$: 3-regime TVECM with an unknown cointegrating vector.

Figure 1: Model classes $\mathcal{H}_i$

The VECM model class $\mathcal{H}_1$ represents markets or periods, where the basis only marginally deviates from zero. The markets are perfect and frictionless for arbitrageurs to step in immediately to correct the pricing differential between the spot and the derivatives market. Model classes $\mathcal{H}_2$ and $\mathcal{H}_3$ are classical multi-regime models, where outside the neutral regime arbitrageurs engage in basis trades in the outer regimes (lower and upper regimes). Note that a model belonging to class $\mathcal{H}_2$ can also have a negative basis and consequently a negative threshold.
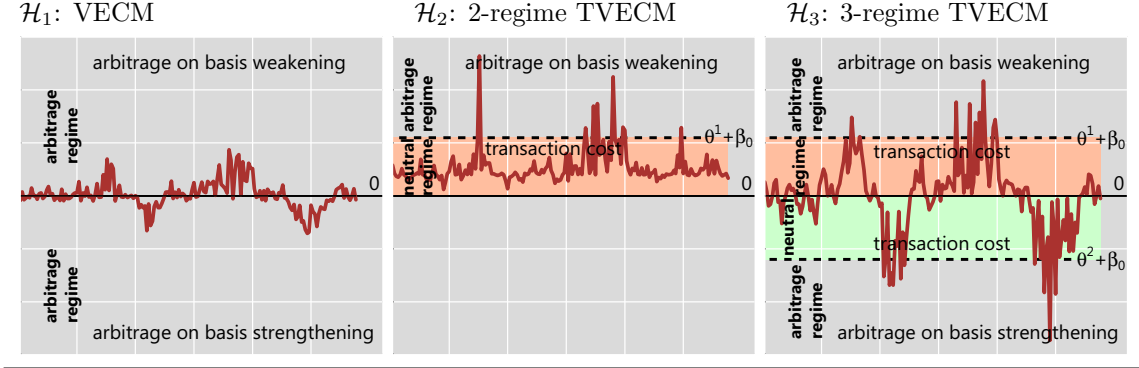
From Equation (1), the VECM (model class $\mathcal{H}_1$) can be rewritten as:

$$
\begin{aligned}
\Delta y_t &= \left[\lambda_1^1 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_1^1(L)\Delta y_t\right] d_t(\beta_0, \beta_1, -\infty, \infty) + \varepsilon_t \\
&= \lambda_1^1 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_1^1(L)\Delta y_t + \varepsilon_t\,,
\end{aligned} \tag{3}
$$

where the last line simply takes into account that for model class $\mathcal{H}_1$ the function d_t is identical to 1 on the entire t-axis. The subscript k in λ_k^i and Γ_k^i indicates explicitly to which model class the parameters belong, whereas the superscript i denotes the regime.

Equation (1) can for 2-regime TVECM (model class $\mathcal{H}_2$) be rewritten as:

$$
\begin{aligned}
\Delta y_t &= \left[\lambda_2^1 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_2^1(L)\Delta y_t\right] d_t(\beta_0, \beta_1, -\infty, \theta^1) \\
&\quad + \left[\lambda_2^2 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_2^2(L)\Delta y_t\right] d_t(\beta_0, \beta_1, \theta^1, \infty) + \varepsilon_t\,.
\end{aligned} \tag{4}
$$

The 3-regime TVECM (model class $\mathcal{H}_3$) takes the following form:

$$
\begin{aligned}
\Delta y_t &= \left[\lambda_3^1 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_3^1(L)\Delta y_t\right] d_t(\beta_0, \beta_1, -\infty, \theta^1) \\
&\quad + \left[\lambda_3^2 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_3^2(L)\Delta y_t\right] d_t(\beta_0, \beta_1, \theta^1, \theta^2) \\
&\quad + \left[\lambda_3^3 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_3^3(L)\Delta y_t\right] d_t(\beta_0, \beta_1, \theta^2, \infty) + \varepsilon_t.
\end{aligned} \tag{5}
$$

6

All parameters are allowed to switch between the regimes except for β_i. Following Hansen and Seo (2002) we estimate the model classes by imposing the following additional constraint for each regime:

$$\pi_0 \leq P(\theta^{j-1} \leq ec_{t-1}(\beta_0, \beta_1) < \theta^j) \leq 1 - \pi_0\,, \qquad (6)$$

where $\pi_0 > 0$ is a trimming parameter and P is the share of observations in each regime. This constraint allows us to identify a threshold effect only if the share of observations in each regime is large enough, ie is greater than π_0. If this condition is not met the model $\mathcal{H}_i$ reduces to $\mathcal{H}_{i-1}$ for $i > 2$ (eg from $\mathcal{H}_3$ to $\mathcal{H}_2$). Andrews (1993) argues that setting π_0 between 0.05 and 0.15 are typically good choices. We chose as a baseline setup $\pi_0 = 0.1$, but perform robustness tests also for $\pi_0 = 0.05$ and $\pi_0 = 0.15$.

2.1 Estimating the model

The most important statistical issue for threshold models is estimating the unknown threshold(s) and test for their significance. Balke and Fomby (1997) suggest transforming the TVECM into a univariate arranged autoregression, while Tsay (1989) reformulates the problem into a univariate structural change test to the cointegrating residual. However, these approaches are valid only in the univariate case when the cointegrating vector is known.

We contribute to the existing literature by examining the case of an unknown cointegrating vector for a 3-regime TVECM. We implement, as Hansen and Seo (2002) for the 2-regime TVECM, a maximum likelihood estimation of a bivariate TVECM with the assumption of i.i.d. Gaussian error terms. The likelihood function to be maximized for a l regime model takes the form:[2]

$$\mathcal{L}_n(\lambda_l^1, \ldots, \lambda_l^l, \Gamma_l^1, \ldots, \Gamma_l^l, \beta_0, \beta_1, \theta^1, \ldots, \theta^{l-1}, \Sigma) = -\frac{n}{2}\ln|\Sigma| - \sum_{t=1}^{n}\frac{1}{2}\varepsilon_t^{\mathsf{T}}\Sigma^{-1}\varepsilon_t\,, \qquad (7)$$

with $\Sigma = \mathrm{E}(\varepsilon_t\,\varepsilon_t^{\mathsf{T}})$ and n represents the sample size. ε_t and Σ are functions of λ_l^i, Γ_l^i, β_0, β_1 and θ^j, where $j = 1$ to $l - 1$ and $i = 1$ to l.

Hansen and Seo (2002) suggest, that it is computationally convenient to hold β_0 and β_1 as well as θ^j fixed and compute the concentrated maximum likelihood estimations for λ_l^i, Γ_l^i and Σ. Due to the linearity of the model, this is simply an OLS regression. As shown in Hansen and Seo (2002) the concentrated likelihood function for β_0, β_1 and thresholds θ^j for a l regime model is:

---

[2] We will focus on $l = 1$, 2 and 3 in our further analysis.

$$
\begin{aligned}
\mathcal{L}_n(\beta_0,\,\beta_1,\,\theta^j) &= \mathcal{L}_n(\hat{\lambda}_l^i(\beta_0,\,\beta_1,\,\theta^j),\,\hat{\Gamma}_l^i(\beta_0,\,\beta_1,\,\theta^j),\,\hat{\Sigma}(\beta_0,\,\beta_1,\,\theta^j),\,\beta_0,\,\beta_1,\,\theta^j) \\
&= -\frac{n}{2}\ln|\hat{\Sigma}(\beta_0,\,\beta_1,\,\theta^j)| - \frac{n\cdot p}{2}\,, \qquad\qquad (8)
\end{aligned}
$$

where again $j = 1$ to $l-1$, $i = 1$ to l in a l regime case and all variables with a hat are OLS estimators. We consider the bi-variate case with $p = 2$. The remaining task of finding the maximum likelihood estimation of β_0, β_1 and the thresholds is therefore to minimize $\ln|\hat{\Sigma}(\beta_0,\,\beta_1,\,\theta^j)|$, subject to the constraint in Equation (6).

Unfortunately, the function in Equation (8) is not smooth (see for example the left-hand panel of Figure 2), hence conventional hill-climbing algorithms cannot be used to find the minimum. Therefore, Hansen and Seo (2002) suggest a joint grid search. We present evidence of the good performance of the proposed grid search in Section 2.3.

Two issues remain to be discussed with respect to the parameter estimation. The 2-regime model used by Hansen and Seo (2002) requires a two-dimensional grid search over $(\beta_1,\,\theta^1)$. Our proposed model requires a search over a three dimensional grid $(\beta_0,\,\beta_1,\,\theta^1)$ in the 2-regime case and a four-dimensional grid search $(\beta_0,\,\beta_1,\,\theta^1,\,\theta^2)$ in the 3-regime case. The 3-regime model requires the evaluation of Equation (8) at 100 million grid points, if we evaluate each variable at 100 grid points. In order to keep the computation feasible, we suggest a sequential threshold search as it was proven to be consistent by Bai (1997) and Bai and Perron (1998).

A standard sequential search would require 2×100^3 grid point evaluations for the two thresholds, ie 100^3 for β_0, β_1 and θ^1 during the first grid search with 100 grid points each and then another 100^3 grid point evaluations for β_0, β_1 and θ^2 during the second search with 100 grid points each. We will however show that it is efficient to fix β_1 in the second threshold search to the value found in the first threshold search and hence reduce the second search to a two-dimensional space. This reduces the computational burden dramatically to 100^3 grid points for the first search and 100^2 grid points for the second search. We will discuss and justify this proposal in a comprehensive simulation exercise in Section 2.3. In the same section we will also show that we cannot fix β_0 in the threshold search for θ^2, unlike β_1, as the β_0 estimate suffers from a large uncertainty.

The second remaining issue to be addressed is the setup of the "correct" search area for each parameter. The search region $[\theta_L,\,\theta_U]$ for the thresholds is straightforward as it must be identical to the interval $[\min(\mathrm{ec}_t(\beta_0,\,\beta_1)),\,\max(\mathrm{ec}_t(\beta_0,\,\beta_1))]$ given by the error correction term for β_0 and β_1. The region for the β_0 and β_1 parameters can be calibrated based on the estimates of the linear VECM model and the theoretical values $\beta_1 = 1$ and $\beta_0 = 0$, which would constitute the observed basis $(S - \beta_1 D - \beta_0)_t = (S - D)_t$. It is important to keep the search area for β_0 and β_1 large enough to include the minimum,

but not too large to reduce the precision of the grid search. The grid search for θ^2 will be reduced by the constraint $\theta^2 + \beta_0 > 0$ if the first search resulted in $\theta^1 + \beta_0 < 0$ and vice versa if $\theta^1 + \beta_0 > 0$. This is purely based on our previous arbitrage discussion, where we expect to find at most two transaction costs, one for a positive basis trade and one for a negative basis trade.

Obviously, the precision of the estimated parameters will depend on the distance between two neighbouring grid points. We will test in Section 2.3 various grid sizes, such as 10, 50 or 100 grid points for each search dimension as well as a dynamic grid setting. The dynamic grid setting has no fixed number of grid points, but a precision parameter fixes the distance between two neighbouring grid points. For our simulation exercise we have chosen $\Delta\beta_0 = 0.5$, $\Delta\beta_1 = 0.01$ and $\Delta\theta^j = 0.5$. This choice is motivated by the applications in Section 3 and the data generating process used in Section 2.3, which restrains our time series to become larger in absolute terms than 500. That means, our basis can become as large as $\pm$ 1,000 in extreme cases, leading to approximately 4,000 grid points for β_0 and θ^j. β_1 is in a range of around 0 and 10, leading to maximal 1,000 grid points based on $\Delta\beta_1 = 0.01$.[3] The dynamic grid setting is computationally more expensive but we will show that, as expected, it yields the best results.

2.2 Testing for a threshold

The next step is to determine whether the estimated thresholds $\hat{\theta}^j$ are statistically significant. As the model class $\mathcal{H}_1$ is nested in $\mathcal{H}_2$ and $\mathcal{H}_2$ is nested in class $\mathcal{H}_3$, we start with the discussion of a 1-threshold model $\mathcal{H}_2$. In that case, under the null hypothesis, there is no threshold, so the model reduces to a conventional linear VECM model class $\mathcal{H}_1$ where $\lambda_2^1 = \lambda_2^2 = \lambda_1^1$ and $\Gamma_2^1(L) = \Gamma_2^2(L) = \Gamma_1^1(L)$. The 1-threshold TVECM model class $\mathcal{H}_2$ is detected under the alternative hypothesis with $\lambda_2^1 \neq \lambda_2^2$ under the constraint in Equation (6). As the models are linear, a regular LM test with an asymptotic $\chi^2(N)$-distribution can be calculated from a linear regression on the model in Equation (4). However, the LM test statistic can only be applied if β_i and the threshold variable θ^1 are known a priori (Hansen and Seo; 2002). While the point estimates of β_0 and β_1 under the null hypothesis are $\tilde{\beta}_0$ and $\tilde{\beta}_1$ from the linear model, there is no estimate of θ^1 under the null hypothesis. This implies that there is no distribution theory for the parameter estimates and no conventionally defined LM-like statistic.

We follow Hansen and Seo (2002) who derive the LM-like statistic for the one threshold case. We test for a linear VECM $\mathcal{H}_1$ under the null and a 1-threshold model $\mathcal{H}_2$ under the alternative hypothesis. Our models $\mathcal{H}_1$ and $\mathcal{H}_2$ are defined in Equations (3) and (4).

---

[3] The β_1 grid is gauged around the estimate of the VECM and the theoretical value of the observed basis, which is equal to 1.

The subscripts k in λ_k^i, Γ_k^i and in the below defined functionals denote explicitly to which model class $\mathcal{H}_k$ they belong.

We can derive the LM-like statistic for the first threshold search as:

$$\text{LM}(\beta_0, \beta_1, \theta^1) = \text{vec}(\hat{A}_2^1 - \hat{A}_2^2)^{\mathsf{T}}(\hat{V}_2^1 + \hat{V}_2^2)^{-1}\text{vec}(\hat{A}_2^1 - \hat{A}_2^2), \tag{9}$$

with the OLS estimator $\hat{A}_2^i$

$$\hat{A}_2^i(\beta_0, \beta_1, \theta^1) = \left(\sum_{t=1}^{n} Y_{t-1}(\beta_0, \beta_1)Y_{t-1}(\beta_0, \beta_1)^{\mathsf{T}} d_t(\beta_0, \beta_1, \theta^{i-1}, \theta^i)\right)^{-1}$$

$$\times \left(\sum_{t=1}^{n} Y_{t-1}(\beta_0, \beta_1)\Delta y_t^{\mathsf{T}} d_t(\beta_0, \beta_1, \theta^{i-1}, \theta^i)\right), \tag{10}$$

where $i \in \{1, 2\}$. n is the length of the time series. $\hat{A}_2^i$ are $(1 + pm) \times p$ matrices, with m denoting the number of VAR lags in the model and $p = 2$ for the bi-variate case. $\hat{V}_2^i$ is defined via the moment functionals M_2^i and Ω_2^i:

$$\hat{V}_2^i(\beta_0, \beta_1, \theta^1) = M_2^i(\beta_0, \beta_1, \theta^1)^{-1}\Omega_2^i(\beta_0, \beta_1, \theta^1)M_2^i(\beta_0, \beta_1, \theta^1)^{-1}. \tag{11}$$

The moment functionals are defined as:

$$M_2^i(\beta_0, \beta_1, \theta^1) = \mathbb{1}_p \otimes Y^i(\beta_0, \beta_1, \theta^1)^{\mathsf{T}}Y^i(\beta_0, \beta_1, \theta^1), \tag{12}$$

$$\Omega_2^i(\beta_0, \beta_1, \theta^1) = \xi^i(\beta_0, \beta_1, \theta^1)^{\mathsf{T}}\xi^i(\beta_0, \beta_1, \theta^1), \tag{13}$$

which are both $p \cdot (1 + p \cdot m) \times p \cdot (1 + p \cdot m)$ matrices. Y^i is a short form of the matrices of the stacked rows $Y_{t-1}(\beta_0, \beta_1) \circ d_t(\beta_0, \beta_1, \theta^{i-1}, \theta^i)$, with

$$Y_{t-1}(\beta_0, \beta_1) = \begin{pmatrix} ec_{t-1}(\beta_0, \beta_1) \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-m} \end{pmatrix}. \tag{14}$$

Hence, Y^i is a $t \times (1 + p \cdot m)$ matrix of the following form:

$$Y^i(\beta_0, \beta_1, \theta^1) = \begin{pmatrix} \vdots & & & \\ ec_{t-1}(\beta_0, \beta_1) & \Delta y_{t-1}^{\mathsf{T}} & \cdots & \Delta y_{t-m}^{\mathsf{T}} \\ \vdots & & & \end{pmatrix} \circ d_t(\beta_0, \beta_1, \theta^{i-1}, \theta^i), \tag{15}$$

10

where $\circ$ denotes elementwise multiplication. Y^i contains only non-zero entries if $\theta^{i-1} \leq$ $\text{ec}_{t-1} < \theta^i$. For the here considered 2-regime or 1-threshold TVECM model class $\mathcal{H}_2$ we have $i \in \{1, 2\}$ with $\theta^0 = -\infty$ and $\theta^2 = \infty$.

ξ^i is defined as $\tilde{\varepsilon}_t \otimes Y_{t-1}(\beta_0, \beta_1) \circ d_t(\beta_0, \beta_1, \theta^{i-1}, \theta^i)$, with $\tilde{\varepsilon}_t$ is the OLS estimate of the residual vector from the linear model and $\otimes$ is the Kronecker product.

Equation (9) can be evaluated at the point estimates of the null, which is the model $\mathcal{H}_1$, if the parameters β_0, β_1 and θ^1 would be known. However, there is no estimate of θ^1 for model class $\mathcal{H}_1$. Based on the union-intersection principle, Davies (1987) proposes:

$$\text{SupLM} = \sup_{\theta^1_L \leq \theta^1 \leq \theta^1_U} \text{LM}(\tilde{\beta}_0, \tilde{\beta}_1, \theta^1) \tag{16}$$

with $\tilde{\beta}_i$ being the point estimates obtained under the null hypothesis (linear VECM, model class $\mathcal{H}_1$).

According to the constraint in Equation (6) we set the search region $[\theta_L, \theta_U]$ such that θ_L is the π_0 percentile of the error correction term, and θ_U is the $(1 - \pi_0)$ percentile. This grid evaluation over $[\theta_L, \theta_U]$ is necessary to implement the maximisation defined in Equation (16) as the function $\text{LM}(\tilde{\beta}_0, \tilde{\beta}_1, \theta^1)$ is non-differentiable in θ^1 and hence conventional hill-climbing algorithm cannot be used to find the extremum.

The value of θ^1 which maximizes Equation (16) is different from the MLE $\hat{\theta}^1$ in Section 2.1, as Equation (16) are LM tests that are based on parameter estimates obtained under the null hypothesis, ie $\mathcal{H}_1$. Also, the test statistic is calculated with HAC-consistent covariance matrix estimates which leads to differing estimates compared to the estimate in Section 2.1 (see also the discussion in Hansen and Seo (2002)).

For three sample simulations we present the maximum likelihood function and the supremum LM estimator in Figure 2, as well as the corresponding estimators of β_0 and β_1 used to compute the functions presented there. The left-hand side of Figure 2 shows the β_i estimates under the alternative hypothesis, ie in this case the TVECM (class $\mathcal{H}_2$), and therefore the minimum is very close to the theoretical threshold. The right-hand side shows the β_i estimates obtained from the linear VECM ($\mathcal{H}_1$) estimation, which are far off from the values used in the simulation ($\beta_0 = 10$, $\beta_1 = 1.1$), and therefore the maxima of the LM estimator are also far from the theoretical threshold, which is 3.

This issue was also discussed by Hansen and Seo (2002). The displayed difference in the estimated thresholds is generic and not special to threshold cointegration. It is important to stress that the supremum LM test is performed to compute the critical value, which is later used to decide if the computed threshold is statistically significant. The supremum LM test is not used to estimated the model parameters.

Figure 2: Maximum likelihood estimator versus supremum LM estimator

---

The graphs show the maximum likelihood estimation and the supremum LM estimation for three sample simulations. Each graph contains also the corresponding estimators of β_0 and β_1 for each simulation. The left-hand side of Figure 2 shows the β_i values estimated under the alternative hypothesis, ie in this case the TVECM (class $\mathcal{H}_2$), and where the minimum is very close to the theoretically expected threshold. The right-hand side shows β_i obtained from the VECM estimation, which are far off from the values used in the simulation ($\beta_0 = 10$, $\beta_1 = 1.1$), and so are the maxima of the supremum LM estimator far from the theoretical threshold, which was chosen to be 3.

Maximum Likelihood estimator

Supremum LM estimator
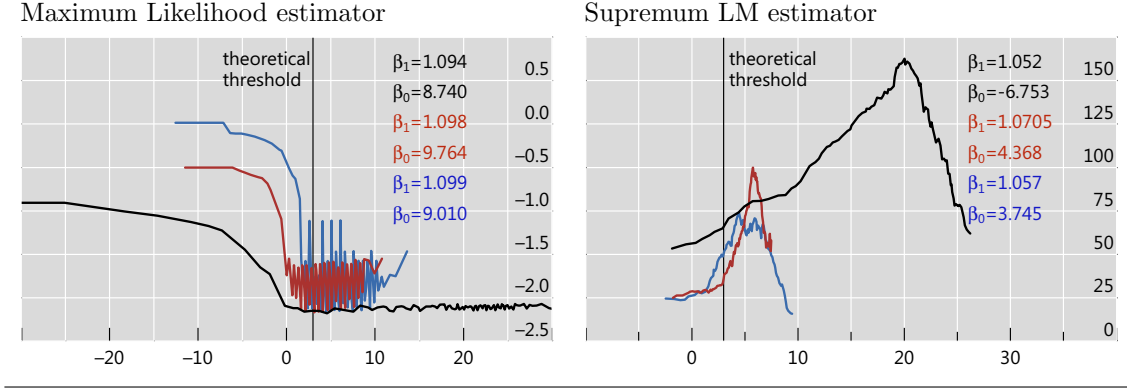
Just like in the one threshold case ($\mathcal{H}_2$), the next step is to determine whether the estimated 2-threshold TVECM ($\mathcal{H}_3$) is statistically significant. Under the null hypothesis, there is one threshold θ^1, so the model reduces to a TVECM with two regimes ($\mathcal{H}_2$) described in Section 2.1 with $\lambda_3^1 \neq \lambda_3^2 = \lambda_3^3$. The 2-threshold TVECM is detected under the alternative hypothesis $\mathcal{H}_3$ with $\lambda_3^1 \neq \lambda_3^2 \neq \lambda_3^3$. The constraint in Equation (6) is again applied. To be explicit, now the null hypothesis is the model $\mathcal{H}_2$:

$$
\begin{aligned}
\Delta y_t \;=\; & \left[\lambda_2^1 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_2^1(L)\Delta y_t\right] d_t(\beta_0, \beta_1, -\infty, \theta^1) \\
& + \left[\lambda_2^2 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_2^2(L)\Delta y_t\right] d_t(\beta_0, \beta_1, \theta^1, \infty) + \varepsilon_t
\end{aligned}
\tag{17}
$$

and the alternative hypothesis is model $\mathcal{H}_3$:

$$
\begin{aligned}
\Delta y_t \;=\; & \left[\lambda_3^1 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_3^1(L)\Delta y_t\right] d_t(\beta_0, \beta_1, -\infty, \theta^1) \\
& + \left[\lambda_3^2 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_3^2(L)\Delta y_t\right] d_t(\beta_0, \beta_1, \theta^1, \theta^2) \\
& + \left[\lambda_3^3 \mathrm{ec}_{t-1}(\beta_0, \beta_1) + \Gamma_3^3(L)\Delta y_t\right] d_t(\beta_0, \beta_1, \theta^2, \infty) + \varepsilon_t.
\end{aligned}
\tag{18}
$$

Following the same steps and arguments discussed before we can define the following LM-like statistics:

$$\text{LM}(\beta_0, \beta_1, \theta^1, \theta^2) = \text{vec}(\hat{A}_3^2 - \hat{A}_3^3)^\mathsf{T} (\hat{V}_3^2 + \hat{V}_3^3)^{-1} \text{vec}(\hat{A}_3^2 - \hat{A}_3^3). \tag{19}$$

Following the proposition by Davies (1987) on the union-intersection principle we get:

$$\text{SupLM} = \sup_{\theta_L^2 \leq \theta^2 \leq \theta_U^2} \text{LM}(\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\theta}^1, \theta^2) \tag{20}$$

with the point estimates $\tilde{\beta}_i$ and $\tilde{\theta}^1$ obtained under the null (1-threshold TVECM). There is no point estimate of θ^2 under the null hypothesis. We perform again a grid search with the search region for θ^2 subject to the constraint in Equation (6). Furthermore, based on our arbitrage discussion in Section 1 and the assumption, that we have at most one positive (basis weakening trade) and one negative (basis strengthening trade) transaction cost, we further impose for the grid search of θ^2 and β_0 the constraints that $\theta^2 + \beta_0 > 0$ if the first search resulted in transaction costs $\theta^1 + \beta_0 < 0$ and $\theta^2 + \beta_0 < 0$ if the first search resulted in transaction costs $\theta^1 + \beta_0 > 0$.

As there is no formal distribution theory in the case under discussion we follow the proposition by Hansen and Seo (2002) and perform two different bootstrap methodologies in order to estimate the asymptotic distribution for our model specification in Equation (1).

2.2.1 Fixed regressor bootstrap

We implement a non-parametric bootstrap of the residuals, called the "fixed regressor bootstrap", which resamples (Monte-Carlo) the residuals from the estimated linear VECM or 1-threshold TVECM, in the case of the threshold search for θ^1 or θ^2, respectively.

We follow the discussion of Hansen and Seo (2002), Hansen (2000) and Hansen (1996). We take estimates under the null hypothesis of β_i, denoted as $\tilde{\beta}_i$, and define $\tilde{\text{ec}}_{t-1} \equiv \text{ec}_{t-1}(\tilde{\beta}_i)$ and $\tilde{Y}_{t-1} \equiv Y_{t-1}(\tilde{\beta}_i)$, whereby ec_{t-1} denotes the error correction term (see definition below Equation (2)) and Y_{t-1} is defined in Equation (14). Further, $\tilde{\varepsilon}_t$ are the residuals of the null. The name "fixed regressor bootstrap" conveys the message that $\tilde{\beta}_i$, $\tilde{\varepsilon}_t$, $\tilde{\text{ec}}_{t-1}$ and $\tilde{Y}_{t-1}$ are kept fixed at their sample values.

Next, we compute a large number of times (eg 1,000) $y_{bt} = \tilde{\varepsilon}_t e_{bt}$, whereby e_{bt} is i.i.d. N(0,1) and in each draw e_{bt} is independently chosen. For each draw (identified by the index b), we perform an LM test. $\tilde{\varepsilon}_{bt}$ is computed by regressing y_{bt} on $\tilde{Y}_{t-1}$. $\hat{A}_2^j(\tilde{\beta}_i, \theta^1)_b$ (see Equation (10) for the 1-threshold case)[4] and $\hat{\varepsilon}_{bt}(\tilde{\beta}_i, \theta^1)$ are computed by regressing

---

[4]   For the 2-threshold TVECM we need to compute $\hat{A}_3^j(\tilde{\beta}_i, \tilde{\theta}^1, \theta^2)_b$, whereby the values of the null $\tilde{\beta}_i$ and $\tilde{\theta}^1$ are kept fixed.

y_{bt} on $\tilde{Y}_{t-1}d_{1t}(\tilde{\beta}_i, \theta^1)$ and $\tilde{Y}_{t-1}d_{2t}(\tilde{\beta}_i, \theta^1)$, whereby $\tilde{\beta}_i$ is kept fixed. Further, for each draw b we compute $\hat{V}_2^j(\tilde{\beta}_i, \theta^1)_b$ (see Equation (11) for the 1-threshold TVECM)[5], whereby $\tilde{\beta}_i$ is kept fixed again. Similar to Equations (9) and (16)[6] we compute for each draw b:

$$\text{SupLM}^* = \sup_{\theta_L^1 \leq \theta^1 < \theta_U^1} \text{vec}(\hat{A}_{2b}^1 - \hat{A}_{2b}^2)^\mathsf{T}(\hat{V}_{2b}^1 + \hat{V}_{2b}^2)^{-1}\text{vec}(\hat{A}_{2b}^1 - \hat{A}_{2b}^2), \tag{21}$$

where $\hat{A}_{2b}^i$ and $\hat{V}_{2b}^i$ are functions of the fixed $\tilde{\beta}_i$ and θ^1.

Hansen (1996) has shown that SupLM* is a valid first-order approximation to the asymptotic null distribution of SupLM. Despite having the computational cost of a bootstrap, it only approximates the asymptotic distribution.

The p-value is calculated as the percentage of SupLM* values which exceed the actual SupLM value of the original time series.

2.2.2 Residual bootstrap

This method is fully parametric with respect to the data generating process, that means for the one threshold case we use the complete specification for the null as given by $\mathcal{H}_1$ (single regime VECM versus 1-threshold TVECM as alternative) and for the 2-threshold TVECM we use the complete specification for the null as given by $\mathcal{H}_2$ (1-threshold TVECM versus 2-threshold TVECM as alternative). We further assume ε_t to be i.i.d. from an unknown distribution and fixed initial conditions. To be specific, random draws are made from $\tilde{\varepsilon}_t$, which are the residuals under the null. Using the given initial conditions from the data, and the parameters estimated under the null ($\tilde{\lambda}^i$, $\tilde{\beta}_0$, $\tilde{\beta}_1$ and $\tilde{\Gamma}^i$) we recursively generate the bivariate vector series x_{bt} for the given model ($\mathcal{H}_1$ or $\mathcal{H}_2$). For each draw the SupLM* value is computed and then again the percentage of the SupLM* values which exceed the actual SupLM value (computed from the original time series) gives the p-value.

Hansen and Seo (2002) conjecture that this bootstrap method gives better finite sample performance at the computational cost of being fully parametric with respect to the data generating process.

2.3 Simulation

We perform numerous simulations in order to test the power of the proposed LM tests using different data generating processes for the VECM and TVECM model specifications from Equation (1), using $\text{ec}_{t-1}(\beta_0, \beta_1) = (S - \beta_1 D - \beta_0)_{t-1}$. In particular, we aim to test the power of the two different bootstrap methodologies, the fixed regressor bootstrap

---

[5]   For the 2-threshold TVECM we need to compute $\hat{V}_3^j(\tilde{\beta}_i, \tilde{\theta}^1, \theta^2)_b$, whereby the values of the null $\tilde{\beta}_i$ and $\tilde{\theta}^1$ are kept fixed.
[6]   In the 2-threshold TVECM we need to compute Equation (20).

and the residual bootstrap as discussed in Section 2.2. Figure 3 represents graphically the sequential search for the two thresholds.

Figure 3: Null versus alternative hypothesis in the sequential search for θ^1 and θ^2

The fixed regressor bootstrap and the residual bootstrap test assume that there is no threshold in the threshold search for θ^1, ie for the first threshold search the null hypothesis is a VECM ($\mathcal{H}_1$). For the threshold search for θ^2, the null is a 2-regime TVECM ($\mathcal{H}_2$), with the 3-regime TVECM ($\mathcal{H}_3$) as the alternative hypothesis.
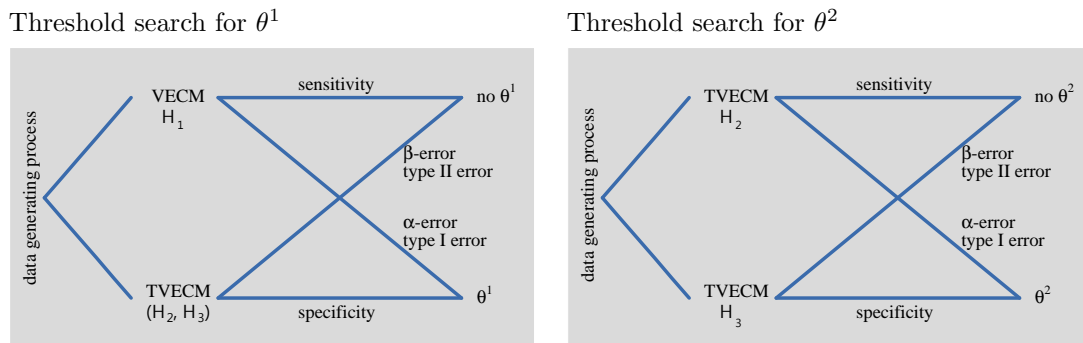
In this subsection we use the following notation: β_i and θ^i are the parameters fixed in the data generating process, while $\hat{\beta}_i$ and $\hat{\theta}^i$ are the estimates from our simulations. For each simulation we generate 1,000 estimators $\hat{\beta}_i$ and $\hat{\theta}^i$ and compute the mean as well as the standard deviation. The mean is compared to β_i and θ^i and the standard deviation is used as a measure of the precision of the estimations.

The aim of the simulation is to understand how precise we can estimate β_i and θ^i, as well as to understand the power of the fixed regressor and the residual bootstrap.

2.3.1 Model class $\mathcal{H}_2$: 2-regime TVECM

We begin our simulation to test the proposed methodology for the 2-regime TVECM $\mathcal{H}_2$ as specified in Equation (4). The aim of the simulations is to understand how precise we can estimate β_0, β_1 and θ^1, and measure the power of the fixed regressor and the residual bootstrap test. We choose as an exemplification $\beta_1 = 1.10$, $\beta_0 = 10$ and $\theta^1 = 3$ for our data generating process. We introduce additional restrictions regarding the generated time series: firstly, we expect that the data generating process in Equation (4) produces time series that are I(1) and a basis $ec_{t-1}(\beta_0, \beta_1)$ that is I(0) at 90% confidence level (CL) using the augmented Dickey-Fuller test. We do also not allow for the time series to become large in absolute terms, ie we set the maximal allowed absolute value of each generated data point to 500. For cross checking purposes we also test a simulation where

we do not enforce these restrictions. We label this simulation in the following tables as "unrestricted" as opposed to the "restricted" cases. We investigate different time series lengths, however our baseline sample size is 1,000 periods. The grid search is performed using an equidistant grid size of 10, 50 and 100 as well as a dynamic grid setting. The dynamic grid setting is determined by a minimum distance, between two individual grid points. The dynamic setting is of course potentially very expensive, as it may lead to a large number of grid points. In our simulation exercise, we have chosen 0.01 as the distance between two grid points for the β_1 grid and 0.5 for the β_0 and θ^1 grid. In the restricted case, where we have the spot and derivative times series confined in absolute terms to 500, the basis can vary between -1,000 and 1,000 in extreme situations.[7] In this case we can expect a grid size in the order of 2,000/0.5=4,000 for the β_0 grid as well as for the θ^1 grid. The grid size for the β_1 grid does usually not reach such extreme values. From an economic viewpoint, the value of β_1 in our data generating process is chosen to be close to 1. Then, the grid is build around the theoretical value $\beta = 1$ (observed basis) and the value found in the initial VECM estimation. Assuming that the VECM estimation is 4, then the grid is gauged from around 0.8 and 4.2, to include the VECM estimate and the theoretical value of 1. This leads to a grid size of (4.2-0.8)/0.01=340. The advantage of the dynamic grid setting is that the precision of the estimation process is predefined and not the number of grid points.

All parameters in the data generating process given in Equation (4) are generated with a random number generator, except for β_0, β_1 and the threshold θ^1 (in case of a threshold model). The means of the estimates $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\theta}^1$ from the 1,000 simulations are then compared to the values β_0, β_1 and θ^1 chosen in the data generating process. We have fixed the VAR lag in the data generating process to one. Relaxing this restriction to more VAR lags is straight forward and yields the same results, however with lower precision due to the larger number of parameters to be estimated.

The detailed results of the 1-threshold TVECM model specification are presented in Table 1. Table 1 yields two immediate observations, the mean of $\hat{\beta}_1$ is practically independent of the grid setup and the outcomes of the bootstrap methodologies are also independent. The fixed regressor and residual bootstrap have very small β-errors, below 2%. As expected the standard deviation of $\hat{\beta}_1$ is getting larger for coarser grid settings. In other words, we find a clear dependence of the quality of the estimation precision on the grid size. The dynamic grid setting shows the best performance. Independent of the grid size, the mean values of $\hat{\beta}_0$ and $\hat{\theta}^1$ from our 1,000 simulations are imprecise, ie far-off from their theoretical values, fixed in the data generating process. No relationship between the grid setup and the precision can be inferred.

---

[7]   This is strictly correct only if β_1 is around 1.

Table 1: 2-regime TVECM $\mathcal{H}_2$ - estimation precision

The 2-regime TVECM model with one threshold θ^1 belongs to model class $\mathcal{H}_2$. The data generating process contains one threshold with the theoretical value 3. "Restricted" means that the underlying time series are restricted to the maximum value of 500 in absolute terms and must be I(1) while their basis must be I(0). The table shows the means and the standard deviations (figures in brackets) of 1,000 estimates $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\theta}^1$ for different grid settings. We also report how often the null hypothesis of a linear VECM model class $\mathcal{H}_1$ is incorrectly not rejected at different confidence levels (CL) by the fixed regressor and the residual bootstrap. "dyn" stands for the dynamic grid setting based on a minimum distance, while the rest are fixed, equidistant grid settings.

	grid	$\hat{\beta}_1$ (stdev)	$\hat{\beta}_0$ (stdev)	$\hat{\theta}^1$ (stdev)	fixed regressor β-error, CL			residual β-error, CL		
					90%	95%	99%	90%	95%	99%
theor. value		1.10	10	3	90%	95%	99%	90%	95%	99%
restricted	10	1.10	7.49	4.79	0.002	0.004	0.007	0.004	0.005	0.005
		(0.29)	(50.56)	(47.48)						
restricted	50	1.09	5.42	7.4	0.002	0.002	0.004	0.003	0.006	0.007
		(0.10)	(14.74)	(14.92)						
restricted	100	1.10	5.42	7.49	0.003	0.005	0.007	0.004	0.005	0.012
		(0.04)	(16.63)	(16.00)						
restricted	dyn	1.10	8.39	4.46	0.002	0.002	0.007	0.004	0.005	0.012
		(0.02)	(32.99)	(33.00)						
unrestricted	50	1.05	-1.67	14.28	0.002	0.005	0.008	0.002	0.004	0.010
		(0.84)	(111.16)	(121.17)						

However and most importantly, if we compute the sum of $\hat{\beta}_0$ and $\hat{\theta}^1$ for each simulation and compute the average and standard deviation, we find that the results of the point estimate of the sum $\hat{\beta}_0 + \hat{\theta}^1$, which are the estimated transaction costs, have a much higher estimation precision compared to the point estimates of the individual parameter estimates. The results are displayed in Table 2.

The best performance is again found for the dynamic grid setting for the important restricted case. This result is also graphically shown in Figure 4. This is a very positive result, which we have tested for arbitrary parameters, because for arbitrageurs the sum $\hat{\beta}_0 + \hat{\theta}^1$, which represents the transaction costs, is important and not the individual parameter estimates $\hat{\beta}_0$ and $\hat{\theta}^1$.

Table 2: 2-regime TVECM $\mathcal{H}_2$ - $\hat{\beta}_0 + \hat{\theta}^1$ estimation precision

The table shows for the same simulations as in Table 1 the sum $\hat{\beta}_0 + \hat{\theta}^1$, which corresponds to the transaction costs based on our no-arbitrage argument. The sum $\hat{\beta}_0 + \hat{\theta}^1$ is very close to the theoretical value and the standard deviation is small, with a clear dependence on the grid setup. We find the highest precision for the dynamic grid setup.

	grid	$\hat{\beta}_0 + \hat{\theta}^1$	stdev
theor. value		13	
restricted	10	12.28	4.37
restricted	50	12.82	3.14
restricted	100	12.91	1.33
restricted	dyn	12.85	1.17
unrestricted	50	12.25	7.75

Figure 4: 2-regime TVECM $\mathcal{H}_2$ - distribution of parameter estimates

The first three graphs show the distribution of the three estimates based on 1,000 MC simulations for the model in Equation (4) for the dynamic grid setup. The graph on the very right-hand side shows the sum $\hat{\beta}_0 + \hat{\theta}^1$, which represents the estimator of the transaction costs. The distribution of $\hat{\beta}_1$ is very narrow, whereas we find large outliers for the individual parameter estimates of $\hat{\beta}_0$ and $\hat{\theta}^1$. The distribution of $\hat{\beta}_0 + \hat{\theta}^1$ is again very narrow, with a similar relative error as for $\hat{\beta}_1$.
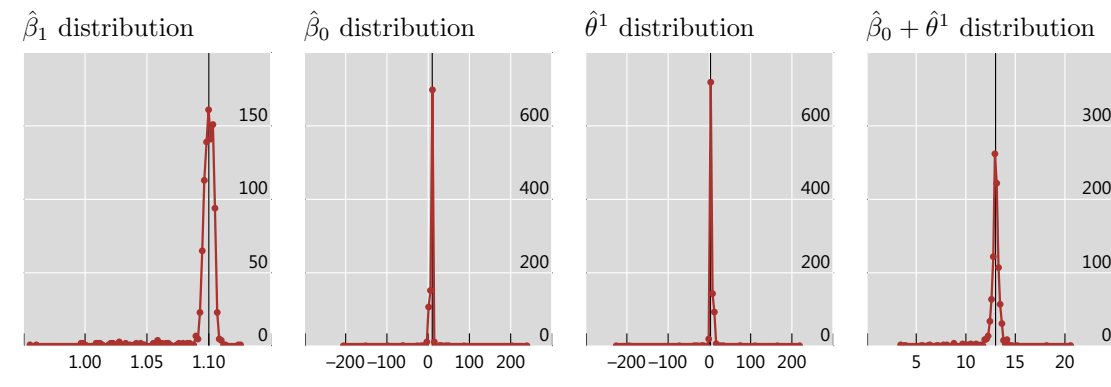
The graphs in Figure 4 contain, despite their similarity, several interesting features. As already shown in the Table 1, the means generated from our 1,000 simulations are close to the theoretically expected value. The range of the distribution is partly determined by the chosen grid. That is why the plot for the estimator of β_1 is shifted to the right, because we have calibrated the grid to include the value $\hat{\beta}_1 = 1$ (observed basis). For

the $\hat{\beta}_0$ and $\hat{\theta}^1$ estimators we find that the weight of the distribution is heavily centred around the values of the data generating process. However, we find several outliers far away from that theoretical value fixed in the data generating process. The most important and interesting finding is, that if we look at the very right graph in Figure 4, namely the sum of $\hat{\beta}_0 + \hat{\theta}^1$, we see that the dispersion is dramatically lower compared to its individual parameter estimates (two graphs in the middle). The individual parameter estimates of $\hat{\beta}_0$ and $\hat{\theta}^1$ have large outliers ranging from around -200 to +200.

The negative log-likelihood functions in Figure 5, which need to be minimized, show a non-smooth behaviour, hence the proposed grid search is necessary. The graphs give also an immediate explanation why the estimation of β_1 has a high degree of precision (assuming a sufficiently fine grid is chosen) and why the precision of the estimators of β_0 and θ^1 are poor.

Figure 5: 2-regime TVECM $\mathcal{H}_2$ - negative log-likelihood functions

The figures show the negative log-likelihood function for three out of our 1,000 simulations. It is obvious that the functions are non-differentiable. It also shows why we find a very high degree of precision for β_1 and a poor precision for β_0 and θ^1.
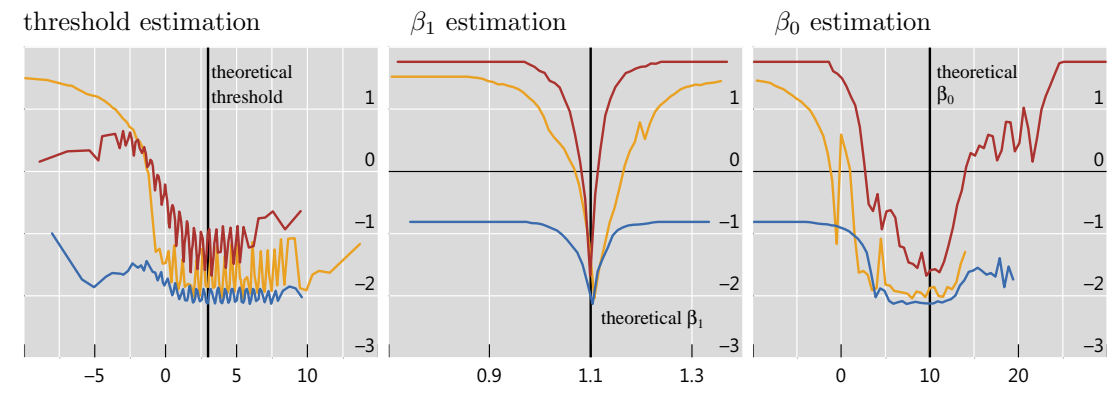
2.3.2 Testing for a second threshold in model class $\mathcal{H}_2$

As a next and important step we generalize the simulation to search and test for two thresholds. We are making one assumption, based on a purely economic viewpoint, namely that, we have at most one positive and/or one negative threshold, which correspond to the two transaction costs for a positive basis trade and a negative basis trade. This assumption relates to our no-arbitrage discussion in Section 1 which in turn links straight to the three classes of models $\mathcal{H}_i$ (see Figure 1). We want to test for the performance of the methodology when searching for a second threshold in a data generating process that only contains one threshold.

The previous analysis has shown that the dynamic grid point setting is the most precise methodology to estimate the transaction costs (see Table 1 for $\hat{\beta}_1$ and Table 2 for $\hat{\beta}_0 + \theta^1$), even though computationally more costly. Therefore, we will from now on use this grid setting only.

We use the same model parameters and the same model, a 1-threshold TVECM (model class $\mathcal{H}_2$), as in the previous subsection, however, test if our two bootstrap methodologies reject the existence of a second threshold. The data generating process has no second threshold, ie θ^2 does not exist and the estimate $\hat{\theta}^2$ should be distributed around zero.

Our previous simulation has shown that β_1 is estimated with a high degree of precision if we use a dynamic grid setup. In order to save computational time we fix β_1 during the second threshold search to the parameter value found in the first threshold search and hence reduce the search for the second threshold to a 2-dimensional grid search for β_0 and θ^2.

In order to improve the understanding of the reliability of the methodology for the second threshold search, we test different time series lengths as well as the restricted versus the unrestricted case. However, we always use the dynamic grid point setting. The results are presented in Table 3.

Table 3: 2-regime TVECM $\mathcal{H}_2$ - estimation precision (fixed β_1)

This table shows the parameter estimates for the threshold search for θ^1 and θ^2 as well as the theoretically expected values. All results are generated with a dynamic grid setup and β_1 is fixed in the threshold search for θ^2. The theoretical value for θ^2 is denoted with a "n/a" as the data generating process only contains one threshold.

		$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\theta}^1$	$\hat{\beta}_0 + \hat{\theta}^1$	$\hat{\beta}_0$	$\hat{\theta}^2$
theor. value		1.10	10	3	13	10	n/a
type	periods	(stdev)	(stdev)	(stdev)	(stdev)	(stdev)	(stdev)
restricted	1,000	1.10	8.39	4.46	12.85	8.64	-10.38
		(0.02)	(32.99)	(33.00)	(1.17)	(19.90)	(19.82)
restricted	2,000	1.10	9.35	3.57	12.93	8.09	-9.51
		(0.01)	(7.93)	(7.91)	(0.72)	(13.83)	(13.72)
unrestricted	1,000	1.10	11.05	2.00	13.05	16.29	-17.83
		(0.01)	(26.29)	(26.33)	(1.86)	(49.91)	(49.78)

As in the previous simulations, the estimators $\hat{\beta}_1$ as well as $\hat{\theta}^1 + \hat{\beta}_0$ are very precise, with the highest precision achieved in the restricted case and especially in the case where the tested time series has a length of 2,000. We use a dynamic grid setup, which is the reason

why we find similar standard deviations for the estimators $\hat{\beta}_0$ and $\hat{\theta}^1$ in the unrestricted case compared to the restricted case for a time series length 1,000. In contrast, we have tested fixed and dynamic grid points in Table 1 which results in much wider standard deviations of the unrestricted case compared to the figures of the restricted case. The estimator $\hat{\theta}^2$ is distributed around zero, as expected, because we do not have a second threshold in our data generating process.

The results of the reliability of the bootstrap algorithms for both threshold searches are given in Table 4. Again, we find that both tests, the fixed regressor and the residual bootstrap, produce consistent and precise estimates in terms of α- and β-errors.

The β-errors are very small, as expected from the results in Section 2.3.1. The α-errors of the second threshold search are higher, however still acceptable.

Table 4: 2-regime TVECM $\mathcal{H}_2$ - bootstrap precision

This table shows the β-errors for the threshold search for θ^1 and the α-errors of the threshold search for θ^2. We use the same ordering of model setups as in the previous Table 3.

first threshold bootstrap						second threshold bootstrap					
fixed regressor			residual			fixed regressor			residual		
90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.002	0.002	0.007	0.004	0.005	0.012	0.21	0.15	0.07	0.25	0.15	0.06
0.003	0.003	0.006	0.004	0.004	0.014	0.35	0.30	0.23	0.36	0.30	0.18
0.000	0.000	0.000	0.001	0.001	0.001	0.23	0.16	0.09	0.27	0.16	0.09

In order to test the robustness of the method, where we keep β_1 fixed in the threshold search for θ^2, we repeat the simulation and search now also for an optimal β_1 in the threshold search for θ^2 using the log-likelihood method. The threshold search for θ^1 is performed in the usual manner, hence we only need to present the figures for the threshold search for θ^2. As an illustration, we focus on the restricted case and a time series length of 1,000 in our data generating process. The results presented in Table 5 lead to two immediate observations: firstly, the estimator $\hat{\beta}_1$ is dramatically less precise compared to the results with a fixed β_1 (Table 3) and the α-errors are very similar to the ones presented in the first row and the last six columns of Table 4.

We can conclude that the performance of the tests is higher for the second threshold search if β_1 is kept fixed to the value found in search for θ^1. This also reduces the computational burden.

Table 5: 2-regime TVECM $\mathcal{H}_2$ - estimation precision (β_1 variable)

This table shows the parameter estimates for the threshold search for θ^2 as well as the theoretically expected values. All results are generated with a dynamic grid setup and this time β_1 is kept variable, ie a grid search is applied also to β_1 while searching for θ^2. The theoretical value for θ^2 is denoted with "n/a" as the data generating process only contains one threshold.

	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\theta}^2$	fixed regressor			residual		
	(stdev)	(stdev)	(stdev)	α-error, CL			α-error, CL		
theor. value	1.10	10	n/a	90%	95%	99%	90%	95%	99%
restricted	1.02	8.74	-9.94	0.28	0.19	0.09	0.28	0.19	0.09
	(0.13)	(17.25)	(17.22)						

2.4   Model class $\mathcal{H}_3$: 3-regime TVECM

Finally, we present the results for a data generating process with two thresholds. Without loss of generality we present results for the following choice of parameters: $\beta_1 = 1.1$, $\beta_0 = 1$ as well as two thresholds $\theta^1 = -4$ and $\theta^2 = 6$. We use again the dynamic grid point method and a time series of length 1,000 as the baseline setup. Further, as part of our baseline case, we keep β_1 for the threshold search for θ^2 fixed, as we have shown in Section 2.3.1 that this produces best results. However, we perform a variety of further robustness checks, such as we extend the time series length to 5,000, compute a case with a very short time series length of 100, analyse the unrestricted case, as well as test results where we keep β_1 variable in the threshold search for θ^2. In addition, we vary the precision parameter settings in the dynamic grid search. As previously, we use the standard setting $\Delta\beta_0 = 0.5$, $\Delta\beta_1 = 0.01$ and $\Delta\theta^j = 0.5$, but also consider cases where $\Delta\beta_0 = 1$ or 0.1, $\Delta\beta_1 = 0.1$ or 0.005 and $\Delta\theta^j = 1$ or 0.1. We will start with the presentation of the results for the threshold search for θ^1 and proceed with the threshold search for θ^2. Depending on the generated time series either of the two thresholds may be found in the first round. Table 6 shows that the estimation of β_1 is again very precise, whereas the estimate of β_0 is poor. For the estimation of β_1, we find a clear improvement of quality with increasing time series length and smaller grid spacing. The estimate of β_0 shows a large standard deviation, which is not dependent on the time series length and the grid spacing. The threshold estimate is poor, for two reasons, firstly, we know from the one threshold case that the estimation of the threshold is imprecise and secondly, in the two threshold case the estimation procedure may find either the negative or the positive threshold. Again, the most important finding is, that the sum of $\hat{\beta}_1$ and the threshold is for the negative values very close to $\hat{\beta}_0 + \hat{\theta}^1 = -3$ and for the positive values very close to $\hat{\beta}_0 + \hat{\theta}^2 = 7$. We

find a clear dependence of the quality of the estimator on the time series length and the grid spacing.

Table 6: 3-regime TVECM $\mathcal{H}_3$ - first threshold search estimation precision

This table shows the results of the first threshold search for our 2-threshold respective 3-regime TVECM. The point estimates of θ^i are imprecise, simply because it is possible to find either of the two thresholds in the first search, -4 or 6. The estimates of the two transactions costs (positive basis trade and negative basis trade) show again a high estimation precision.

type	periods	precision	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\theta}$	$\hat{\beta}_0 + \hat{\theta}^1$	$\hat{\beta}_0 + \hat{\theta}^2$
		theor. value	1.10	1	-4 or 6	-3	7
		precision	(stdev)	(stdev)	(stdev)	(stdev)	(stdev)
restricted	100	0.01/0.5	1.09	0.53	1.85	-2.98	6.44
			(0.08)	(16.09)	(17.00)	(0.95)	(1.78)
restricted	1,000	0.1/1	1.10	-0.89	3.00	-3.35	6.89
			(0.04)	(16.84)	(18.14)	(2.16)	(2.54)
restricted	1,000	0.01/0.5	1.10	0.08	2.43	-2.98	6.94
			(0.02)	(6.15)	(8.69)	(0.65)	(1.59)
restricted	1,000	0.005/0.1	1.10	-0.24	2.69	-2.94	6.86
			(0.01)	(5.89)	(8.43)	(0.34)	(0.89)
restricted	5,000	0.01/0.5	1.10	-0.10	2.65	-3.00	6.75
			(0.01)	(21.03)	(21.90)	(0.89)	(1.33)
unrestricted	1,000	0.01/0.5	1.10	-1.02	4.00	-3.16	6.82
			(0.01)	(17.47)	(18.38)	(1.51)	(1.83)

The power of the two bootstrap tests is very high, as expected from the previous simulation results (see Table 7), except for the case where the tested series have a length of 100 data points. Even though the tests have still an acceptable power in this case, it is visibly worse compared to the test statistics where we consider time series with a length of 1,000 and more observations.

The figures for the second threshold search, presented in Table 8, are extremely promising, because the two transaction costs are well estimated. In addition to the simulations presented in Tables 6 and 7, we add for the second threshold search also the simulation, where β_1 is kept variable in the search for the second threshold. In this simulation we do not fix β_1 to the value found in the threshold search for θ^1, but perform a grid search to find the "optimal" value based on the maximum likelihood method.

We find in Table 8 a similar picture as in the first threshold search, however as expected with slightly bigger standard deviations. The threshold parameter estimate is poor. However, the sum of $\hat{\beta}_0$ and the threshold is for the negative values very close to $\hat{\beta}_0 + \hat{\theta}^1 = -3$ and for the positive values very close to $\hat{\beta}_0 + \hat{\theta}^2 = 7$. In general, we find higher precisions for longer time series length and for smaller grid spacing.

The β_1 estimation in the variable case has a low degree of precision (last simulation in Table 8). This result is in line with previous simulation findings from model class $\mathcal{H}_2$.

The bootstrap precisions are displayed in Table 9. We find that the test performance of our proposed methodology is generally satisfying, however as expected not as high as in the search for the first threshold (see Table 7). The power of the tests increases with the time series length.

Table 7: 3-regime TVECM $\mathcal{H}_3$ - bootstrap precision for first threshold search

This table shows the β-errors of the two bootstrap methods. The parameter estimates can be found in the previous Table 6.

| | | | fixed regressor | | | residual | | |
| | | | β-error, CL | | | β-error, CL | | |
type	periods	precision	90%	95%	99%	90%	95%	99%
restricted	100	0.01/0.5	0.018	0.028	0.069	0.017	0.036	0.099
restricted	1000	0.1/1	0.000	0.001	0.001	0.001	0.001	0.002
restricted	1000	0.01/0.5	0.000	0.001	0.001	0.000	0.000	0.002
restricted	1000	0.005/0.1	0.000	0.000	0.000	0.001	0.001	0.002
restricted	5000	0.01/05	0.000	0.000	0.000	0.000	0.000	0.000
unrestricted	1000	0.01/0.5	0.000	0.001	0.002	0.001	0.001	0.004

The negative log-likelihood function for 3 different sample simulations is shown in Figure 6. The negative log-likelihood function for the threshold estimation is now more complex, reflecting the fact that we have two thresholds. The functional form shows immediately the non-differentiable structure. Further, it is again evident, why the individual parameter estimates of the threshold and β_0 are imprecise. The estimation of β_1 is as in the previous cases very accurate.

Table 8: 3-regime TVECM $\mathcal{H}_3$ - second threshold search estimation precision

This table shows the results of the second threshold search for our 2-threshold TVECM, respective 3-regime TVECM. The point estimate of θ is imprecise, simply because we have a chance to find either of the two thresholds (-4 or 6). In addition to the first threshold search presented in Table 6 we also include a simulation where we keep β_1 variable (last two rows), ie where we have not fixed β_1 to the estimate found in the first threshold search. This simulation is denoted by the appreciation "var" for variable. The precision of the estimates of the two transactions costs is again very high. β_1 is not reestimated, except for the simulations in the last two rows, hence we filled in "n/a".

			$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\theta}$	$\hat{\beta}_0 + \hat{\theta}^1$	$\hat{\beta}_0 + \hat{\theta}^2$
		theor. value	1.10	1	-4 or 6	-3	7
type	periods	precision	(stdev)	(stdev)	(stdev)	(stdev)	(stdev)
restricted	100	0.01/0.5	n/a	0.62	0.13	-3.32	6.10
				(8.18)	(10.44)	(2.46)	(3.00)
restricted	1,000	0.1/1	n/a	0.53	0.55	-2.99	5.60
				(15.53)	(16.85)	(3.00)	(2.90)
restricted	1,000	0.01/0.5	n/a	0.84	0.43	-2.91	6.28
				(4.28)	(7.56)	(1.80)	(2.22)
restricted	1,000	0.005/0.1	n/a	0.99	0.25	-2.96	6.42
				(6.43)	(9.25)	(2.48)	(2.04)
restricted	5,000	0.01/0.5	n/a	2.36	-1.07	-2.63	6.34
				(19.75)	(20.90)	(1.84)	(2.10)
unrestricted	1,000	0.01/0.5	n/a	1.65	-1.48	-3.44	5.94
				(18.51)	(19.97)	(4.15)	(2.82)
restricted	1,000/var	0.01/0.5	1.10	1.06	-0.21	-2.95	5.55
			(0.05)	(8.37)	(10.52)	(2.09)	(2.63)

The distribution of the estimators $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\theta}^i$ and the transaction costs for the first and second threshold search are presented in Figure 7. The distribution of $\hat{\beta}_0$ and the thresholds is wide in line with the form of the negative likelihood function in Figure 6. In each search we can either find θ^1 or θ^2, which cannot be distinguished due to the imprecise estimation. However, looking at the transaction costs (right-hand panel), which are estimated at a good precision, we can clearly unravel the two different transaction costs.

Table 9: 3-regime TVECM $\mathcal{H}_3$ - bootstrap precision for the second threshold search

This table shows the β-errors of the two bootstrap methods. The parameter estimates can be found in the previous Table 8.

			fixed regressor β-error, CL			residual β-error, CL		
type	periods	precision	90%	95%	99%	90%	95%	99%
restricted	100	0.01/0.5	0.053	0.070	0.102	0.052	0.073	0.113
restricted	1,000	0.1/1	0.015	0.016	0.021	0.021	0.021	0.027
restricted	1,000	0.01/0.5	0.007	0.008	0.009	0.006	0.007	0.009
restricted	1,000	0.005/0.1	0.015	0.017	0.020	0.015	0.016	0.017
restricted	5,000	0.01/0.5	0.008	0.009	0.010	0.008	0.009	0.012
unrestricted	1,000	0.01/0.5	0.015	0.016	0.018	0.014	0.019	0.021
restricted	1,000/var	0.01/0.5	0.011	0.011	0.013	0.009	0.010	0.014

Figure 6: 3-regime TVECM $\mathcal{H}_3$ - negative log-likelihood function for the second threshold

The figures show the negative log-likelihood functions for three sample simulations out of our 1,000 simulations for the second threshold search. The functions are non-differentiable. A high degree of precision can only be expected for β_1. We find evidence for a low individual precision of β_0 and θ^i. The two vertical lines in the left-hand graph represent the two thresholds θ^1 and θ^2.

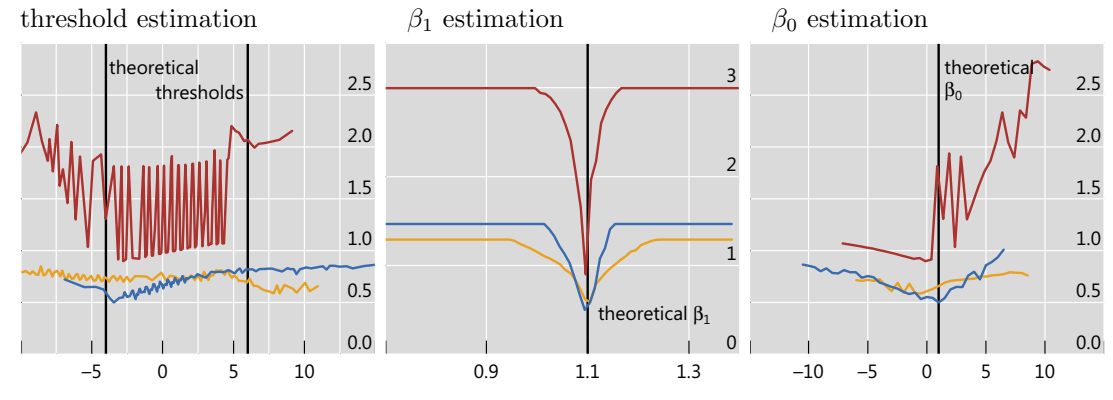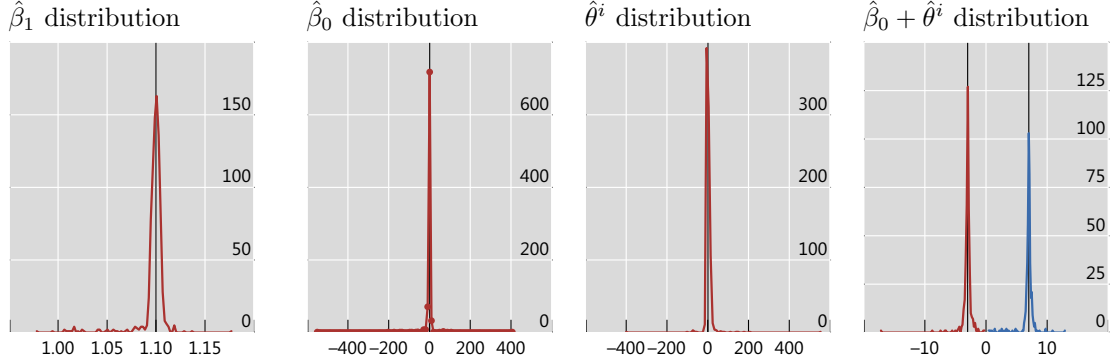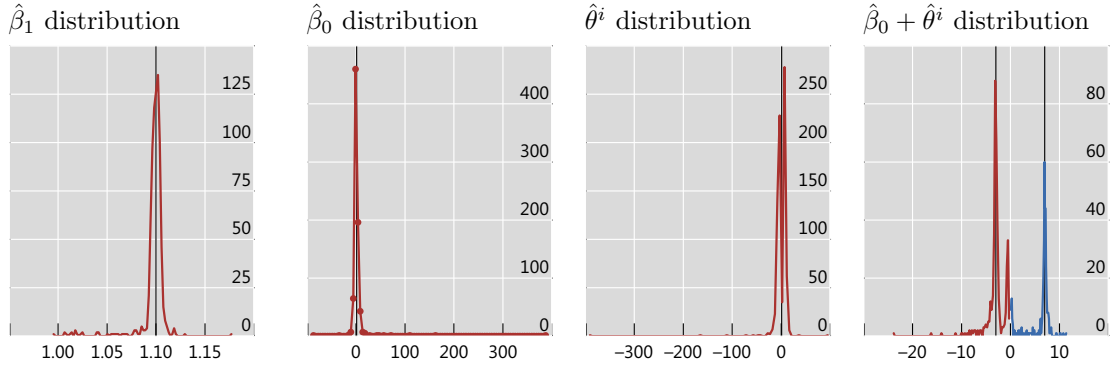threshold estimation β_1 estimation β_0 estimation

26

Figure 7: 3-regime TVECM $\mathcal{H}_3$ - distribution of TVECM estimates

The first three graphs in the upper and lower panel show the distribution of the three estimates based on 1,000 MC simulations for our model in Equation (18). The graph on the very right shows the estimation of the transaction costs $\hat{\beta}_0 + \hat{\theta}^i$. The distribution of $\hat{\beta}_1$ is very narrow, whereas we find large outliers for $\hat{\beta}_0$ and $\hat{\theta}^i$. The distribution of $\hat{\beta}_0 + \hat{\theta}^i$ is also very narrow, with a similar relative error as for $\hat{\beta}_1$.

Panel A: first search

$\hat{\beta}_1$ distribution $\hat{\beta}_0$ distribution $\hat{\theta}^i$ distribution $\hat{\beta}_0 + \hat{\theta}^i$ distribution

Panel B: second search

$\hat{\beta}_1$ distribution $\hat{\beta}_0$ distribution $\hat{\theta}^i$ distribution $\hat{\beta}_0 + \hat{\theta}^i$ distribution

2.4.1 Conclusion from the simulation exercise

The comprehensive simulation exercise has revealed the following outcomes: The method discussed in Sections 2.1 and 2.2 leads to stable and robust results for the 1- and 2-threshold TVECM. The transaction costs $\theta^i + \beta_0$ as well as β_1 can be estimated with high accuracy, whereas the distribution of $\hat{\theta}^i$ and $\hat{\beta}_0$ is very wide. The dynamic grid search strategy, where the grid size is determined with an accuracy parameter (distance between two grid points of the equidistant grid), has turned out to be superior. The β_1 can be fixed in the second threshold search to the value found in the first search. This yields best results and reduces the computational burden dramatically.

3 Application: Index trading and Commodity arbitrage

We apply our proposed 3-regime TVECM methodology to different markets that exhibit a non-zero basis (either in contango or backwardation) for two similar financial market instruments traded in the spot and the derivatives market.

We empirically estimate unknown transaction costs for the S&P 500 and for palladium by analysing the basis defined as the spot price minus the price of the future on the same underlying.

In our application we present results for our dynamic grid setup, for different periods and we run the model for 1 lag up to 5 lags in the VAR term. We have used the Schwarz information criterion (SIC) to determine the model with the best fit. We furthermore have also performed unit root, stationarity and cointegration tests. We find cointegration for the spot and futures prices of the S&P 500 and palladium, while the individual time series are I(1).

Further as an additional illustration of the usage of (T)VECM models we introduce briefly two measures of price finding and the terminology of half-life of shock absorption. From the speed of adjustments in Equation (1) we can compute the HAS and GG measures of price discovery. The two relevant and independent Hasbrouck (HAS) measures (Man et al.; 2013) are defined for one individual regime as follows:

$$\text{HAS}_1 = \frac{\lambda_2^2\left(\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}\right)}{\lambda_2^2\sigma_1^2 - 2\lambda_1\lambda_2\sigma_{12} + \lambda_1^2\sigma_2^2} \qquad \text{and} \qquad \text{HAS}_2 = \frac{\left(\lambda_2\sigma_1 - \lambda_1\frac{\sigma_{12}}{\sigma_1}\right)^2}{\lambda_2^2\sigma_1^2 - 2\lambda_1\lambda_2\sigma_{12} + \lambda_1^2\sigma_2^2}. \quad (22)$$

To estimate the Equations (22) we rely on the estimated covariance matrix from the VECM to capture the terms σ_1^2, σ_{12} and σ_2^2. In the following we define HAS as the average of HAS_1 and HAS_2. For simplicity, we have skipped the superscripts indicating the regime.

The second indicator for price discovery, the GG measure (Gonzalo and Granger; 1995) decomposes the common factor itself, but ignores the correlation of the innovations in the two markets. The following two measures exist

$$\text{GG}^{\text{spot}} = \frac{-\lambda_2}{\lambda_1 - \lambda_2} \qquad \text{and} \qquad \text{GG}^{\text{futures}} = \frac{\lambda_1}{\lambda_1 - \lambda_2}, \quad (23)$$

whereby it is obvious that $\text{GG}^{\text{spot}} + \text{GG}^{\text{futures}} = 1$.[8] Therefore, we will use GG^{spot} only and skip the superscript.

---

[8] Unlike the HAS measure, the GG measure is not mathematically confined to the closed interval [0,1] which seems to make an interpretation similar to the Hasbrouck measure difficult. GG measures below 0 and above 1 should be interpreted as 0 and 1, respectively (see Gyntelberg et al. (2013) for more discussion).

HAS and GG measures greater than 0.5 imply that more than 50% of the price discovery occurs in the spot market. When the measures are close to 0.5 both markets contribute to price discovery without evidence on which market is dominant. GG and HAS below 0.5 suggest price leadership of the futures market.

Finally, we are interested in examining the speed of adjustment towards the long-term equilibrium. As the spot and futures time series in the bivariate VECM share a common stochastic trend, the impulse response function for the cointegrating residual can be used to determine the speed of adjustment to the long-run equilibrium (Zivot and Wang (2006)). The vector error correction mechanism directly links the speed of adjustments to the cointegration residual u_t which follows an implied AR(1) process:

$$u_t = (1 + \lambda_1 - \beta_1 \lambda_2) u_{t-1} + \varepsilon_t^{\text{spot}} - \beta_1 \varepsilon_t^{\text{futures}} \equiv \phi u_{t-1} + \varepsilon_t^{\text{spot}} - \beta_1 \varepsilon_t^{\text{futures}} . \qquad (24)$$

The half-life of a shock, n, can now be calculated from the AR(1) coefficient ϕ as:

$$n = \frac{ln(0.5)}{ln(\phi)} . \qquad (25)$$

Regulators can use the model in a variety of different ways. They can study which market moves first, how large are transactions costs and how long does it take for the markets to return to the long-run equilibrium after an exogenous shock. Changes in the speed of adjustment or half-life might be especially interesting, after regulatory changes have been introduced, in order to study the impact of the implemented changes on market participants behaviour and also on relative or absolute liquidity of both markets. In markets, where high frequency data is available (eg. index-trading), a pre and post-event analysis could be performed to evaluate the effectiveness of policy actions.

3.0.1   S&P 500

As an illustration we present the analysis of the Standard & Poor's 500, abbreviated as the S&P 500 index, and its futures index for the period starting in 2000. The S&P 500 is an American stock market index based on the market capitalizations of 500 large companies having common stock listed on the New York stock exchange or NASDAQ. The index and the future prices, as well as the basis are presented in Figure 8.

Figure 8: S&P 500 spot, futures price and spot-future basis

The left-hand graph shows the S&P 500 index and its futures for the period starting in 2000. The prices are given in index points. The right-hand graph shows the observed basis (spot-future), as well as the two transaction costs for a positive and a negative basis trade. Source: Bloomberg, authors' calculations.

Prices

Basis
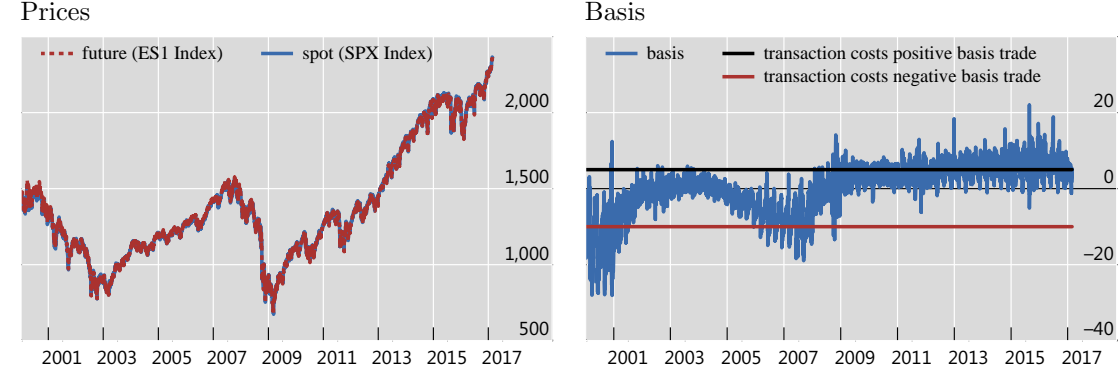
We find evidence for two thresholds within the considered period. Detailed results can be found in Tables 10 (threshold search for θ^1) and 11 (threshold search for θ^2).

Table 10: Threshold search for θ^1 - S&P 500

The table shows the results for the threshold search for θ^1 for the period April 2000 until end February 2017. The estimated transaction costs are significant at lag=3, for which the SIC also has a local minimum. The transaction costs are around -10 index points. The superscript *, ** and *** denotes the 90%, the 95% and 99% CL respectively. λ_i^j are the regime dependent speed of adjustments from Equation (1) with the subscript L as the lower regime and U as the upper regime.

		Bootstrap								%
lag	SIC	fixed	residual	$\hat{\beta}_1$	$\hat{\beta}_0 + \hat{\theta}^1$	λ_{1L}^1	λ_{2L}^1	λ_{1U}^1	λ_{2U}^1	lower regime
1	7.47	0.00	0.00	1.007	-8.18	0.02	0.06*	0.37**	0.68***	44
2	7.43	0.00	0.01	0.998	-1.78	0.02	0.09	0.12*	0.17***	21
3	7.38	0.04	0.07	1.001	-10.16	-0.01	0.04	0.17***	0.23***	10
4	7.39	0.05	0.11	1.003	-12.40	-0.04	0.01	0.21***	0.27***	11
5	7.37	0.02	0.06	1.003	-12.21	-0.01	0.04	0.19***	0.26***	11

The threshold search for θ^1 estimates transaction costs for a negative basis trade of around $\hat{\beta}_0 + \hat{\theta}^1 = -10$ index points. The transaction costs for a negative basis trade have to be interpreted in absolute terms. The local minimum of the SIC is at lag=3, for which

the threshold is also significant. The estimated transaction costs are very stable, except for lag 2.

The results for the threshold search for θ^2 are shown in Table 10. No results are presented for lag=1 as the 1-threshold TVECM is a better model fit compared to the 2-threshold TVECM model.

Table 11: Threshold search for θ^2 - S&P 500

The table shows the results for the threshold search for θ^2 for the period April 2000 until end February 2017. We do not show the result for one lag, as a 1-threshold TVECM is the best model fit. The local minimum value of the SIC is found at three VAR lags, for which the threshold is also significant. The transaction costs for a positive basis trade are around 5 index points. The superscript * and ** denotes the 90% and 95% CL respectively. λ_i^j are the regime dependent speed of adjustments from Equation (1) with the subscript L as the lower regime and U as the upper regime.

		Bootstrap								%
lag	SIC	fixed	residual	$\hat{\beta}_1$	$\hat{\beta}_0 + \hat{\theta}^2$	λ_{1L}^2	λ_{2L}^2	λ_{1U}^2	λ_{2U}^2	upper
										regime
2	7.44	0.09	0.11	0.998	9.24	0.05	0.09**	0.07	0.15	14
3	7.40	0.08	0.11	1.001	5.38	0.04	0.08*	0.08	0.18	11
4	7.41	0.68	0.70	1.003	2.30	0.03	0.07	0.28	0.40	11
5	7.39	0.90	0.93	1.003	2.24	0.04	0.09*	0.20	0.29	11

Similar to the threshold search for θ^1, we find a local minimum of the SIC at 3 lags, for which the threshold is significant. This leads to transaction costs of a positive basis trade of around 5 index points.

Further, we show price discovery measures for the different regimes at three lags in Table 12.

Table 12: Measures of price discovery - S&P 500

The table shows price discovery measures for lag=3. The GG measures with the superscript $^+$ should be interpreted as 1. The price discovery measures of the neutral regime are means of the estimation from the first and second threshold search.

regime	HAS	GG	leading market	half-life (in days)
upper	0.53	1.72$^+$	spot	6.6
neutral	0.56	2.90$^+$	spot	14.05
lower	0.51	0.82	spot	13.50

The spot market is leading in the two arbitrage regimes and as expected, the half-life in the neutral regime is the longest. The HAS, GG and half-lives for the upper and lower regime in Table 12 are based on the point estimates of insignificant speed of adjustments. This seems to be contradicting to the general notion, that speed of adjustments must be unequal to zero in the extreme regimes (arbitrage regimes). However, the upper and lower regimes contain only around 10% of the data points and therefore it is challenging to achieve statistical significance. This issue could be avoided by using intraday data (see for example Gyntelberg et al. (2013) and Gyntelberg et al. (2017)).

3.1 Palladium

Palladium is a precious metal, which has similar properties like Platinum and belongs to the same chemical group. It is used as an investment and industrial commodity. Its main industrial use is in catalytic converters.

We have analysed several periods and present the period starting in 2000 for which we have found evidence for two thresholds, resulting in the two transaction costs, for a negative basis trade $\hat{\beta}_0 + \hat{\theta}^1 = -5.7$USD per contract size of 100troy oz and a positive basis trade $\hat{\beta}_0 + \hat{\theta}^2 = 1.1$USD per contract size of 100troy oz for a model with one VAR lag. The transaction costs for the negative basis trade have to be interpreted in absolute terms. The time series of the spot, futures as well as the basis are graphically presented in Figure 9. The two transactions costs are shown as horizontal lines in the right-hand panel of that figure.

The estimation procedure shows the model fit based on SIC is not straight forward as the value of the SIC is declining with a weak slope, reaching a plateaux at around 7 lags before it declines further with increasing number of lags. The most parsimonious model suggests the existences of two thresholds with high significance (see Tables 13 and 14) for the period starting in 2000. Apart from lag=2 the transaction costs for a negative basis trade are stable and around $\hat{\beta}_0 + \hat{\theta}^1 = -5.7$USD for a contract size of 100troy oz which have to be interpreted in absolute terms. The lower regime is an arbitrage regime, which as expected contains only 11% of all observations.

Figure 9: Palladium spot, future prices and spot-future basis

The figure on the left-hand side shows the palladium spot and the future prices. The prices are given in USD for a contract size of 100troy oz. The figure on the right-hand side shows the basis as well as the two transaction costs for a positive and a negative basis trade. Source: Bloomberg, author's calculations
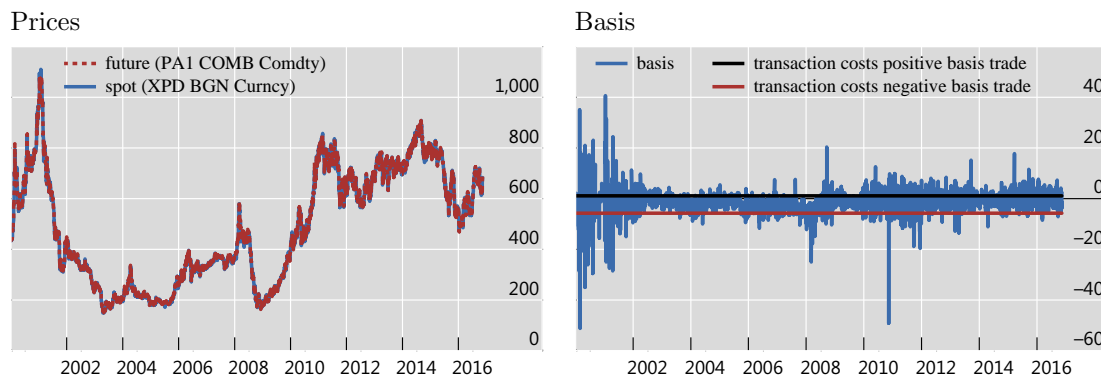
Table 13: Threshold search for θ^1 - Palladium

The table shows the results for the threshold search for θ^1 for the period April 2000 until end December 2016. It is difficult to determine the best model fit as the SIC is declining until it reaches a local minimum at lag =7 before it declines even further. The most parsimonious model has a significant threshold. The transaction costs $\beta_0 + \theta^1$ for a negative basis trade is -5.7USD. The superscript ** and *** denote the 95% and the 99% CL respectively. λ_i^j are the regime dependent speed of adjustments from Equation (1) with the subscript L as the lower regime and U as the upper regime.

| | | Bootstrap | | | | | | | | % |
lag	SIC	fixed	residual	$\hat{\beta}_1$	$\hat{\beta}_0 + \hat{\theta}^1$	λ_{1L}^1	λ_{2L}^1	λ_{1U}^1	λ_{2U}^1	lower regime
1	7.22	0.06	0.08	1.002	-5.75	-0.14	0.60***	-0.37***	0.28***	11
2	7.19	0.44	0.52	1.003	0.67	-0.18	0.39***	-0.29***	0.39***	88
3	7.16	0.67	0.74	1.003	-6.36	0.12	0.50**	-0.19***	0.31***	10
4	7.08	0.28	0.34	1.002	-5.66	0.20	0.51**	-0.17	0.25**	10
5	7.07	0.62	0.69	1.002	-5.77	0.15	0.54**	-0.19	0.16**	11

The threshold search for θ^2 finds a significant threshold at lag 1 for a positive basis trade. The estimated transaction costs are $\hat{\beta}_0 + \hat{\theta}^2 = 1.1$USD for a contract size of 100troy oz. We also find that the upper regime appears to be an arbitrage regime as it contains

12% of the overall observations. All other VAR specifications beyond lag 1 lead to a 1-threshold TVECM as a better model choice over the 2-threshold TVECM. However, as the lag=1 setup is the most parsimonious model and the threshold is highly significant we have strong evidence for a 2-threshold TVECM. The results are presented in Table 14.

Table 14: Threshold search for θ^2 - Palladium

The table shows the results for the threshold search for θ^2 for the period April 2000 until end December 2016. It turns out that for lag>1 the threshold search for θ^2 is not successful, because the 1-threshold TVECM is a better model fit. The superscript ** and *** denotes the 95% and the 99% CL, respectively. λ_i^j are the regime dependent speed of adjustments from Equation (1) with the subscript L as the lower regime and U as the upper regime.

lag	SIC	Bootstrap fixed	Bootstrap residual	$\hat{\beta}_1$	$\hat{\beta}_0 + \hat{\theta}^2$	λ_{1L}^2	λ_{2L}^2	λ_{1U}^2	λ_{2U}^2	% upper regime
1	7.23	0.02	0.03	1.002	1.14	-0.28***	0.44***	-0.49***	0.30**	12

In addition to the estimated transaction cost we have computed measures of price discovery for the relevant lag=1, which are presented in Table 15.

Table 15: Measures of price discovery - Palladium

The table shows price discovery measures for lag=1. The price discovery measures of the neutral regime are means of the estimation from the first and second threshold search.

regime	HAS	GG	leading market	half life (in days)
upper	0.48	0.38	futures	0.44
neutral	0.46	0.56	-	0.60
lower	0.56	0.81	spot	0.51

In the upper regime we find some indication that the futures market leads the price discovery process, whereas in the lower regime there is leadership of the spot market. Both markets seem to absorb one unit shocks rather quickly. The neutral regime has as expected a longer half life, however the difference is not very large. Furthermore, as expected we have no clear indication as to which market leads the price finding in the neutral regime.

4   Conclusion

The identification of thresholds in non-linear vector correction models is of rather complex origin with several unresolved problems. In this paper, we present the solution for one unresolved issue which is the estimation of a 3-regime TVECM with an unknown cointegrating vector. Our proposed methodology extends the 2-regime TVECM model as proposed by Hansen and Seo (2002). In contrast to Hansen and Seo (2002) we also introduce an intercept β_0 in the cointegrating relation $(S - \beta_1 D - \beta_0)_t$ to account for a distorted parity relationship, eg the non-zero basis. Using a sequential grid search for the first and the second threshold search, we estimate the cointegration relationship as well as the two thresholds (θ^1 and θ^2) by employing a maximum likelihood approach. As there are practically no empirical studies for TVECMs even for the 1-threshold case we present a comprehensive simulation study to investigate the reliability of our proposed method. Hansen and Seo (2002) have suggested a grid point search to estimate the variables β_i and θ^i with a fixed number of grid points. We show that a dynamic grid point setting, where the distance between two grid points is set via a precision parameter, instead of defining a fixed number of grid points, yields the best results, albeit at potentially high computational costs. The slope β_1 in the cointegrating relationship is estimated at a very high degree of precision. We show that the value $\hat{\beta}_1$ found in the first threshold search can be fixed in the second search. This lowers the dimension of the grid space and hence reduces the computing time in the second grid search. The thresholds θ^i and the shift in the cointegration relation β_0 are estimated poorly, but the sum $\theta^i + \beta_0$ is estimated with a very high degree of precision.

Our proposed methodology is particularly appealing for the analysis of distorted parity relationships in economics, such as no-arbitrage relationships with a non-zero basis. A persistent non-zero basis between two similar financial market instruments traded in the spot and in the derivative markets points towards the presence of transaction costs on arbitrage trades that prevent a complete adjustment of market prices to the theoretical no-arbitrage condition of a zero basis. In our applications, we show that our proposed 3-regime TVECM methodology can estimate unknown transaction costs on arbitrage trades (sum of $\theta^i + \beta_0$) with a very high degree of precision. Further, we have computed measures of price discovery as well as half-lives of shocks. These measures could help regulators evaluate which markets are more dynamic, and hence provide them with a better picture of market activity.

References

Adams, K. and van Deventer, D. (1993). Comment on intra-day arbitrage opportunities and price behavior of the Hang Seng index futures, *Review of Futures Markets* .

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point, *Econometrica* **61**: 821–1414.

Bai, J. (1997). Estimating multiple breaks one at a time, *Econometric Theory* **13**(3): 315–352.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes, *Econometrica* **66**(1): 47–78.

Baillie, R. T., Booth, G. G., Tse, Y. and Zabotina, T. (2002). Price discovery and common factor models, *Journal of Financial Markets* **5**: 309–321.

Balke, N. S. and Fomby, T. B. (1997). Threshold cointegration, *International Economic Review* **38**(3): 627–645.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika* **74(1)**: 33–43.

Fama, E. F. and French, K. R. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage, *The Journal of Business* **60**: 55–73.

Forbes, C. S., Kalb, G. R. J. and Kofhian, P. (1999). Bayesian arbitrage threshold analysis, *Journal of Business & Economic Statistics* **17**: 364–372.

Gonzalo, J. and Granger, C. W. J. (1995). Estimation of common long-memory components in cointegrated systems, *Journal of Business & Economic Statistics* **13**(1): 27–36.

Gyntelberg, J., Hoerdahl, P., Ters, K. and Urban, J. (2013). Intraday dynamics of euro area sovereign cds and bonds, *BIS Working Papers 423*, Bank for International Settlements.

Gyntelberg, J., Hoerdahl, P., Ters, K. and Urban, J. (2017). Arbitrage costs and the persistent non-zero cds-bond basis: Evidence from intraday euro area sovereign debt markets, *BIS Working Papers 631*, Bank for International Settlements.

Hansen, B. E. (1996). Inference when an nuisance parameter is not identified under the null hypothesis, *Econometrica* **64**: 413–430.

Hansen, B. E. (2000). Testing for structural change in conditional models, *Journal of Econometrics* **97**: 93–115.

Hansen, B. E. and Seo, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models, *Journal of Econometrics* **110**(2): 293–318.

Johansen, S. (1988). Statistical analysis of cointegrated vectors, *Journal of Economic Dynamics and Control* **12**: 231–254.

Lien, D. and Yang, L. (2008). Asymmetric effect of basis on dynamic futures hedging: Empirical evidence from commodity markets, *Journal of Banking and Finance* **32**: 187–198.

Lütkepohl, H. (2006). *New Introduction to Multiple Time Series Analysis*, Springer.

Man, K., Wang, J. and Wu, C. (2013). Price discovery in the u.s. treasury market: Automation versus intermediation, *Management Science* **59**: 695–714.

McMillan, D. G. (2005). Threshold adjustment in spot-futures metal prices, *Applied Financial Economics Letters* **1**: 5–8.

Stevens, J. (2015). Do transaction costs prevent arbitrage in the market for crude oil? Evidence from a threshold autoregression, *Applied Economic Letters* **22**: 169–172.

Sutcliffe, C. M. S. (2006). *Stock Index Futures*, Ashgate Publishing.

Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes, *Journal of the American Statistical Association* **84**(405): 231–240.

Zivot, E. and Wang, J. (2006). *Modeling Financial Time Series with S-PLUS*, Springer-Verlag New York, Inc.

Previous volumes in this series

All volumes are available on our website www.bis.org.