



On Using Markov Switching Time Series Models to Verify Structural Identifying Restrictions and to Assess Public Debt Sustainability

Anton Velinov

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, June 2013

European University Institute
Department of Economics

On Using Markov Switching Time Series Models to Verify
Structural Identifying Restrictions and to Assess Public Debt
Sustainability

Anton Velinov

Thesis submitted for assessment with a view to obtaining the degree of Doctor of
Economics of the European University Institute

Examining Board

Prof. Helmut Lütkepohl, DIW Berlin and Freie Universität (External Supervisor)

Prof. Peter Hansen, European University Institute

Prof. Ralf Brüggemann, University of Konstanz

Prof. Luca Fanelli, University of Bologna

© Anton Velinov, 2013

No part of this thesis may be copied, reproduced or transmitted without prior
permission of the author

On Using Markov Switching Time Series Models to Verify Structural Identifying Restrictions and to Assess Public Debt Sustainability

Abstract

The first paper in this thesis deals with the issue of whether there are bubble components in stock prices. This is joint research with Wenjuan Chen (Free University Berlin). We investigate existing bivariate structural vector autoregressive (SVAR) models and test their identifying restriction by means of a Markov switching (MS) in heteroskedasticity model. We use data from six different countries and find that, for five of the country models, the structural restriction is supported at the 5% level. Accordingly, we label the two structural shocks as fundamental and non-fundamental. This paper illustrates the virtue of being able to test structural restrictions in order to justify the relevant shocks of interest.

The second paper proceeds in the spirit of the first paper. In particular, five trivariate structural VAR or vector error correction (VEC) versions of the dividend discount model are considered, which are widely used in the literature. A common structural parameter identification scheme is used for all these models, which claims to be able to capture fundamental and non-fundamental shocks to stock prices. A MS-SVAR/SVEC model in heteroskedasticity is used to test this identification scheme. It is found that for two of the five models considered, the structural identification scheme appropriately classifies shocks as being either fundamental or non-fundamental. These are models which use real GDP and real dividends as proxies of real economic activity. The findings are supported by a series of robustness tests. Results of this paper serve as a good guideline when conducting future research in this field.

The third thesis paper addresses the question of how sustainable a government's current debt path is by means of a Markov switching Augmented Dickey-Fuller (MS-ADF) model. This model is applied to the debt/GDP series of 16 different countries. Stationarity of this series implies that public debt is on a sustainable path and hence, the government's present value borrowing constraint holds. The MS specification also allows for unit root and explosive states of the debt/GDP process. Two different criteria are used to test the null hypothesis of a unit root in each state. The countries with a sustainable debt path are found to be Finland, Norway, Sweden, Switzerland and the UK. The model indicates that France, Greece, Ireland and Japan have unsustainable debt trajectories. The remaining seven countries, (Argentina, Germany, Iceland, Italy, Portugal, Spain and the US) are all found to have uncertain debt paths. The model is robust to the sample size and number of states used. It is shown that this model is an improvement to existing models investigating this subject.

This work would not be able to resemble its present state without the help and co-operation of various people. Firstly, I would like to thank my supervisor, Helmut Lütkepohl for very useful advice and guidance. I also thank Wenjuan Chen for working with me on one of the thesis papers, this was a very helpful and educative experience. My thanks also go to Aleksei Netsunajev for very useful discussions on the technical aspects and programming issues involved in this work. Finally, many thanks to Norbert Metiu, Tomasz Wozniak, Reppa Zoltan and everyone who listened to me presenting the thesis papers at seminars and made comments on them. I dedicate this thesis to my family and friends.

Contents

1 Are there Bubbles in Stock Prices? Testing for Fundamental Shocks	5
1.1 Introduction	6
1.2 The Model	8
1.3 The MS-SVAR Model Results	12
1.4 Empirical Simulations of the SVAR Model	24
1.5 Conclusion	30
1.6 Appendix	31
2 Can Stock Price Fundamentals Really be Captured? Using Markov Switching in Heteroskedasticity Models to Test Identification Restrictions	36
2.1 Introduction	37
2.2 The Models	38
2.3 Estimation and Testing Procedure	43
2.4 Model Results	45
2.5 Robustness Analysis	59
2.6 Conclusion	65

3 Assessing the Sustainability of Government Debt - on the Different States of the Debt/GDP Process	67
3.1 Introduction	68
3.2 Related Literature	70
3.3 The model	72
3.4 The Data and Countries Investigated	74
3.5 Diagnostic Tests and Model Selection	76
3.6 A Note on Estimation and Testing	82
3.7 Empirical results	83
3.8 Robustness Analysis	98
3.9 Conclusion	104

Chapter 1

Are there Bubbles in Stock Prices? Testing for Fundamental Shocks

Anton Velinov (EUI), Wenjuan Chen (FU Berlin)

Abstract

This paper makes use of a Markov switching-SVAR (MS-SVAR) model in heteroskedasticity as in Lanne et al. (2010) and Herwartz and Lütkepohl (2011) to test a long-run restriction that would just-identify a bivariate SVAR model in the conventional sense. In particular, we investigate the SVAR model considered by Binswanger (2004a) and Groenewold (2004), which tries to answer the question of how well stock prices reflect their fundamentals. We use data from six different countries and find that, for five of the country models, the long-run restriction is supported at the 5% level. Accordingly, we label the two structural shocks as fundamental and non-fundamental. From conventional SVAR analysis, impulse responses show that in most cases there are significant long-run effects of a fundamental shock to real stock prices at the 95% confidence level. Historical decompositions indicate a potential bubble in real stock prices for Japan and pricing according to fundamentals for the UK and the US. This paper illustrates the virtue of being able to test structural restrictions in order to justify the relevant shocks of interest.

1.1 Introduction

There is a wide range of literature investigating stock prices and their relation to other macroeconomic variables. A very popular tool in this area is the vector autoregressive (VAR) or the so-called reduced form model. Once estimated this model can be used to draw inferences about the economic relationships through empirical simulations such as impulse response analysis, forecast error variance decompositions and historical decompositions.

Studies applying the VAR model to share prices and macroeconomic variables include some early work by Campbell and Shiller (1988), who try to forecast stock returns and find that using historical averages of real earnings is one of the most important predictors. Later applications include work by Gjerde and Sættem (1999) who analyze the relations among stock returns and many other macroeconomic variables for the Norwegian economy. They find that the real interest rate and real activity are important variables for explaining returns. Cheung and Ng (1998) conduct a similar analysis for several countries making use of a vector error correction (VEC) model to control for cointegrating relationships.

However, because the error terms are often correlated in such models and due to their non-theoretical nature, results from reduced form models could be subject to ambiguous interpretation. This is because there is no clear distinction of the shocks. Hence, it is not always possible to classify shocks in a desired way.

Many studies therefore make use of the structural VAR (SVAR) model, so that shocks can be defined according to economic theory. This would make model simulation results potentially easier to interpret. Contributions in this area - applying the SVAR model to share prices and other macro data - include work by Lee (1995b), who finds that stock prices respond equally strong to both permanent and temporary shocks to dividends. In a follow up analysis, Lee (1998) introduces a non-fundamental component to the model and finds that stock prices tend to deviate from fundamentals only in the short-run and then gradually reach their price according to fundamentals. He therefore draws the conclusion that there are fads rather than bubbles present in stock prices.¹ Lastrapes (1998) analyzes the effect of money supply shocks on stock and bond prices and concludes that there is a real liquidity effect for both.

¹ A fad is defined by a gradual change in stock prices rather than a sudden shift as in a bubble.

Rapach (2001) uses a four-variable SVAR model to characterize the effects of different macro shocks on stock prices. Relying on macro theory he imposes restrictions on the model to identify the structural shocks. He finds that the surge of share prices in the late 1990s was driven by the expansionary effects of these macro shocks. Slightly more recent publications are those by Binswanger (2004a) and Groenewold (2004), both of whom use a bivariate SVAR model with industrial production and real share prices. They try to determine whether stocks are priced above their fundamentals, i.e. whether there are bubble components in stock prices. They conclude that from the mid 1990s this has indeed been the case.

Although the SVAR model does solve some of the interpretation problems presented by the reduced form model, it has a drawback in that it requires identification restrictions. This is caused by the fact that only the parameters of the VAR model can be estimated consistently and efficiently by OLS. The reduced form parameter estimates can therefore be used to obtain the structural parameter estimates. There are however, more structural parameters than reduced form parameters hence $K(K-1)/2$ identification restrictions need to be imposed, where K is the number of endogenous variables. The restrictions are imposed on the structural parameters in various ways, below we discuss this in more detail. For example, in a four-variable model there would need to be 6 restrictions in order to identify the structural shocks. Since these restrictions solely rely on economic theory and can hence not be tested, they may not be altogether innocuous.

A way to go around this problem is proposed in Lanne et al. (2010) and slightly more recently in Herwartz and Lütkepohl (2011); the latter of which deals with a structural vector error correction model. The basic idea is to let the data decide whether the restrictions needed for identifying the structural model are supported or not. This is done through a Markov switching-SVAR (MS-SVAR) model in heteroskedasticity. Even with just two Markov states, this model is potentially exactly identified, and hence any further restrictions would be over-identifying and could be tested by means of conventional statistical tests.

The main contribution of this paper is to make use of this testing technique and to apply it to an already existing SVAR model, namely the bivariate model of Binswanger (2004a) and Groenewold (2004). Their model investigates the role of fundamentals in

stock prices. It is therefore very important to make sure that these fundamentals are truly captured, especially when using a simple bivariate model. Hence, we use the MS-SVAR model to allow the data to tell us whether the structural shocks are correctly identified. We find that in the majority of cases the structural identification scheme is accepted, thereby providing a formal justification to the relevant shocks of interest. We then use the properly identified SVAR model to derive impulse responses, forecast error variance decompositions and historical decompositions.

The paper is structured as follows: Section 2 introduces the basic SVAR model and the MS-SVAR model. Section 3 discusses the MS-SVAR model results and Section 4 displays the empirical results of the SVAR model, having concluded with the MS part. Finally, Section 5 concludes.

1.2 The Model

The focus of this paper is on a bivariate SVAR model used in Binswanger (2004a) and Groenewold (2004). The two variables are the log of industrial production (IP_t) and the log of real stock prices, (s_t). Since industrial production and real GDP are closely related, both measures are commonly used as proxies of real economic activity. Studies, such as James et al. (1985) and Chen et al. (1986), find that industrial production is a significant factor in explaining share prices. The rest of this section elaborates on the theoretical side of the model.

The following reduced form VAR(p) model in first differences is considered:

$$\Delta y_t = \nu + A_1 \Delta y_{t-1} + A_2 \Delta y_{t-2} + \cdots + A_p \Delta y_{t-p} + u_t. \quad (1.1)$$

Δy_t is a $(K \times 1)$ vector of the endogenous variables, in our case, $\Delta y_t = [\Delta IP_t, \Delta s_t]'$, so that $K = 2$. Δ is the first difference operator (such that $\Delta y_t = y_t - y_{t-1} = (1 - L)y_t$, where L is the lag operator). ν is a (2×1) vector of constants, $A_i, i = 1, \dots, p$ are (2×2) parameter matrices and u_t is a (2×1) vector of unobservable error terms with $E[u_t] = 0$ and $E[u_t u_t'] = \Sigma_u$, not necessarily diagonal (where E denotes the expectation operator). The above equation can be rewritten as

$$A(L) \Delta y_t = \nu + u_t, \quad (1.2)$$

where $A(L) = I_K - A_1L - A_2L^2 - \dots - A_pL^p$. Provided that $A(L)^{-1}$ exists, the Wold MA representation for the stationary Δy_t process is

$$\Delta y_t = \mu + \sum_{s=0}^{\infty} \Phi_s u_{t-s} = \mu + \Phi(L) u_t, \quad (1.3)$$

where $\mu = (I_K - A_1 - A_2 - \dots - A_p)^{-1}v = A(1)^{-1}v$, $\Phi(L) \equiv A(L)^{-1}$ and $\Phi_0 = I_K$. Using the B-model, structural shocks are identified as $u_t = B\varepsilon_t$, where B is the contemporaneous impact matrix. Further, it is usually assumed that $E[\varepsilon_t \varepsilon_t'] = \Sigma_\varepsilon$ which is a diagonal covariance matrix (usually the identity matrix). Hence, the structural representation of the model is

$$\Delta y_t = \mu + \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s} = \mu + \Psi(L) \varepsilon_t, \quad (1.4)$$

here $\Psi_i \equiv \Phi_i B$, for $i = 0, 1, 2, \dots$. The accumulated long-run effects of the structural shocks over all periods are given by the long-run impact matrix, $\Psi \equiv \Phi B$, where $\Phi \equiv \sum_{s=0}^{\infty} \Phi_s = A(1)^{-1}$.

1.2.1 Identifying restrictions

As already discussed, OLS estimation only yields consistent and efficient estimates of the reduced form parameters in (1.1). From the assumptions made above it follows that $\Sigma_u = BB'$. Since the covariance matrix is symmetric, it has $K(K+1)/2$ non-repeating diagonal and off diagonal elements. The B matrix on the other hand consists of K^2 elements. Hence, we need to impose $K^2 - K(K+1)/2 = K(K-1)/2$ restrictions to identify the structural parameters of the model. In our case this amounts to one restriction since $K = 2$.

Restrictions can be imposed directly on the B matrix or indirectly through the long-run impact matrix, Ψ as is proposed by Blanchard and Quah (1989). Long-run restrictions are used in Binswanger (2004a) and Groenewold (2004), who both set the upper right element, $\Psi_{1,2}$, of the long-run impact matrix to zero making it lower triangular. Hence,

$$\Psi = \begin{bmatrix} \star & 0 \\ \star & \star \end{bmatrix} \quad (1.5)$$

where \star can take on any value. Consequently the structural shocks, $\varepsilon_t = [\varepsilon_t^F, \varepsilon_t^{NF}]'$ can be interpreted as fundamental and non-fundamental shocks respectively. Hence,

it is assumed that a fundamental shock can have a permanent effect on the economy and on the stock market, whilst a non-fundamental shock, although having a permanent effect on real stock prices, can only have a transitory effect on the economy.

This way the model is just-identified, however, it is important to be able to test this restriction. Since we only have two variables it is hard to be certain whether we have indeed isolated the fundamental and the non-fundamental components of stock prices. Any subsequent empirical results and conclusions are highly dependent on this assumption. That is why in this case it would be good to test whether the above identification restriction is supported by the data.

1.2.2 The MS-SVAR model

Testing such identifying restrictions is proposed in Lanne et al. (2010) and Herwartz and Lütkepohl (2011) by means of a MS-SVAR in heteroskedasticity model. We briefly present the model and the basic testing strategy here.

The MS-VAR in heteroskedasticity model is exactly the same as the conventional VAR model given by equation (1.1) with the exception that the residuals are assumed to be normally and independently distributed conditional on a Markov state. In other words,

$$u_t|S_t \sim \text{NID}(0, \Sigma(S_t)), \quad (1.6)$$

where S_t follows a first order discrete valued Markov process. Normality is assumed so that it is possible to use maximum likelihood estimation to estimate the parameters. As is demonstrated in Lanne et al. (2010), this assumption is not restrictive and a wide class of unconditional distributions, other than just the normal, are captured. The discrete stochastic process S_t is assumed to take on M regimes with transition probabilities given by

$$p_{ij} = P(S_t = j | S_{t-1} = i), \quad i, j = 1, \dots, M,$$

which can be arranged in an $(M \times M)$ matrix of transition probabilities,

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix}. \quad (1.7)$$

Note that the probabilities add up to one row-wise, hence $p_{iM} = 1 - p_{i1} - p_{i2} - \dots - p_{iM-1}$, $i = 1, \dots, M$.

In order to test the identifying restriction it is necessary to decompose the covariance matrices in the following way:

$$\Sigma(1) = BB', \quad \Sigma(2) = B\Lambda_2B', \quad \dots \quad \Sigma(M) = B\Lambda_MB', \quad (1.8)$$

where B is the contemporaneous impact matrix and $\Lambda_i = \text{diag}(\lambda_{i1}, \lambda_{i2})$, $i = 2, \dots, M$ can be interpreted as relative variances. The underlying assumptions are that the contemporaneous effects matrix, B , stays constant across regimes and that shocks are orthogonal across regimes, i.e. Λ_i , $i = 2, \dots, M$ is diagonal. The assumption of a regime-invariant B matrix means that impulse responses are constant throughout different time periods so that empirical simulation results are as in a common SVAR model. Orthogonality of the shocks implies that the covariance matrices are different across regimes, which is necessary to identify the K^2 parameters of the B matrix and the $(M-1)K$ parameters of the diagonal Λ_i , $i = 2, \dots, M$ matrices.² As shown in Proposition 1 of Lanne et al. (2010), provided that the pairwise diagonal elements of one of the Λ_i , $i = 2, \dots, M$ matrices are distinct, the B matrix is identified up to sign changes and column ordering.

Although a state invariant B matrix may seem somewhat of a restrictive assumption, it can be tested, provided that three or more Markov states are used. This is done by a standard likelihood ratio (LR) test, which has an asymptotically χ^2 distributed test statistic with $(1/2)MK(K+1) - K^2 - (M-1)K$ degrees of freedom.

Given the conditional normality assumption of the residuals, it is possible to use maximum likelihood estimation and in particular, the Expectation Maximization (EM) algorithm. This algorithm estimates the MS-SVAR parameters along with the unrestricted transition probabilities in (1.7) and shows the smoothed probabilities depicting which state prevails at what time period. The EM algorithm was initially popularized in Hamilton (1994) for the univariate case and was later extended to multivariate models in Krolzig (1997). The parameters of the decomposition in (1.8) are estimated

²For example in case of two states we have two reduced form covariance matrices with $2 \times (K(K+1)/2) = K^2 + K$ unique diagonal and off diagonal elements. This is enough to identify the K^2 elements of B and the K diagonal elements of Λ_2 .

using a non-linear optimization procedure in the Maximization step of the EM algorithm. Details of the EM algorithm used in this paper are given in the Appendix.

Once the EM algorithm has converged, standard errors of the point estimates of all unrestricted parameters are obtained through the inverse of the negative of the Hessian matrix evaluated at the optimum. The standard errors enable the use of Wald tests (and in addition LR tests are used) to determine whether the diagonal elements of at least one of the $\Lambda_i, i = 2, \dots, M$ matrices are distinct. The test distributions are asymptotically χ^2 , with the degrees of freedom depending on the number of joint hypotheses being tested, which depends on the number of Markov states.

Finally, provided that the B matrix is identified up to changes in sign, any additional restrictions on it are over-identifying. Hence, we are in a position to test the identifying restriction in (1.5). This is done by estimating the restricted model (i.e. the one with the restriction in (1.5)) and obtaining its maximum log-likelihood. A standard LR test is used to determine whether the identifying restriction is accepted. The test distribution is χ^2 with 1 degree of freedom since one restriction is being tested. The next section describes the data and the results of the MS-SVAR model.

1.3 The MS-SVAR Model Results

1.3.1 The Data

Following Binswanger (2004a), we obtain most of the data from the IMF International Financial Statistics (IFS) database; but Global Financial Data (GFD) and Datastream are also used. The series consist of a seasonally adjusted industrial production index, a stock price index and a consumer price index (CPI), all of which are normalized to a base year of 2005. The stock price series is converted to real terms by dividing by the percentage CPI, hence the CPI series is not used directly in the analysis. In addition, all series are in logs. The data range is quarterly from 1960:I-2012:III; the frequency and the starting date being inline with Binswanger (2004a) and Groenewold (2004).

We use data on France, Germany, Italy, Japan, the UK and the US. Figure 1.1 plots both variables for each country. The UK and the US are also covered in Binswanger (2004a). Together these countries have the highest nominal GDP in the world and

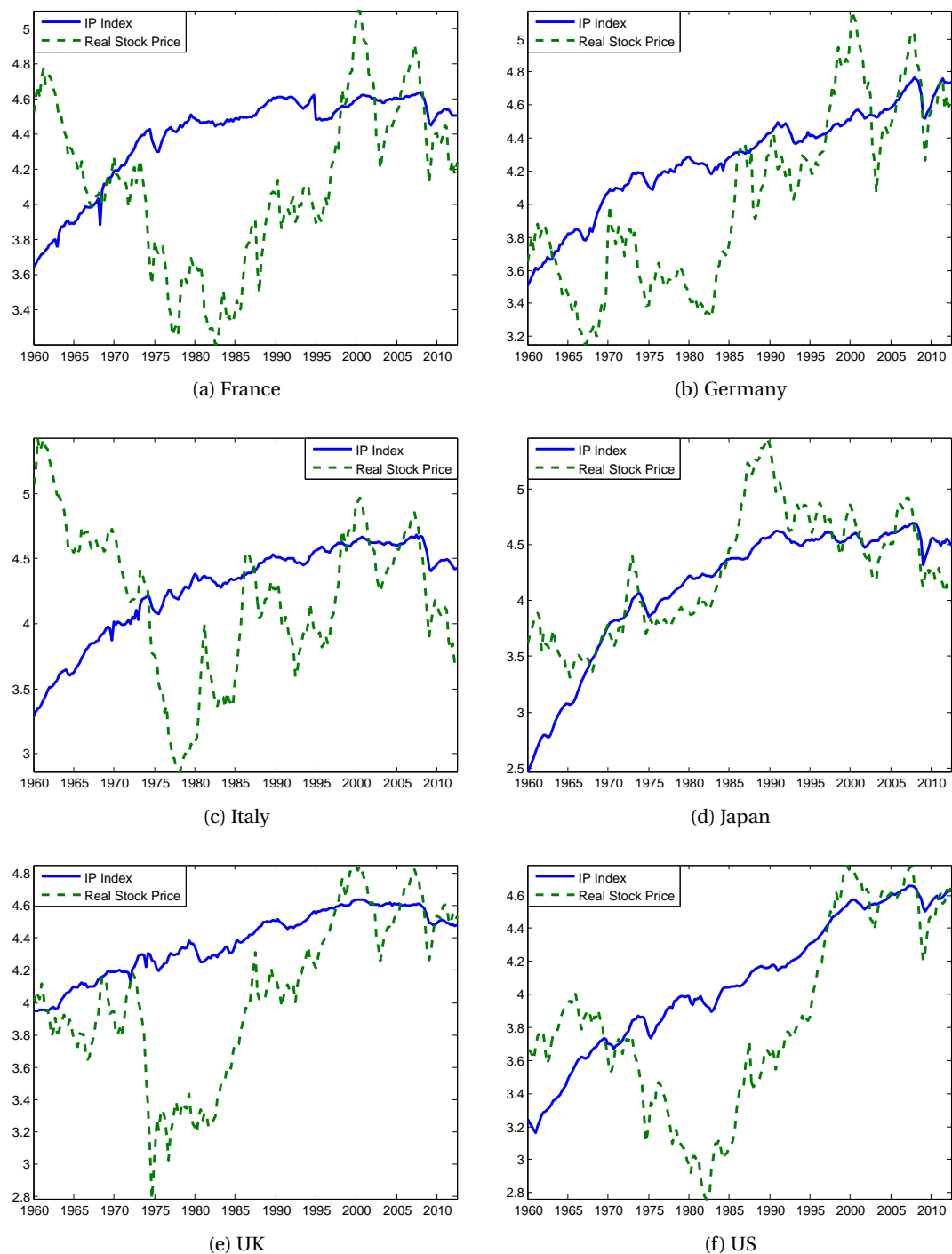


Figure 1.1: Industrial production and real stock price series in log levels per country.

consequently among the highest GDP per capita³.

Augmented-Dickey-Fuller (ADF) unit root tests show that the levels series for all countries are integrated of order one, meaning that the first differences are stationary. When testing for cointegrating relationships on the levels VAR model using a constant and trend term, both the Saikkonen and Lütkepohl (2000) and the Johansen (1995) trace test cannot reject the null hypothesis of no cointegrating relations at the 1% level. Hence, it would appear to be justified to use the standard VAR model in first differences, as given in equation (1.1). Multivariate conditional heteroskedasticity tests show that in most cases the null hypothesis of no heteroskedasticity can be rejected at conventional statistical levels.⁴ All of the above tests are carried out using the Lütkepohl and Krätzig (2004) JMulTi software.

1.3.2 MS-SVAR Model Specifications and Results

Before we begin with testing the identification restriction, we need to determine the type of MS-SVAR model to be used per country. In other words, we need to select the appropriate number states and lags that the model should have. In order to select the states we make use of the information criteria developed by Psaradakis and Spagnolo (2006), which are also used by Herwartz and Lütkepohl (2011) and are found to deliver reasonable results. In particular, we focus on the Akaike Information Criterion (AIC) and the Schwartz Criterion (SC). The AIC is calculated as $-2(\log\text{-likelihood} - n)$ and the SC is calculated as $-2\log\text{-likelihood} + \log(T)n$, where T is the sample size and n stands for the number of free parameters of the model. While these criteria can jointly help select the number of states and lags, we prefer to rely on residual Portmanteau tests to help choose the lag length so as to be certain that there is no presence of residual autocorrelation.

The log-likelihood and the information criteria values for unrestricted models - without the long-run restriction in (1.5) - are reported in Table 1.1. The minimum values of the criteria are shown in bold. Both criteria agree on a model with two Markov states for France and the US. For all the other country models, the AIC tends to opt for

³According to the IMF, the World Bank and the CIA World Factbook

⁴More precisely, a multivariate ARCH-LM test with four lag orders is used. Lag lengths for the individual country models are determined according to the Akaike Information Criterion (AIC), which opts for the same lag orders as the ones chosen for the MS models discussed in Section 3.2.

Table 1.1: Log-likelihood, AIC and SC per country for two and three-state models.

Country	States	Log-likelihood	AIC	SC
France	1	690.125	-1360.250	-1326.827
	2	802.746	-1577.492	-1530.699
	3	807.048	-1574.097	-1507.250
Germany	1	753.854	-1487.709	-1454.286
	2	782.865	-1537.731	-1490.938
	3	793.581	-1547.163	-1480.316
Italy	1	674.503	-1313.006	-1253.018
	2	720.129	-1396.258	-1322.938
	3	734.264	-1412.528	-1319.212
Japan	1	726.489	-1432.979	-1399.556
	2	790.009	-1552.019	-1505.226
	3	799.900	-1559.800	-1492.953
UK	1	805.791	-1567.582	-1494.369
	2	865.552	-1679.104	-1592.580
	3	875.643	-1687.287	-1580.795
US	1	942.415	-1864.831	-1831.408
	2	970.008	-1912.016	-1865.223
	3	974.119	-1908.239	-1841.297

more states than the SC. Without skipping too far ahead, it is sufficient to say at this point that models with three volatility states meet the necessary requirement when it comes to distinction of the relative variance elements, $\lambda_{ij}, i = 2, \dots, M, j = 1, 2$. Hence, in case of disagreement among the criteria, a model with three states is chosen, according to the AIC. This means that Germany, Italy, Japan and the UK are all modeled with three volatility states. Finally, no criteria favors a conventional SVAR model, not subject to regime switches, i.e. in all cases, the AIC and SC values are fairly high for a 1-state model.

It is probably notable that a maximum of three Markov states are considered in the table. In fairness, for Germany and Japan the AIC is marginally in favor for a model with four states. However, when using more Markov states the probability of ending up with very few observations in a given state increases. This can lead to highly imprecise estimates and rather meaningless smoothed probabilities. Further,

due to the specific non-linear nature of the models considered, when more states are used it may be a time consuming and cumbersome exercise to find a global optimum. Once such an optimum is found, there is a fair chance that it may give meaningless results. In light of these considerations a maximum number of three states are taken into account.

As discussed, the model lag lengths are chosen such that there is no residual auto-correlation present according to standard Portmanteau tests. In the interest of parsimony, as few lags are chosen as possible whilst ensuring that the mentioned requirement is met. Therefore, we use one lag order for France, Germany, Japan and the US, three for Italy and four for the UK.

In summary, a MS(2)-SVAR(1) model is used for France and the US, a MS(3)-SVAR(1) model for Germany and Japan, a MS(3)-SVAR(3) model for and Italy and a MS(3)-SVAR(4) model for the UK. Here MS(M)-SVAR(p) stands for a SVAR model with p lags and M Markov switching (MS) volatility states.

Once the models are estimated, the main parameters of interest for our analysis are the relative variance parameters, $\lambda_{ij}, i = 2, \dots, M, j = 1, 2$. These require to be distinct in at least one $\Lambda_i, i = 2, \dots, M$ matrix so that the contemporaneous impact matrix, B in (1.8) is identified up to changes in sign. Other relevant parameters are the transition probabilities, $p_{ij}, i, j = 1, \dots, M$ these indicate how persistent a given state is and can be used to calculate the durations of states. It is easiest to present the estimation results of these parameters along with other model results for two and three-state models separately.

1.3.3 Results of 2-state models

Models with two MS states are used for the data from France and the US. The relevant parameter estimates and their standard errors along with the covariance matrices (scaled by 10^{-3}) are presented in Table 1.2. This table only displays the parameters of the unrestricted models, i.e. without the long-run restriction on the B matrix. Recall that if $\lambda_{21} \neq \lambda_{22}$, then the contemporaneous effects matrix is identified up to sign and ordering. Hence, if that is the case, any additional restrictions can be tested. Upon first glance these parameters seem distinct, however their standard errors are quite large. Distinction is tested later in Table 1.4 by means of standard Wald and likelihood

Table 1.2: Parameter estimates, standard errors and covariance matrices (scaled by 10^{-3}) for 2-state unrestricted models.

	France		US	
Parameter	estimate	σ	estimate	σ
λ_{21}	70.022	31.776	7.731	2.092
λ_{22}	2.538	1.206	3.352	0.914
p_{11}	0.969	0.014	0.784	0.056
p_{22}	0.481	0.171	0.662	0.138
$\Sigma(1)$	$\begin{bmatrix} 0.141 & - \\ 0.013 & 6.580 \end{bmatrix}$		$\begin{bmatrix} 0.046 & - \\ -0.003 & 2.008 \end{bmatrix}$	
$\Sigma(2)$	$\begin{bmatrix} 9.831 & - \\ -0.440 & 16.722 \end{bmatrix}$		$\begin{bmatrix} 0.344 & - \\ 0.305 & 7.267 \end{bmatrix}$	

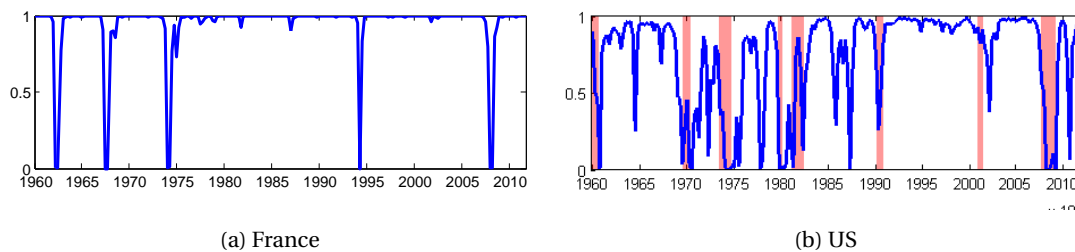


Figure 1.2: Smoothed probabilities of state 1 for the 2-state unrestricted models. US recession dates according to NBER dating given by the shaded bars.

ratio (LR) tests.

To help classify the different volatility states, information on the covariance matrices is needed as well as the on the smoothed probabilities. The smoothed probabilities display the probability of being in a given state at a particular time period. The ones for state 1 for the unrestricted models are shown in Figure 1.2. Note that in the case of two Markov states, the probability of being in state 2 is the mirror image of that of state 1.

For both models, the variances (the diagonal elements of the covariance matrices) increase with the state. In that sense state 2 can be thought of as depicting more

volatile periods than state 1. This can be seen as well when looking at the smoothed probabilities, especially the ones for the US in Panel (b) of Figure 1.2. To help clarify the time periods further, US recession dates are given by the shaded bars according to NBER dating. There is a clear tendency for the model to indicate a switch to the more volatile state in times of recessions. This is in line with literature studying stock market volatility and stock returns. For instance, Schwert (1989) and Hamilton and Lin (1996) find that volatility increases are a result of economic downturns such as recessions or crises.

The duration of the states can be inferred from the transition probabilities - the closer the value is to one, the more persistent the state is likely to be. The exact formula for the duration of any of the M states is $1/(1 - p_{ii}), i = 1, \dots, M$. Clearly, state 1 is the more long-lasting state as we also would expect since non-recession periods tend to last longer than recessionary ones.

Having concluded with the 2-state models for now, we turn to the models with three Markov states.

1.3.4 Results of 3-state Models

The SVAR models, for Germany, Italy, Japan and the UK make use of three MS in volatility states. Estimates of the same types of parameters as for the 2-state models are shown in Table 1.3. With three volatility states there are two Λ_i matrices. Their diagonal elements are not all above one as in the two state case. This means that some of the relative variances in the other states are lower than those of state 1. This can also directly be seen from the diagonal entries in the covariance matrices - they are not always increasing with a given state. In most cases however, state 1 tends to have the lowest volatility, whilst state 3 can be characterized as being most volatile.

The smoothed probabilities for the same 3-state models are displayed in Figure 1.3. Clearly, the third volatility state always captures the period of the late 2000s crisis, indicating that this was indeed a globally very volatile event. State 1, usually most prevalent with the exception of Italy, tends to capture more stable time periods, which are characterized by lower volatility.

As previously noted, the assumption of a state invariant B matrix can now be tested since three Markov states are used. We therefore formally test this assumption

Table 1.3: Parameter estimates, standard errors and covariance matrices (scaled by 10^{-3}) for 3-state unrestricted models. Tests for a state-invariant B matrix at the bottom.

Model	Germany		Italy		Japan		UK	
Parameter	estimate	σ	estimate	σ	estimate	σ	estimate	σ
λ_{21}	0.180	0.128	1.412	0.404	4.844	1.523	4.155	2.015
λ_{22}	7.028	2.928	5.427	1.732	0.931	0.392	0.773	0.198
λ_{31}	9.196	4.194	21.973	9.825	59.258	41.264	34.703	22.246
λ_{32}	1.026	0.578	6.637	2.995	6.177	3.721	9.129	5.595
p_{11}	0.937	0.033	0.905	0.048	0.982	0.012	0.929	0.036
p_{22}	0.593	0.161	0.991	0.078	0.918	0.174	0.862	0.078
p_{33}	0.742	0.158	0.745	0.124	0.605	0.310	0.544	0.168
$\Sigma(1)$	$\begin{bmatrix} 0.224 & - \\ 0.272 & 5.342 \end{bmatrix}$		$\begin{bmatrix} 0.191 & - \\ 0.107 & 2.028 \end{bmatrix}$		$\begin{bmatrix} 0.184 & - \\ 0.125 & 5.223 \end{bmatrix}$		$\begin{bmatrix} 0.056 & - \\ 0.019 & 3.698 \end{bmatrix}$	
$\Sigma(2)$	$\begin{bmatrix} 0.020 & - \\ 0.863 & 58.187 \end{bmatrix}$		$\begin{bmatrix} 0.270 & - \\ 0.098 & 10.705 \end{bmatrix}$		$\begin{bmatrix} 0.883 & - \\ 0.960 & 5.861 \end{bmatrix}$		$\begin{bmatrix} 0.235 & - \\ 0.048 & 2.863 \end{bmatrix}$	
$\Sigma(3)$	$\begin{bmatrix} 2.020 & - \\ 1.837 & 5.917 \end{bmatrix}$		$\begin{bmatrix} 4.188 & - \\ 2.546 & 14.617 \end{bmatrix}$		$\begin{bmatrix} 10.789 & - \\ 12.212 & 45.817 \end{bmatrix}$		$\begin{bmatrix} 1.958 & - \\ 0.422 & 33.800 \end{bmatrix}$	
H_0 : state invariant B								
p -value	0.152		0.620		0.680		0.676	

and show the relevant p -values per country at the bottom of Table 1.3. Recall that the test statistic follows an asymptotic χ^2 distribution with $(1/2)MK(K+1) - K^2 - (M-1)K$ degrees of freedom, which is 1 in our case. We see that in all cases the p -values are above the 10% critical level meaning that the assumption of a state invariant contemporaneous effects matrix has support from the data.

1.3.5 Testing the Model Restrictions

We now continue the analysis with both the two and three state models and test whether the identifying restriction in (1.5) is supported. First, however, we need to make sure that the state-invariant and unrestricted B matrix in (1.8) is identified up to changes in sign and ordering. In other words we need to test whether the B matrix is identified through heteroskedasticity. If that is the case then any additional restric-

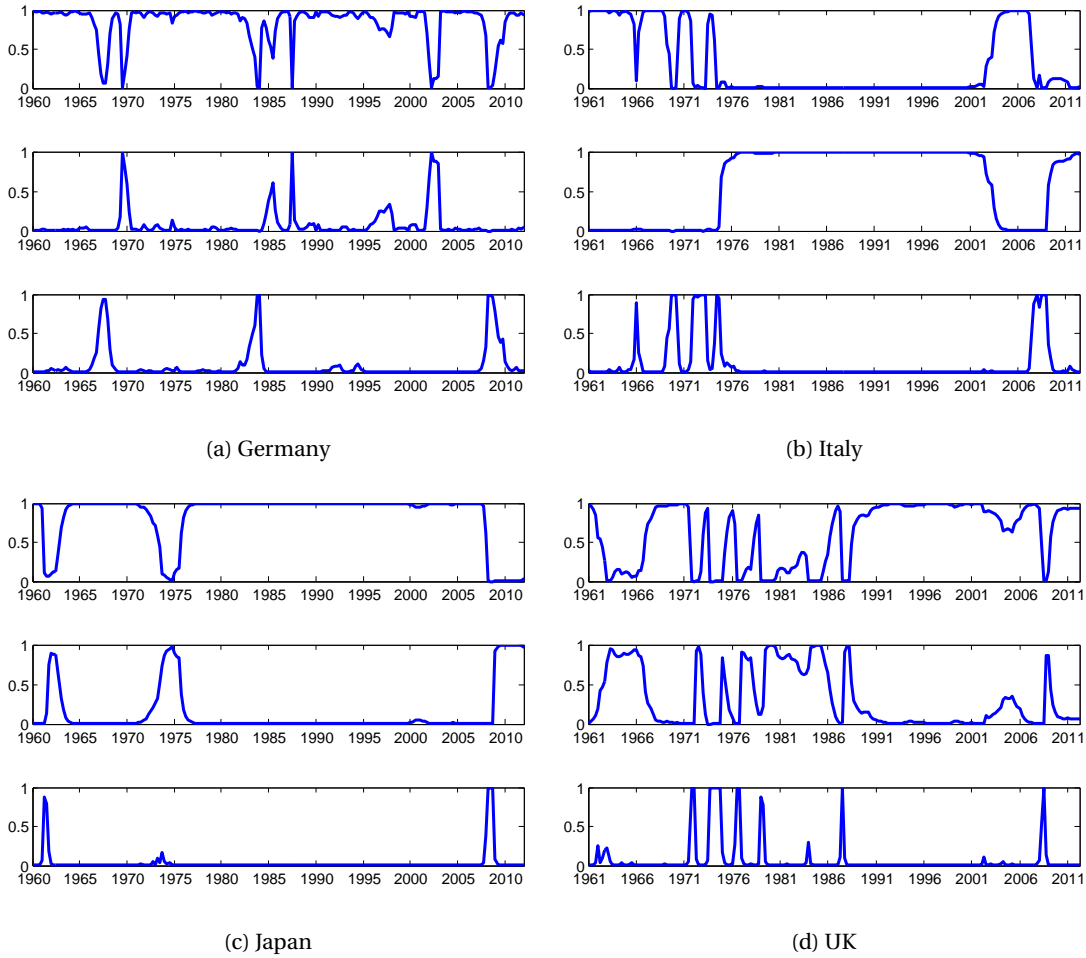


Figure 1.3: Smoothed probabilities of State 1 (top), State 2 (middle) and State 3 (bottom) for the 3-state unrestricted models.

tions short or long-run become over-identifying and are in a position to be tested.

To make sure that B is identified through heteroskedasticity, it is necessary for the pair of diagonal elements in at least one of the $\Lambda_i, i = 2, \dots, M$ matrices to be distinct. This can most readily be tested by means of a Wald test as the standard errors of the parameters are available. We also use LR tests however, and usually both tests come to the same conclusion. The test distribution is asymptotically χ^2 with $M - 1$ degrees of freedom, since there are $M - 1$ Λ_i matrices not equal to the identity matrix and a joint test over the equality of their pairwise elements is made.

The relevant null hypotheses and their p -values of both Wald and LR tests are shown in Table 1.4. Since the models for France and the US make use of two volatility states, their null hypothesis only concerns the elements in Λ_2 . The 3-state models also have a Λ_3 matrix, the elements of which need to be taken into account as well. Hence, a joint hypothesis test is carried out.

Table 1.4: Null hypotheses and p -values for two and three state models.

$H_0 : \lambda_{21} = \lambda_{22}$	Wald test	LR test
France	0.034	0.000
US	0.051	0.028
$H_0 : \lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}$	Wald test	LR test
Germany	0.012	0.000
Italy	0.003	0.000
Japan	0.021	0.001
UK	0.696	0.004

It can be seen that in most cases the Wald tests reject the null hypotheses at the 5% level, while the LR tests usually reject them at the 1% level. Hence, we can be fairly confident that a pair of diagonal elements in at least one of the $\Lambda_i, i = 1, \dots, M$ matrices are distinct. This means that the B matrix in all of the models considered thus far is identified up to changes in sign and ordering.

Given that the long-run restriction in (1.5) is now over-identifying, it is in a position to be tested. This is done by means of an LR test, where the maximum log-likelihood of the unrestricted model is compared to the maximum log-likelihood of the model with the long-run restriction. The test distribution is asymptotically χ^2 with 1 degree of freedom since one restriction is being tested.

The p -values of the LR tests are summarized in Table 1.5. Note that the null hypothesis is the long-run restriction given in (1.5), while the alternative is an unrestricted state invariant B matrix. At the 5% critical level the restriction is accepted for all models except for the one for the UK. The restriction for Japan and the US is accepted at the 10% level. Overall, there seems to be credible evidence in favor of the

long-run restriction in (1.5). Hence, we can characterize the shocks as being fundamental and non-fundamental as was our initial objective.

Table 1.5: p -values of LR tests of the long-run restriction in (1.5).

H_0 : (1.5) H_1 : unrestricted state invariant B					
France	Germany	Italy	Japan	UK	US
0.051	0.068	0.089	0.132	0.001	0.294

1.3.6 A Small Robustness Check

To see whether the results are to some extent driven by the data range or the specific model used, a robustness check is in order. Firstly, although it is not a unique event according to the smoothed probabilities, the period of the financial crisis marks a rather turbulent time. There were large falls in stock prices worldwide as well as a drop in the industrial production (IP) index. This event is also depicted by the smoothed probabilities for all country models, making it a truly global occurrence. To see whether the financial crisis does indeed influence the results, the sample period is shortened to exclude it, and the same analysis is performed with the shortened sample. Since the starting date of this crisis is not very clear for all countries considered, we decide to be on the safe side and cut the sample from 2006:II onwards, hence the last observation is 2006:I. This way we are sure to avoid the event entirely in the smoothed probabilities of all countries.

Secondly, the specific type of MS model chosen could potentially influence the results. For instance, with data on stock markets (and to some extent IP), the intercept could also be subject to Markov switches. It is well known that in times of high (low) volatility stock prices tend to go down (up). In other words, there is no reason to assume that the intercept of the SVAR models is also not state-dependent. The autoregressive parameters could potentially also be switching, however the case for them to switch is harder to justify and to interpret. Further, switching autoregressive parameters may cause estimation issues; in that the number of parameters to be estimated increases and the data range may be too limited to give accurate estimates

of all these parameters when using many MS states. Hence, we decide to investigate a model only with a further switching intercept term in addition to the switching covariance matrix. The model then looks as follows

$$\Delta y_t = v(S_t) + A_1 \Delta y_{t-1} + A_2 \Delta y_{t-2} + \dots + A_p \Delta y_{t-p} + u_t, \quad (1.9)$$

where u_t has the same distributional assumption as in (1.6).

For the ease of comparison, the models used for robustness analysis keep the same number of states and lags as the original model specification. Residual autocorrelation is avoided in all cases. The analysis for both robustness specifications proceeds in the same way as before; first, the parameters of the unrestricted MS models are estimated, then for the 3-state models the hypothesis of a state-invariant B matrix is tested as in Table 1.3. The distinction of the relative variance parameters is then tested as in Table 1.4. Finally, provided that those tests yield favorable results the long-run restriction in (1.5) can be tested. It turns out that the assumption of a state-invariant B matrix is once again accepted for all 3-state models; and the B matrix is also identified through heteroskedasticity in that the relative variance parameters satisfy the distinction requirement.

Table 1.6 presents the p -values of the LR tests on the long-run restriction in (1.5) for both alternative specifications along with the original p -values at the bottom for comparison. We can see that the shorter range - excluding the financial crisis - usually leads to greater acceptance of the long-run restriction. Though the results do not considerably change, only the restriction for the UK would be accepted at the 5% level, while that of Japan would be rejected at that same level. With a switching intercept term we do see more changes in the results, though the original conclusion still holds for Japan, the UK and the US.

We therefore conclude that the restriction in (1.5) is overall supported by the data and that the original conclusions are not merely subject to the data range and exact model specification used. Relying on these conclusions, we next conduct a standard SVAR analysis with the identified shocks.

Table 1.6: p -values of LR tests of the long-run restriction in (1.5) for different robustness specifications.

	H_0 : (1.5) H_1 : unrestricted state invariant B					
	France	Germany	Italy	Japan	UK	US
Short rage	0.175	0.727	0.977	0.036	0.069	0.210
Intercept	0.027	2.548×10^{-5}	0.001	0.096	0.370	0.699
Original	0.051	0.068	0.089	0.132	0.001	0.294

1.4 Empirical Simulations of the SVAR Model

Having found support for the long-run restriction in (1.5), we label the SVAR shocks as fundamental and non-fundamental. Using conventional SVAR analysis, we investigate to what extent stock prices reflect their underlying fundamentals. This section proceeds as in Binswanger (2004a) and Groenewold (2004). We investigate the impact of fundamental shocks on stock prices by means of impulse responses (IRs), forecast error variance decompositions (FEVDs) and historical decompositions (HDs). The original lag lengths as in the MS-SVAR models are also kept for the conventional SVAR models. This is because standard information criteria, such as the AIC, for instance, suggest the same lag orders as the ones used for the MS case. These lag lengths are again sufficient to remove any residual autocorrelation.

We begin with an impulse response (IR) analysis of the systems. More precisely, we are interested in the response of stock prices to a fundamental shock. We also need to construct appropriate confidence intervals for the IRs. In particular, since we argue that a MS model in heteroskedasticity is appropriate for the type of data we are investigating and in addition we confirm the presence of conditional heteroskedasticity through tests; it is necessary to take the heteroskedastic properties of the residuals into account when constructing the relevant confidence intervals. Accordingly, we use the fixed design wild bootstrap technique as in Goncalves and Kilian (2004) to construct the IR confidence intervals. The series are bootstrapped as

$$\Delta y_t^* = \hat{v} + \hat{A}_1 \Delta y_{t-1} + \hat{A}_2 \Delta y_{t-2} + \cdots + \hat{A}_p \Delta y_{t-p} + u_t^*,$$

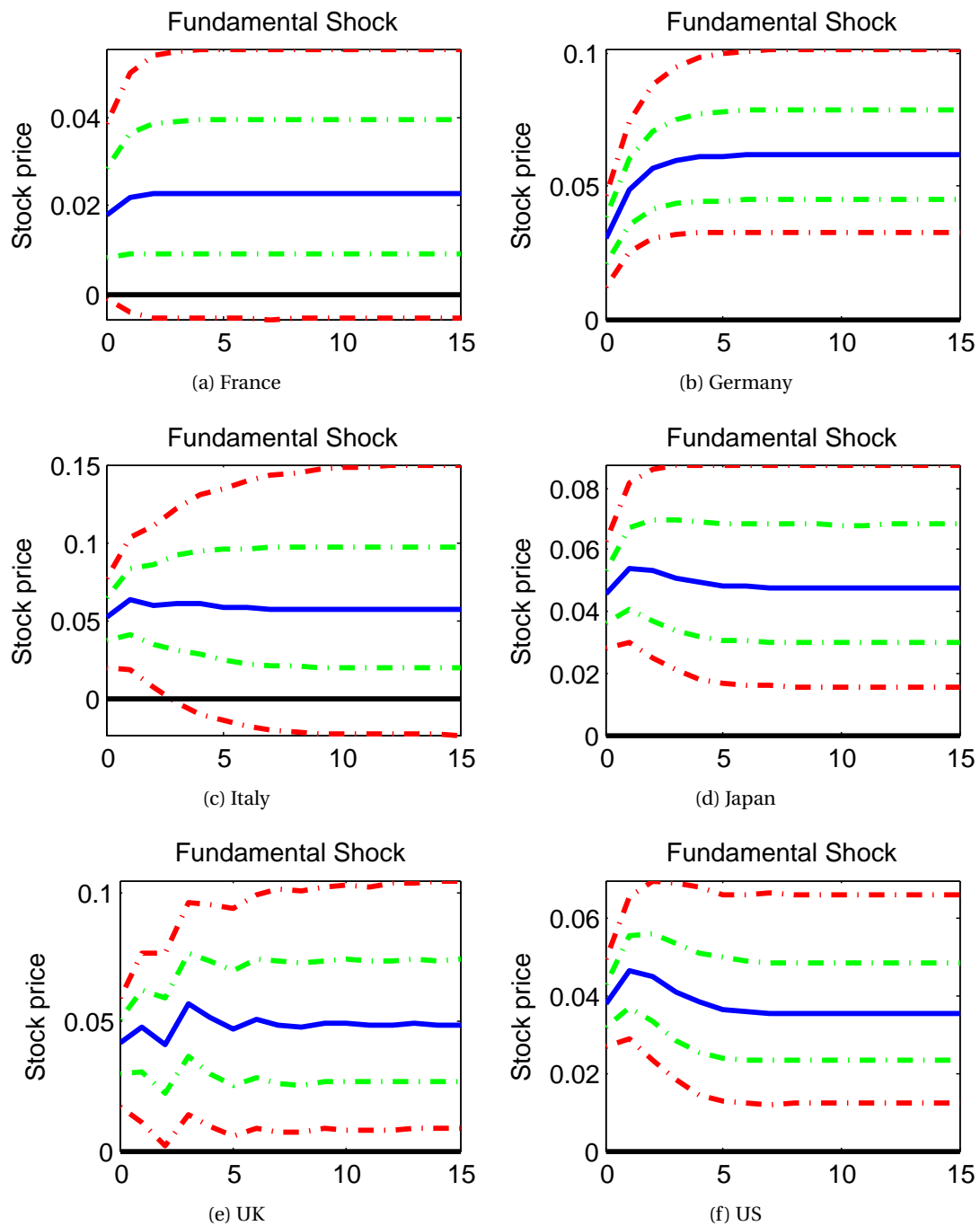


Figure 1.4: Accumulated impulse responses of a fundamental shock to real stock prices. Broken lines indicate Efron confidence intervals according to the fixed design wild bootstrap technique at the 95 and 68 percentiles with 2000 replications.

where $u_t^* = \varphi_t \hat{u}_t$ and where φ_t is a random variable, independent of y_t following a Rademacher distribution. In other words, φ_t is either 1 or -1 with a 50% probability. The hat denotes estimated parameters. Note that since this is a fixed design bootstrap method, in formulating the bootstrap series the lagged values are taken from the original data and not from the lagged bootstrap series.

The accumulated IRs of a fundamental shock on stock prices are shown in Figure 1.4. The response itself is positive in all cases, as one would expect. This is also found by Binswanger (2004a) and Groenewold (2004). Only for France and Italy is the long-run effect of the shock insignificant at the 95% level. However, all IRs in Figure 1.4 are significant at the 68% interval. Hence, there is some evidence that fundamental shocks influence stock prices.

As in Binswanger (2004a), we next conduct forecast error variance decompositions (FEVDs). The FEVD tells us to what extent the structural shocks account for the forecast error of a specific variable. More precisely, we are interested in the extent the forecast error variance of real stock prices is accounted for by fundamental and non-fundamental shocks. Results of the FEVDs of real stock prices for all country SVAR models are given in Table 1.7.

For most country models, the proportion of the forecast error variance of real stock prices accounted for by fundamental shocks is between 30% to 40%. For Germany this is slightly lower and for France this proportion is even under 5%. After about 10 quarters usually, the FEVDs stabilize and the proportion of the forecast error variance explained by each shock stays at a constant level.

The FEVDs for Japan and the US are somewhat similar to the ones in Binswanger (2004a), who also uses data on these countries. More precisely, when using his whole sample he finds a slightly lower percentage of the forecast error variance of real stock prices that can be explained by a fundamental shock for Japan and a slightly higher percentage for the US. Binswanger (2004a) also argues for the existence of a structural break in the early 1980s and therefore splits the sample in two and analyzes the FEVDs of both sub-samples as well. In our analysis however, we do not split the sample into two parts. We choose to work with the whole sample since the purpose of our analysis is to investigate the empirical results of the SVAR models having first determined whether the identifying restriction is supported by the data. In other words,

Table 1.7: FEVD's of the real stock price for all countries. Values are in percent

Quarters ahead	Percentage of variance attributable to:		Quarters ahead	Percentage of variance attributable to:	
	Fundamental shock	Non-fundamental shock		Fundamental shock	Non-fundamental shock
France			Germany		
1	4.51	95.49	1	12.67	87.33
2	4.44	95.56	2	16.30	83.70
3	4.44	95.56	3	16.96	83.04
4	4.44	95.56	4	17.06	82.94
5	4.44	95.56	5	17.07	82.93
10	4.44	95.56	10	17.07	82.93
15	4.44	95.56	15	17.07	82.93
20	4.44	95.56	20	17.07	82.93
Italy			Japan		
1	32.81	67.19	1	32.63	67.37
2	30.24	69.76	2	30.90	69.10
3	29.97	70.22	3	30.71	69.29
4	28.93	71.07	4	30.75	69.25
5	28.65	71.35	5	30.77	69.23
10	28.58	71.42	10	30.78	69.22
15	28.58	71.42	15	30.78	69.22
20	28.58	71.42	20	30.78	69.22
UK			US		
1	31.27	68.73	1	42.23	57.77
2	28.58	71.42	2	39.00	61.00
3	28.97	71.03	3	38.56	61.44
4	31.74	68.26	4	38.75	61.25
5	32.04	67.96	5	38.88	61.12
10	32.32	67.68	10	38.92	61.08
15	32.32	67.68	15	38.92	61.08
20	32.32	67.68	20	38.92	61.08

after we test whether we can label the structural shocks as fundamental and non-fundamental, we then conduct a standard SVAR analysis with these shocks in mind. Further, smoothed probabilities from our MS-SVAR models show weak evidence at best of any globally significant event occurring around the early 1980s. If there indeed is such a phenomenon, it is less pronounced than the financial crisis. Hence, since we have not split the sample to determine whether the FEVDs change over the sub-samples, the results of our FEVDs cannot really be interpreted as being in favor or against a bubble hypothesis in stock prices.

Finally, as in Binswanger (2004a) and Groenewold (2004), we perform a historical decomposition (HD) to see to what extent fundamentals influence real stock prices. The HD technique is not new in the multivariate time series literature. One of the first papers to make use of this approach is Burbidge and Harrison (1985). In this paper we follow their methodology. First, we obtain the estimates of the structural shocks from $\hat{\varepsilon}_t = \hat{B}^{-1} \hat{u}_t$, then we set the non-fundamental shocks to zero after a specified time period and finally we forecast the model from its moving average (MA) representation. The exact formula for the decomposition is given as

$$\Delta y_{T_H+j} - \hat{\mu} = \sum_{s=0}^{j-1} \hat{\Psi}_s \hat{\varepsilon}_{T_H+j-s} + \sum_{s=j}^{\infty} \hat{\Psi}_s \hat{\varepsilon}_{T_H+j-s}, \quad (1.10)$$

where the Ψ s are the structural moving average coefficient matrices defined in (1.4) and T_H is the starting period of the decomposition.

So as to clarify the procedure in a little more detail, we set the non-fundamental shocks to zero since we are only interested in the effect of the fundamental shocks on real stock prices. The starting period of the HD, T_H is unfortunately somewhat arbitrary and does influence the results to a large extent. This problem is also noted by Binswanger (2004a) and Groenewold (2004). Since we cannot forecast (1.10) by going backwards until infinity, we begin with the start of the sample period. Naturally, the first several forecasts will differ from the actual data because of that constraint. After about 20 quarters or 5 years, we find that the forecasted series matches the actual series (even though we do not have an infinite number of observations). Hence, we choose 1965:I as a starting date for our historical decomposition; recall that it is 5 years after the beginning of the sample period of 1960:I. This makes the choice of a starting date somewhat less arbitrary and guarantees that the historical series matches the actual series at T_H . Finally, for better comparison, HDs in this paper are constructed by integrating the historical series and the demeaned actual series forward.

Results of the historical decompositions for all country models are displayed in Figure 1.5. The fundamentals series refers to the decomposed series in which only fundamental shocks influence stock prices, i.e. with the non-fundamental shocks set to 0. The demeaned actual real stock price series is also given for comparison. From the figure there are two quite apparent things. First, from the early 1970s through

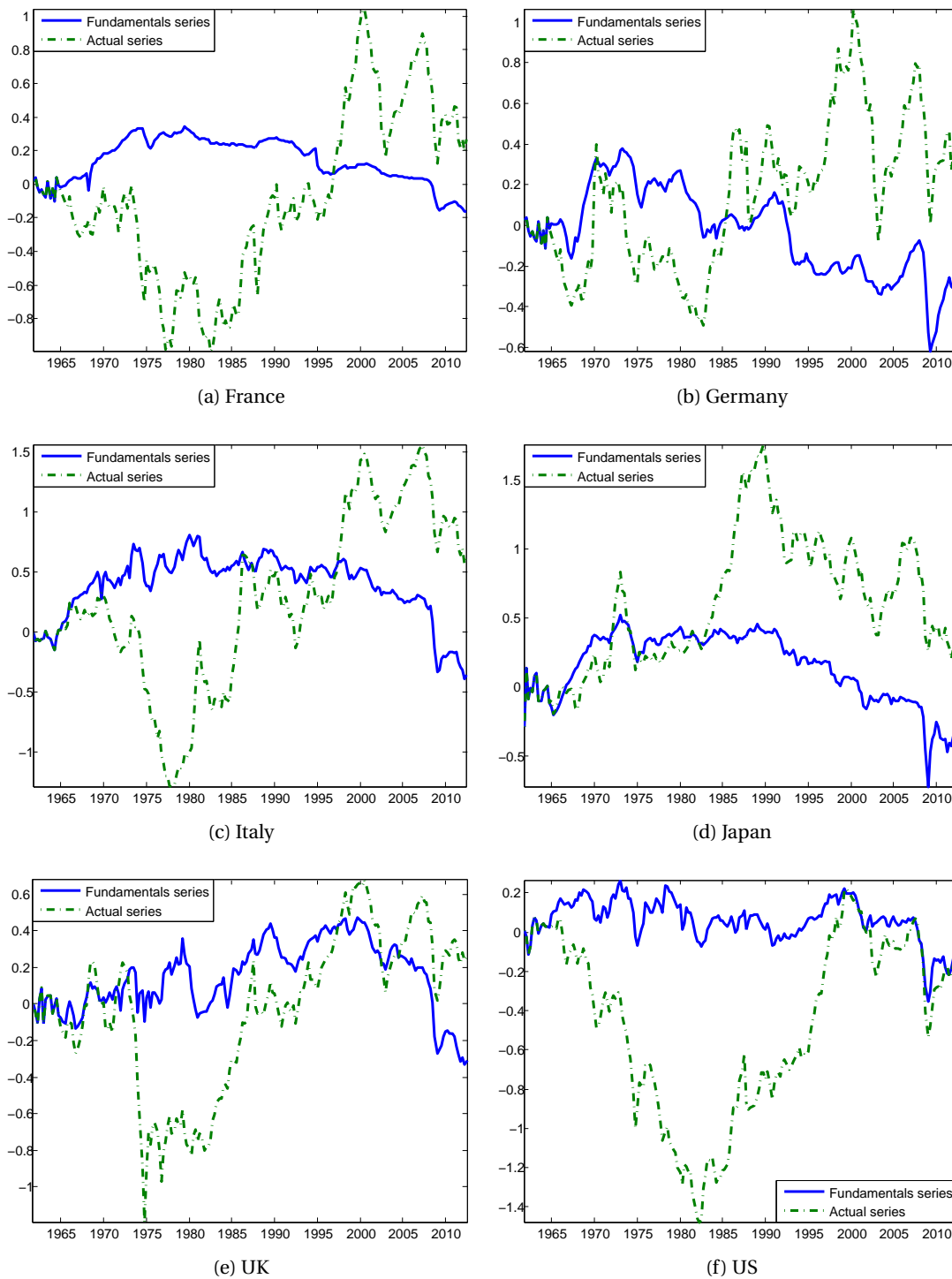


Figure 1.5: Historical decompositions of the real stock price according to fundamentals along with the actual demeaned stock price series for each country model

the mid 1980s, in most cases the fundamentals series is above the level of the actual series. For France and the US this phenomenon lasts well until the mid 1990s. This could indicate that stock prices may have been undervalued during that time period, which is also the time of the oil crises (and the early 1980s and 1990s recession in the US). Japan is an exception to this in that during that time period the fundamentals and the actual series move closely together. Second, after the time period in which stock prices appear to be undervalued, the actual series in most cases starts to go above the series that can be explained only by fundamentals. This effect is especially pronounced for Japan and could indicate that stock prices have started to become overvalued or even of a potential bubble in stock prices. For the UK and the US this effect is not observable, in the sense that the actual series appears to be inline with the fundamentals series. We therefore reach the same conclusion as Binswanger (2004a) for Japan and a different one for the US, more similar to Rapach (2001).

The results of the HDs however may not be all that reliable. As noted before, the starting date of the decompositions plays a large role in the final outcome. For instance, had we started the HDs in the early 1980s, as is done by Binswanger (2004a), we would have come to the same conclusion as him for the US; namely that the actual stock price series is well above the fundamentals series. Hence, it is not an easy task to draw conclusions solely based on the results of historical decompositions. Nevertheless, given our results, we do not find real evidence of a bubble in stock prices for the UK and the US, while there may be one such for Japan. It is less clear for the remaining three countries.

1.5 Conclusion

Structural VAR (SVAR) models are a popular tool in the multivariate time series literature since they allow for the investigation of economic shocks of interest. They rely however on identifying restrictions in order to estimate the structural parameters. This paper makes use of a Markov switching-SVAR (MS-SVAR) model in heteroskedasticity as in Lanne et al. (2010) and Herwartz and Lütkepohl (2011) to test a long-run restriction that would just-identify a SVAR model in the traditional sense.

More precisely, we investigate a bivariate SVAR model already considered in the

literature by Binswanger (2004a) and Groenewold (2004), which tries to answer the question of how well stock prices reflect their fundamentals. One restriction on the long-run impact matrix as in Blanchard and Quah (1989) is enough to identify the structural parameters of the model.

Using data from six different countries, we find that for five of the country models the long-run restriction is supported at the 5% level. Two robustness specifications largely reinforce our conclusion of accepting the long-run restriction. Therefore, we decide that there is enough support from the data to warrant the identifying restriction. Accordingly, we label the two structural shocks as fundamental and non-fundamental.

Having identified the structural shocks we analyze the conventional SVAR empirical simulation results. Impulse responses show that in most cases there are significant long-run effects of a fundamental shock to real stock prices at the 95% confidence band, constructed using the fixed design wild bootstrap method. Forecast error variance decompositions indicate that most of the forecast error variance of real stock prices is explained by non-fundamental shocks. Finally, historical decompositions indicate a potential bubble in real stock prices for Japan and pricing according to fundamentals for the UK and the US.

Even though the long-run restriction in almost all cases is supported by the data, it does not necessarily have to be the case. Any standard SVAR analysis could potentially suffer from inappropriate identification restrictions. Hence, this paper illustrates the virtue of being able to test structural restrictions in order to justify the relevant shocks of interest.

1.6 Appendix

This is a technical appendix explaining the EM algorithm used in this paper in more detail. It is largely based on Krolzig (1997) and for more details the reader is referred to that work and to Chapter 22 of Hamilton (1994). Here T denotes the sample size, K the number of variables, p the number of lags and M the number of states.

1.6.1 Definitions

For the expectation step define

$$\hat{\xi}_{t|t} = \begin{bmatrix} P(S_t = 1|\Delta Y_t) \\ P(S_t = 2|\Delta Y_t) \\ \vdots \\ P(S_t = M|\Delta Y_t) \end{bmatrix}, \quad (1.11)$$

an $(M \times 1)$ vector of conditional probabilities of being in a particular state at time period t given all observations up to that time period, i.e. $\Delta Y_t = [\Delta y_1, \Delta y_2, \dots, \Delta y_t]$. These are also referred to as the filtered probabilities of a MS model. Further, the conditional densities of an observation given a particular state, all past observations and all SVAR parameter estimates, θ are defined as

$$\eta_t = \begin{bmatrix} P(\Delta y_t|S_t = 1, \Delta Y_{t-1}, \theta) \\ P(\Delta y_t|S_t = 2, \Delta Y_{t-1}, \theta) \\ \vdots \\ P(\Delta y_t|S_t = M, \Delta Y_{t-1}, \theta) \end{bmatrix} = \begin{bmatrix} \frac{1}{2\pi|\Sigma(1)|^{1/2}} \exp\left\{-\frac{u_t' \Sigma(1)^{-1} u_t}{2}\right\} \\ \frac{1}{2\pi|\Sigma(2)|^{1/2}} \exp\left\{-\frac{u_t' \Sigma(2)^{-1} u_t}{2}\right\} \\ \vdots \\ \frac{1}{2\pi|\Sigma(M)|^{1/2}} \exp\left\{-\frac{u_t' \Sigma(M)^{-1} u_t}{2}\right\} \end{bmatrix}. \quad (1.12)$$

Here θ consists of the vectorized SVAR parameters, i.e. $v, A_i, \Lambda_j, B, i = 1, \dots, p, j = 2, \dots, M$.

For the maximization step define

- $\Delta y = [\Delta y_1', \dots, \Delta y_T']'$, a $(TK \times 1)$ vector of endogenous variables
- $\bar{Z} = [\mathbf{1}_T, \Delta Y_{-1}, \dots, \Delta Y_{-p}]$, a $(T \times (1 + KP))$ matrix
- $\Delta Y_{-i} = [\Delta y_{1-i}, \dots, \Delta y_{T-i}]', i = 1, \dots, p$, a $(T \times K)$ matrix of lagged regressors
- $\beta = \text{vec}[v, A_1, \dots, A_p]$, a $(K(Kp + 1) \times 1)$ vector of the parameters in (1.1)
- u , a $(TK \times 1)$ vector of residuals, $u_i, i = 1, \dots, T$ distributed according to (1.6).

Then (1.1) can be rewritten as

$$\Delta y = (\bar{Z} \otimes I_K) \beta + u,$$

where \otimes denotes the Kronecker product.

1.6.2 Starting Values

- $\beta_0 = [\bar{Z}'\bar{Z} \otimes I_K]^{-1}(\bar{Z}' \otimes I_K)\Delta y$
- $B_0 = (UU'/T)^{1/2}$, where U is obtained from $u = \Delta y - (\bar{Z} \otimes I_K)\beta_0$, where $u = \text{vec}(U)$
- $P_0 = \mathbf{1}_M \mathbf{1}_M' / M$, $\mathbf{1}_M$ is an $(M \times 1)$ vector of ones
- $\xi_{0|0} = \mathbf{1}_M / M$
- $\Lambda_2 = i \times I_K, \Lambda_2 = i^2 \times I_K, \dots, \Lambda_M = i^M \times I_K, i = 2, \dots$. Different values of i are used to determine which gives the highest sensible log-likelihood.

1.6.3 Expectation Step

Calculate the filtered probabilities, (1.11) from (1.12) as

$$\hat{\xi}_{t|t} = \frac{\hat{\eta}_t \odot \hat{\xi}_{t|t-1}}{\mathbf{1}_M'(\hat{\eta}_t \odot \hat{\xi}_{t|t-1})}, \quad (1.13)$$

where

$$\hat{\xi}_{t|t-1} = \hat{P}' \hat{\xi}_{t-1|t-1}, \quad (1.14)$$

for $t = 1, \dots, T$. This generates an $(M \times 1)$ vector of conditional probabilities for each time period. Here \odot denotes element-by-element multiplication and \hat{P} is defined as in (1.7). The sum of the denominators for $t = 1, \dots, T$ in (1.13) is the likelihood of the MS-SVAR model as noted in Hamilton (1994). Using the estimated filtered probabilities, the smoothed probabilities, conditional on all observations up to time T , $P(S_t = i | \Delta Y_T), i = 1, \dots, M$ are estimated as

$$\hat{\xi}_{t|T} = \left[\hat{P} \left(\hat{\xi}_{t+1|T} \odot \hat{\xi}_{t+1|t} \right) \right] \odot \hat{\xi}_{t|t}, \quad (1.15)$$

for $t = T-1, \dots, 0$ where $\hat{\xi}_{T|T}$ is taken from the last iteration in (1.13). The symbol \oslash denotes element-by-element division.

1.6.4 Maximization Step

Calculate $\rho = \text{vec}(P)$ using (1.13), (1.14) and (1.15) as

$$\hat{\rho} = \hat{\xi}^{(2)} \oslash (\mathbf{1}_M \otimes \hat{\xi}^{(1)}), \quad (1.16)$$

where $\hat{\xi}^{(1)} = (\mathbf{1}'_M \otimes I_M) \hat{\xi}^{(2)}$, $\hat{\xi}^{(2)} = \sum_{t=0}^{T-1} \hat{\xi}_{t|T}^{(2)}$ and

$$\hat{\xi}_{t|T}^{(2)} = \text{vec}(\hat{P}) \odot \left[\left(\hat{\xi}_{t+1|T} \otimes \hat{\xi}_{t+1|t} \right) \otimes \hat{\xi}_{t|t} \right].$$

Using (1.15), estimate B and $\Lambda_i, i = 2, \dots, M$ by optimizing

$$\begin{aligned} l(B, \Lambda_2, \dots, \Lambda_M) &= T \log |\det(B)| + \frac{1}{2} \text{tr} \left((BB')^{-1} \hat{U} \hat{\Xi}_1 \hat{U}' \right) \\ &+ \sum_{m=2}^M \left[\frac{\hat{T}_m}{2} \log(\det(\Lambda_m)) + \frac{1}{2} \text{tr} \left((B \Lambda_m B')^{-1} \hat{U} \hat{\Xi}_m \hat{U}' \right) \right], \end{aligned} \quad (1.17)$$

where \hat{U} is obtained from $\hat{u} = \Delta y - (\bar{Z} \otimes I_K) \hat{\beta}$ where $\hat{u} = \text{vec} \hat{U}$, $\hat{\Xi}_m = \text{diag}(\hat{\xi}_{m1|T}, \dots, \hat{\xi}_{mT|T})$, the smoothed probabilities of regime m and $\hat{T}_m = \sum_{t=1}^T \hat{\xi}_{mt|T}$. To avoid singularity a lower bound of 0.01 is imposed on the diagonal elements of the $\Lambda_m, m = 2, \dots, M$ matrices. Using the estimates from (1.17), the updated covariance matrices are then derived as in (1.8) as

$$\hat{\Sigma}(1) = \hat{B} \hat{B}', \quad \hat{\Sigma}(2) = \hat{B} \hat{\Lambda}_2 \hat{B}', \quad \dots \quad \hat{\Sigma}(M) = \hat{B} \hat{\Lambda}_M \hat{B}'. \quad (1.18)$$

Using (1.15) and (1.18) calculate the remaining SVAR parameters as

$$\hat{\beta} = \left[\sum_{m=1}^M (\bar{Z}' \hat{\Xi}_m \bar{Z}) \otimes \hat{\Sigma}^{-1}(m) \right]^{-1} \left[\sum_{m=1}^M (\bar{Z}' \hat{\Xi}_m) \otimes \hat{\Sigma}^{-1}(m) \right] \Delta y. \quad (1.19)$$

Obtain a new \hat{U} using $\hat{\beta}$ from (1.19) and keep on re-estimating (1.17), (1.18) and (1.19) until a convergence criterion is met.

Finally, using (1.15),

$$\hat{\xi}_{0|0} = \hat{\xi}_{0|T}. \quad (1.20)$$

1.6.5 Convergence

The expectation and maximization steps are iterated until convergence. Recall, the log-likelihood of the MS-SVAR model is given by

$$l(\theta | \Delta Y_T) = \sum_{t=1}^T \ln \left(\mathbf{1}'_M \left(\hat{\eta}_t \odot \hat{\xi}_{t|t-1} \right) \right). \quad (1.21)$$

We use the absolute change in the log-likelihood as a convergence criterion in the maximization step and for the EM algorithm as a whole, i.e.

$$\Delta = \left| l(\bullet)^{j+1} - l(\bullet)^j \right|, \quad (1.22)$$

where $l(\bullet)^j$ is the log-likelihood given by (1.17) or (1.21) for the j -th iteration. Convergence is satisfied when $\Delta \leq 10^{-6}$ or after a specified maximum number of iterations.

1.6.6 Switching Intercept

The model with a switching intercept term as well as a switching covariance matrix, given in (1.9) is easily calculated with a small modification to the \bar{Z} matrix. Namely,

$$\bar{Z}_m = [\mathbf{1}_T \otimes \iota'_m, \Delta Y_{-1}, \dots, \Delta Y_{-p}], m = 1, \dots, M,$$

where ι_m is the m -th column of the M -dimensional identity matrix. The above matrix is of a $(T \times (M + KP))$ dimension. Hence, $\hat{\beta} = \text{vec}[\hat{v}(1), \dots, \hat{v}(M), \hat{A}_1, \dots, \hat{A}_p]$ in (1.19).

1.6.7 Standard Errors

Upon convergence of the EM algorithm, the optimal values of $\beta, B, \Lambda_m, m = 2, \dots, M, \xi_{0|0}$ and the $M(M - 1)$ unrestricted parameters in P are used in a function to calculate the log-likelihood in (1.21) from (1.12), (1.13) and (1.14). Using this function, standard errors of all unrestricted parameters are obtained by the inverse of the negative of the Hessian matrix.

Chapter 2

Can Stock Price Fundamentals Really be Captured? Using Markov Switching in Heteroskedasticity Models to Test Identification Restrictions

Anton Velinov (EUI)

Abstract

This paper tests a commonly used structural parameter identification scheme to try and determine whether it can really capture fundamental and non-fundamental shocks to stock prices. In particular, five trivariate versions of the dividend discount model are considered, which are widely used in the literature. They are either specified in vector error correction (VEC) or in vector autoregressive (VAR) form. Restrictions on the long-run effects matrix as in Blanchard and Quah (1989) are used to identify the structural parameters. These identifying restrictions are tested by means of a Markov switching in heteroskedasticity model as in Lanne et al. (2010) and Herwartz and Lütkepohl (2011). It is found that for two of the models considered, the long-run identification scheme appropriately classifies shocks as being either fundamental or non-fundamental. Those are the models with real GDP and real dividends as proxies of real economic activity. A series of robustness tests are performed, which largely confirm the original findings. Results of this paper serve as a good guideline when conducting future research in this field.

2.1 Introduction

The question of how well stock prices reflect their underlying economic fundamentals is not a new one. This paper aims to test some of the models in the existing literature to determine whether they really do capture fundamental and non-fundamental shocks as they claim. In particular, many popular multivariate structural models are considered. In this sense, the approach of this paper is similar to that of Binswanger (2004b) and Jean and Eldomiaty (2010), though only in so far as the structural models themselves are concerned. In other words, this analysis is only focused on determining whether the structural parameter identification schemes are correct as opposed to considering how well stock prices are reflected by their fundamentals. A proper identification scheme, one that is supported by the data, is crucial in any analysis using structural models.

Structural vector autoregressive (SVAR) and error correction (SVEC) models have extensively been used in the literature to determine stock price fundamentals. Some notable contributions are by Lee (1995a,b, 1998), Chung and Lee (1998), Rapach (2001), Binswanger (2004a,b,c), Allen and Yang (2004), Laopodis (2009) and Jean and Eldomiaty (2010). All of these models make use of long-run restrictions on the parameters as in Blanchard and Quah (1989) to identify the structural shocks. These models are not identical in the number of variables used or in the time periods considered. This makes results difficult to compare, as noted by Binswanger (2004b) and Jean and Eldomiaty (2010). Nevertheless, there are some prominent (subset) models, all based on the dividend discount model (DDM), which are considered here. These models are summarized in Table 2.1.

The DDM is popular in asset pricing. Its basic premise is that an asset's price is the sum of its expected future discounted payoffs (i.e. dividends). These payoffs are necessarily linked to real economic activity such as real GDP or industrial production. Hence, Table 2.1 can be thought of as summarizing multiple variants of the DDM, which is a widely used model in the literature.

Almost all the models considered in Table 2.1 use a lower triangular long-run effects matrix to identify the structural shocks. This paper is concerned with analyzing these identifying restrictions in order to determine whether or not they truly capture fundamental and non-fundamental shocks to stock prices. Even though these mod-

Table 2.1: Popular models used in the literature.

(Subset) Model	Used by
$y_t = [Y_t, r_t, s_t]'$	Lee (1995a), Rapach (2001),* Binswanger (2004b), Jean and Eldomiaty (2010) Lanne and Lütkepohl (2010)
$y_t = [IP_t, r_t, s_t]'$	Binswanger (2004b), Laopodis (2009),* Jean and Eldomiaty (2010)
$y_t = [D_t, r_t, s_t]'$	Lee (1995a), Allen and Yang (2004), Jean and Eldomiaty (2010)
$y_t = [E_t, r_t, s_t]'$	Binswanger (2004b), Jean and Eldomiaty (2010)
$y_t = [E_t, D_t, s_t]'$	Lee (1998), Chung and Lee (1998), Binswanger (2004b), Jean and Eldomiaty (2010)

Here Y_t , IP_t , D_t , E_t , r_t and s_t stand for real GDP, industrial production index, real dividends, real earnings, real interest rates and real stock prices respectively.

* These variables are a subset of the variables used in the original model.

els can all be thought of as similar - in terms of being derived from the DDM - imposing the same type of restrictions when using diverse data may lead to inadequate identification of shocks.

A method of testing whether structural restrictions are supported by the data is provided in Lanne et al. (2010) and in Herwartz and Lütkepohl (2011). Both extend the basic multivariate structural model to allow for switching covariance matrices according to a Markov process. This allows for heteroskedastic error terms across states. Due to the extra covariance parameters, any structural restrictions become over-identifying and can thus be tested. This procedure is explained in detail in section 2 along with the relevant models used.

After elaborating on the models, section 3 discusses the estimation and testing procedure briefly. Section 4 presents the model results along with relevant details on the model selection procedure. Section 5 deals with model robustness issues and finally section 6 summarizes the main conclusions.

2.2 The Models

This section briefly sets out the structural vector autoregressive (SVAR) and error correction (SVEC) models that are considered. Later a regime switching extension is

introduced and the basic methodology for testing the identifying restrictions is explained.

2.2.1 The SVAR and SVEC models

The basic vector autoregressive model with p lags, VAR(p) is summarized in equation (2.1).

$$y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t, \quad (2.1)$$

where y_t is a $(K \times 1)$ vector of stationary endogenous variables, v is a $(K \times 1)$ vector of constants and $A_i, i = 1, \dots, p$ are $(K \times K)$ autoregressive parameter matrices. The $(K \times 1)$ vector of reduced form error terms, u_t is assumed to have an expected value of 0 and a positive definite covariance matrix Σ_u . Hence, $u_t \sim (0, \Sigma_u)$.

A reduced form vector error correction model, (VEC($p-1$)) is given as follows

$$\Delta y_t = v_t + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t, \quad (2.2)$$

where now y_t may include variables with unit roots. Here v_t is a K dimensional deterministic component that can include an intercept and a trend term, hence $v_t = v_0 + v_1 t$. Further, $\Gamma_i, i = 1, \dots, p-1$ are $(K \times K)$ parameter matrices and the residual terms, u_t are assumed to have the same properties as before. Here Δ is the first difference operator (so that $\Delta y_t = y_t - y_{t-1} = (1 - L)y_t$, where L is the lag operator). This means that Δy_t is assumed to be $I(0)$, such that Πy_{t-1} also needs to be stationary. The $(K \times K)$ matrix Π is of rank r , (where $0 < r < K$) and captures the cointegrating relations of the model. More specifically, since Π is singular, it can be decomposed into the product of two $(K \times r)$ matrices of full column rank, α, β so that $\Pi = -\alpha\beta'$. Here β is referred to as the cointegrating matrix and has the r linearly independent cointegrating relations, so that $\beta' y_{t-1}$ is stationary, and α is known as the loading matrix.

In line with the literature, structural shocks are given as $u_t = B\varepsilon_t$, where ε_t is a K dimensional vector of structural residuals such that $\varepsilon_t \sim (0, \Sigma_\varepsilon)$, where Σ_ε is usually assumed to be I_K , the identity matrix. Here B is a $(K \times K)$ matrix depicting contemporaneous effects. According to these assumptions $\Sigma_u = BB'$. The structural parameters can be derived from the reduced form parameters. However, since Σ_u is symmetric,

this only leaves $K(K+1)/2$ reduced form parameters to identify the K^2 structural parameters of the B matrix. Hence, $K^2 - K(K+1)/2 = K(K-1)/2$ restrictions need to be imposed. How this is done for each model is discussed in the following.

Restrictions on the VAR model

All the papers considered in Table 2.1 make use of long run identifying restrictions, as in Blanchard and Quah (1989). How such restrictions are imposed is briefly explained here. Rewriting equation (2.1) in lag polynomial form gives

$$A(L)y_t = v + u_t, \quad (2.3)$$

where $A(L) = I_K - A_1L - A_2L^2 - \dots - A_pL^p$. Provided that $A(L)^{-1}$ exists, the Wold moving average (MA) representation for the stationary y_t process is

$$y_t = \mu + \sum_{s=0}^{\infty} \Phi_s u_{t-s} = \mu + \Phi(L)u_t, \quad (2.4)$$

where $\mu = (I_K - A_1 - A_2 - \dots - A_p)^{-1}v = A(1)^{-1}v$, $\Phi(L) \equiv A(L)^{-1}$ and $\Phi_0 = I_K$. Having defined the structural shocks as $\varepsilon_t = B^{-1}u_t$, the structural representation of (2.4) is

$$y_t = \mu + \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s} = \mu + \Psi(L)\varepsilon_t, \quad (2.5)$$

here $\Psi_i \equiv \Phi_i B$, for $i = 0, 1, 2, \dots$. The accumulated long-run effects of the structural shocks over all time periods are given by the long-run impact matrix, $\Psi \equiv \Phi B$, where $\Phi \equiv \sum_{s=0}^{\infty} \Phi_s = A(1)^{-1}$. It is on the Ψ matrix that Blanchard and Quah (1989) suggest imposing identifying restrictions, usually in the form of zeros. That way some shocks have permanent effects, while others only have transitory effects.

As is common practice, most papers mentioned in Table 2.1 make use of a lower triangular Ψ matrix. In this case the matrix would appear as follows

$$\Psi = \begin{bmatrix} \star & 0 & 0 \\ \star & \star & 0 \\ \star & \star & \star \end{bmatrix}, \quad (2.6)$$

where \star denotes an unrestricted element. Depending on the way the variables are arranged this identification scheme distinguishes between fundamental and non-fundamental shocks. The non-fundamental shock is assumed not to have any permanent effect on any of the variables except the last one (last column of (2.6)). The

other two shocks are assumed to be of a fundamental nature; in that one of them (first column of (2.6)) influences all variables in the long-run, while the other (second column of (2.6)) only leaves a permanent impact on the last two model variables.

¹ The identification scheme in (2.6) is used for testing restrictions on SVAR models throughout this paper.

Restrictions on the VEC model

The long-run effects matrix for a VEC model is not derived in such a straightforward way as that for a VAR model. Fortunately, from Granger's representation theorem, the VEC counterpart of Φ is given as

$$\Xi = \beta_{\perp} \left[\alpha'_{\perp} \left(I_K - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_{\perp} \right]^{-1} \alpha'_{\perp},$$

where \perp stands for the orthogonal complement of a given matrix. For instance, the orthogonal complement of an $(m \times n)$ matrix, A , is given by the $(m \times (m - n))$ matrix, A_{\perp} . The Ξ matrix is computed from the estimates of the reduced form parameters. Hence, the long-run impact matrix is ΞB and restrictions can be imposed on it in a similar way as on the Ψ matrix above.

A quick note on the restrictions of the SVEC model is in order. Since Ξ is a singular matrix, restrictions need to be placed appropriately. In particular, the rank of Ξ is $K - r$ and according to King et al. (1992) there can be at most r transitory shocks, i.e. r columns of ΞB can be 0 and each column of zeros stands for only $K - r$ restrictions. In addition, there need to be $r(r - 1)/2$ restrictions on the B matrix to identify the non-permanent shocks. The remaining restrictions needed to identify the model (exactly) can be placed on the non-zero elements of ΞB or directly on B . A good summary of placing restrictions on a SVEC model can be found in Lütkepohl (2005).

In a similar lower triangular fashion, long-run restrictions on SVEC models in this

¹The zero restriction in the second column of Ψ in (2.6) is left out in Lee (1995a) and Laopodis (2009). The shocks are still identified as fundamental and non-fundamental, even though the model itself is underidentified. Further, the models used in Jean and Eldomiaty (2010) are initially identified according to the Swanson and Granger (1997) identification scheme, however, in a section on model robustness, they mention that a lower triangular long-run impact matrix as in (2.6) performs equally well.

paper are placed as follows

$$\Xi B = \begin{bmatrix} \star & 0 & 0 \\ \star & \star & 0 \\ \star & \star & 0 \end{bmatrix}. \quad (2.7)$$

Here again \star denotes unrestricted elements. It is now assumed that a non-fundamental shock does not have permanent effects on any of the other variables, i.e. the last column of (2.7) contains only zeros. Note that such an assumption cannot be made for the SVAR model restrictions since Ψ in (2.6) cannot be a singular matrix. It may be more realistic to assume that shocks labeled as non-fundamental do not have a permanent impact on any of the model variables. Further, since the rank of Ξ is $K - r$, which is two (as will become apparent later, in all cases $r = 1$), the column of zeros provides two independent restrictions. The identification scheme in (2.7) is therefore enough to just identify the SVEC model in the traditional sense.

2.2.2 The Markov switching SVAR and SVEC models

In order to test identification schemes, as in (2.6) or in (2.7), it is necessary to expand the basic model to allow at least for switching covariance matrices. Further, for estimation convenience it is also assumed that residuals are normally distributed, hence,

$$u_t \sim \text{NID}(0, \Sigma_u(S_t)). \quad (2.8)$$

As is made clear in Lanne et al. (2010), the normality assumption in no way limits the unconditional distribution and it is also not a crucial assumption for the analysis. Here S_t is assumed to follow a first-order discrete valued Markov process with transition probabilities given by

$$p_{ij} = P(S_t = j | S_{t-1} = i).$$

These can be grouped in an $(M \times M)$ matrix of transition probabilities, P such that the rows add up to 1 and where M are the number of different states.

Note that it is also possible to allow for switches in the intercept term, v in the SVAR case and v_0 in the SVEC case. In principle, all the parameters could be subject to regime switches, however such assumptions would need to be justified in the sense of there being structural breaks in the data or some reasonable economic explanation

as to why a certain parameter could be switching. In this analysis it is crucial for the covariance matrices to be switching, it may also be reasonable to assume - given the data used - that the intercept parameter could be subject to regime switches as will be discussed later. All other parameters are assumed to be stable.

As already noted, the Markov switching (MS) model is a convenient way of dealing with data subject to structural breaks. In the literature changes in structural relationships are documented in Lee (1998), Chung and Lee (1998), Binswanger (2000, 2004a,c) and Laopodis (2009) among others. In this sense, a MS model may be better suited to answering the question of how well stock prices are reflected by their fundamentals. However, in this paper a MS model is used solely to test the above-mentioned identifying restrictions.

Due to the many intricacies of the models described thus far it is worth elaborating on how the model parameters are estimated and how the identifying restrictions are tested. This is done in the following section.

2.3 Estimation and Testing Procedure

The VAR parameters are estimated by means of OLS. Since only long-run restrictions are imposed, estimation of the structural parameters is straightforward. With a simple substitution it follows that $\Phi \Sigma_u \Phi' = \Psi \Psi'$. The left hand side of this equation is known, hence for a fully identified model, Ψ is easy to derive. The contemporaneous matrix is then easily obtained as $B = \Phi^{-1} \Psi$.

The VEC parameters are estimated by the method of reduced rank regression in Johansen (1995). Since the cointegrating matrix, β , is not unique it can be identified by a simple normalization such that the first r rows contain an $(r \times r)$ identity matrix, as is shown in Lütkepohl (2005). The structural parameters are estimated by an iterative algorithm proposed by Amisano and Giannini (1997) subject to identifying restrictions placed as in Vlaar (2004).

The parameters of the MS models are estimated using the iterative expectation maximization (EM) algorithm. This algorithm was initially popularized by Hamilton (1994) for univariate processes and later extended to multivariate processes by Krolzig (1997). Since the β matrix in the VEC models symbolizes long-run relation-

ships, it is not re-estimated at each maximization step of the EM algorithm. It is trivial to change this though, so that a reduced rank regression is performed in each maximization step. However, this merely leads to an increase in computational time without really affecting the conclusions obtained below.

In order to test the identifying restrictions it is necessary to decompose the covariance matrices in the following way

$$\Sigma_u(1) = BB', \quad \Sigma_u(2) = B\Lambda_2B', \quad \dots \quad \Sigma_u(M) = B\Lambda_MB'. \quad (2.9)$$

This adds to a nonlinear optimization procedure in the maximization step of the EM algorithm. The underlying assumption is that the contemporaneous effects matrix, B stays the same across states. Here the $\Lambda_i, i = 2, \dots, M$ matrices are diagonal with positive elements, $\lambda_{ij}, i = 2, \dots, M, j = 1, \dots, K$ and can be interpreted as relative variance matrices. In order for the B matrix in (2.9) to be unique up to changes in sign and column ordering, it is necessary for all pairwise diagonal elements in at least one of the $\Lambda_i, i = 2, \dots, M$ matrices to be distinct. For example, for a 3-state model it is required that $\lambda_{ij} \neq \lambda_{il}, i = 2$ and/or $3, j, l = 1, \dots, K, j \neq l$. Hence, even if these elements are equal in one state, they should not be equal in the other state. For a more detailed explanation of the uniqueness of the B matrix the reader is referred to Proposition 1 in the appendix of Lanne et al. (2010). If this distinction requirement is fulfilled, then B is said to be identified through heteroskedasticity.

The assumption of an invariant B matrix may seem rather crucial to this analysis. However, when there are more than two states, this assumption can be tested by means of a likelihood ratio (LR) test. The test statistic has an asymptotic χ^2 distribution with $(1/2)MK(K+1) - K^2 - (M-1)K$ degrees of freedom. Clearly, if $M = 2$ the degrees of freedom would be 0, thus giving a nonsensical result. Note, that the above procedures closely follow Lanne et al. (2010) and Herwartz and Lütkepohl (2011).

Standard errors of the parameter estimates are obtained from the inverse of the negative of the Hessian matrix evaluated at the optimum. Distinction of the $\lambda_{ij}, i = 2, \dots, M, j = 1, \dots, K$ parameters is then determined through Wald tests. It is also possible to use LR tests for this purpose, however, such tests may not give very accurate conclusions since they can potentially converge to the same optimum each time.²

²For instance, the LR test proceeds by forcing two diagonal elements of $\Lambda_i, i = 2, \dots, M$ to be equal and then

If the distinction of the $\lambda_{ij}, i = 2, \dots, M, j = 1, \dots, K$ parameters is fulfilled, the B matrix would be identified up to changes in sign and column ordering. Hence, any restrictions (short or long-run) would be over-identifying and could therefore be tested. This is achieved by estimating the model with and without restrictions on the B matrix and comparing both log-likelihoods. In other words, an LR test is used and the test statistic has an asymptotic χ^2 distribution with degrees of freedom equal to the number of restrictions.

2.4 Model Results

Before the long-run identification schemes can be tested, it is necessary to determine whether a VAR or a VEC model would be best suited to the data.

2.4.1 The Data and Model Specification

Most data are from the Federal Reserve Economic Database (FRED). The dividends and earnings data are from Robert Schiller's webpage.³ In line with many of the papers mentioned in Table 2.1, the data is quarterly. The data range is from 1947:I - 2012:III, with the exception of dividends and earnings, which are until 2012:I. All variables are in real terms (except for the industrial production (IP) index) and in logs (except for the interest rate series). The interest rate is transformed to real terms by subtracting the CPI growth rate. Other variables are transformed to real terms by dividing by the percent level of the CPI. Figure 2.1 plots the data used along with recession periods according to NBER dating marked by the shaded bars.

All variables are $I(1)$, meaning that they contain a unit root. This is according to Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

comparing the log-likelihoods of the restricted and unrestricted models. This is done until all pairwise combinations of elements are exhausted. However, due to the highly nonlinear nature of the models, it is not uncommon for the EM algorithm to converge to the same parameter estimates and log-likelihood values for different pairwise combinations tests. In other words, when testing for the equality of pairwise $\Lambda_i, i = 2, \dots, M$ parameters, the EM algorithm could potentially converge to the same values over different pairwise tests, thereby giving the same results in each LR test. When working with trivariate models, three pairwise combinations exists, and it is usually found that the EM algorithm converges to the same optimum usually for two out of the three cases. Hence, two LR tests would have the same values. Therefore, Wald tests are deemed more reliable.

³Found at <http://www.econ.yale.edu/shiller/data.htm>.

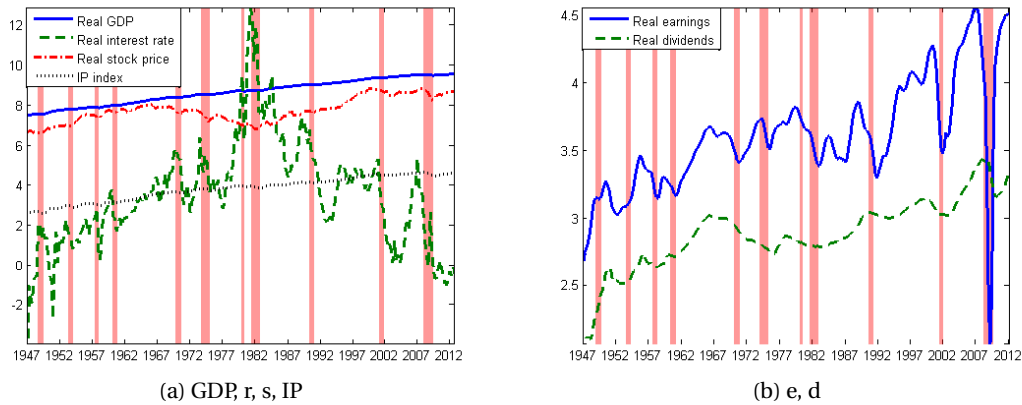


Figure 2.1: Data used with recession dates indicated by the bars.

Table 2.2: Summary of VAR and VEC models.

Model I	$y_t = [Y_t, r_t, s_t]'$	VAR in first differences
Model II	$y_t = [IP_t, r_t, s_t]'$	VAR in first differences
Model III	$y_t = [D_t, r_t, s_t]'$	VAR in first differences
Model IV	$y_t = [E_t, r_t, s_t]'$	VEC, $r = 1$
Model V	$y_t = [E_t, D_t, s_t]'$	VEC, $r = 1$

tests. This is true even for the real interest rate series, although only at the 10% level according to the ADF test. Both the Johansen (1995) trace test and the Saikkonen and Lütkepohl (2000) tests are used to test for cointegration. According to these tests, only models IV and V show signs of cointegration and in each case the cointegrating rank, r , is 1. Table 2.2 summarizes the models considered and which specification best fits them. Unit root and cointegration tests were carried out with the JMulTi software by Lütkepohl and Krätzig (2004).

2.4.2 Model Restrictions

As discussed, the structural models are all identified by means of restrictions on the long-run effects matrix. In particular, the type of restrictions on the SVAR models are all of the form as in (2.6). The restrictions on the SVEC models depend in part on the

number of cointegrating relationships. In this case both models have one such relationship and hence no short-run restrictions are required, since $r(r-1)/2 = 0$. Hence, the long-run restrictions as in (2.7) are used to identify the SVEC models. These identification methods provide three restrictions, as is necessary to just identify the models and are summarized as follows

$$\Psi = \begin{bmatrix} \star & 0 & 0 \\ \star & \star & 0 \\ \star & \star & \star \end{bmatrix} \quad \Xi B = \begin{bmatrix} \star & 0 & 0 \\ \star & \star & 0 \\ \star & \star & 0 \end{bmatrix}.$$

2.4.3 MS Model Specification

It is now necessary to specify the Markov switching (MS) models that are to be used. The number of lags and states can in principle be determined according to the model selection criteria developed by Psaradakis and Spagnolo (2006), who show that they work reasonably well. These criteria are the Akaike Information Criterion (AIC) and the Schwartz Criterion (SC). The AIC is calculated as $-2(\log\text{-likelihood} - n)$ and the SC is calculated as $-2\log\text{-likelihood} + \log(T)n$, where T is the sample size and n is the number of free parameters of the model.

In principle the log-likelihood increases with the number of states used, although at a diminishing rate. When using too many states, however, there are usually convergence and estimation problems and ultimately it is not possible to escape the problem of too few observations for a given state. Therefore, these model selection criteria can provide good judgement as to which model should be used since they penalize over-parameterized models. As already noted, these criteria can also help in selecting the number of model lags, however we prefer to choose model lag orders so as to avoid any residual autocorrelation. Hence, the optimal number of lags are determined by Portmanteau tests.

Table 2.3 shows results of the information criteria along with values of the log-likelihoods, $\ln(L)$ for all unrestricted models, i.e. models without any short or long-run restrictions. Minimum values of the information criteria are in bold. The maximum number of states considered is four. Beyond that no information criteria reaches a minimum and it becomes likely that there will be some states with very few observations. As noted, this causes convergence and estimation problems, often producing

Table 2.3: Information criteria of unrestricted models.

Model	States	AIC	SC	ln(L)
I: $y_t = [Y_t, r_t, s_t]'$	1	1845.537	-1738.716	952.768
	2	-1946.747	-1822.123	1008.374
	3	-1961.999	-1812.451	1023.000
	4	-1980.622	-1798.028	1041.311
II: $y_t = [IP_t, r_t, s_t]'$	1	-1537.906	-1367.364	816.953
	2	-1680.313	-1492.006	893.157
	3	-1727.339	-1514.161	923.669
	4	-1713.095	-1467.941	925.547
III: $y_t = [D_t, r_t, s_t]'$	1	-1629.014	-1490.600	853.507
	2	-1773.642	-1617.482	930.821
	3	-1791.649	-1610.646	946.825
	4	-1798.315	-1585.371	959.157
IV: $y_t = [E_t, r_t, s_t]'$	1	-659.727	-535.373	364.863
	2	-1220.152	-1085.140	648.076
	3	-1297.577	-1137.694	693.789
	4	-1327.046	-1135.186	717.523
V: $y_t = [E_t, D_t, s_t]'$	1	-2775.382	-2619.222	1431.691
	2	-3370.821	-3193.367	1735.410
	3	-3455.565	-3253.268	1784.783
	4	-3493.145	-3258.906	1812.572

meaningless results.

Both information criteria agree for models II and V. For all other models the AIC favors more states than the SC. Unfortunately, convergence problems sometimes occur, even when using 4-state models. For instance, model III in four states turns out to have two states with very few observations in them. With model V it is the case that the restricted 4-state model (i.e. the one with the long-run restrictions) fails to converge in the sense that the decomposition in (2.9) gives a singularity after a certain number of iterations. Given these considerations, two and three states are used for model III and three states are used for model V. For all other models the states suggested by the information criteria are used.

It is also worth noting that models with one state, or simply SVAR and SVEC models, are not supported by any criterion. In particular, for the SVEC models IV and

V the AIC and SC values are very high. Further, although not shown here, the log-likelihoods of models with a varying B matrix over states are only slightly higher than those with a state invariant B matrix; and the AIC and SC values are lower for a model with a state invariant B matrix as opposed to a varying one. This means that the assumption of a state invariant B matrix in (2.9) may well be justified, although this will formally be tested later on.

Portmanteau tests (not reported here) indicate that models I and III have no residual autocorrelation at 2 lags, models IV and V achieve this with 3 lags and model II needs 4 lags. Although, in fairness models IV and V still show signs of autocorrelation at the 10% level, which is also the case even when using higher lag orders. Therefore, in the interest of parsimony, for all models, we choose the lowest reasonable lag length possible.

2.4.4 Estimation results

The results of the MS models can most easily be presented according to number of states.

2-state models

Models I and III are best captured with two states according to the SC in Table 2.3. The most relevant parameter estimates, along with standard errors and the covariance matrices (scaled by 10^{-3}) of the 2-state unrestricted ⁴ models, are shown in Table 2.4. Both the values of the relative variance, $\lambda_{ij}, i = 2, \dots, M, j = 1, \dots, K$ parameters and the diagonal covariance matrix elements confirm that state 1 is the less volatile of the two states. This can also be seen from the smoothed probabilities of state 1 in Figure 2.2. In both models state 1 is never present during severe recessions of 2% or more economic contraction. For instance, the recession of 1958, the 1973-75 recession, both early 1980s recessions and the great recession of the late 2000s are captured by state 2.

Table 2.4: Parameter estimates, standard errors and covariance matrices (scaled by 10^{-3}) for 2-state unrestricted models.

Model	I: $y_t = [Y_t, r_t, s_t]'$		III: $y_t = [D_t, r_t, s_t]'$	
Parameter	estimate	σ	estimate	σ
λ_{21}	4.295	1.088	2.398	0.824
λ_{22}	2.523	0.706	10.004	2.531
λ_{23}	8.084	1.991	6.777	1.730
p_{11}	0.945	0.021	0.939	0.026
p_{22}	0.775	0.081	0.735	0.086
$\Sigma(1)$	$\begin{bmatrix} 0.045 & - & - \\ 0.893 & 278.148 & - \\ 0.033 & -0.082 & 2.288 \end{bmatrix}$		$\begin{bmatrix} 0.067 & - & - \\ -0.110 & 227.717 & - \\ 0.033 & -1.199 & 2.440 \end{bmatrix}$	
$\Sigma(2)$	$\begin{bmatrix} 0.195 & - & - \\ 1.819 & 2065.566 & - \\ 0.296 & -12.278 & 6.423 \end{bmatrix}$		$\begin{bmatrix} 0.559 & - & - \\ -7.257 & 1913.660 & - \\ -0.029 & -11.653 & 5.973 \end{bmatrix}$	

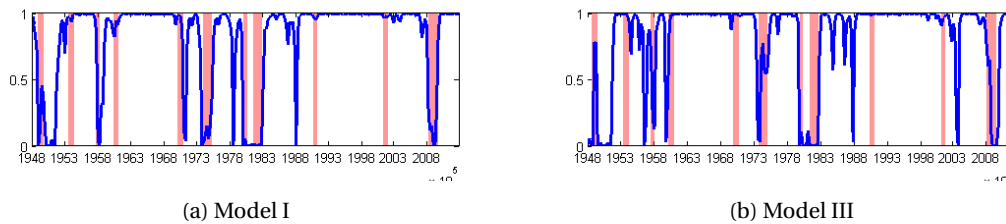


Figure 2.2: Smoothed probabilities of state 1 along with recession dates.

3-state models

A 3-state model is the most prolific one. In particular, given the convergence and interpretation issues mentioned above, three states are used instead of four for some of the models. Hence, models II, III, IV and V are considered in three MS volatility states. As with the 2-state models, the relevant parameter estimates along with standard errors are shown in Table 2.5 and the smoothed probabilities of the states are shown in Figure 2.3. Note that when using more than 2 states, all smoothed probabilities need

⁴Here unrestricted refers to no short or long-run restrictions on the state invariant B matrix.

Table 2.5: Parameter estimates, standard errors and covariance matrices (scaled by 10^{-3}) for 3-state unrestricted models. Tests for a state-invariant B matrix at the bottom.

Model	II: $y_t = [IP_t, r_t, s_t]'$		III: $y_t = [D_t, r_t, s_t]'$		IV: $y_t = [E_t, r_t, s_t]'$		V: $y_t = [E_t, D_t, s_t]'$	
Parameter	estimate	σ	estimate	σ	estimate	σ	estimate	σ
λ_{21}	10.706	3.328	1.326	0.446	5.702	1.251	3.570	0.725
λ_{22}	0.152	0.035	11.311	3.744	5.925	1.311	8.868	1.922
λ_{23}	0.460	0.112	0.458	0.172	2.215	0.518	0.700	0.149
λ_{31}	11.615	2.830	3.401	1.027	1050.264	620.878	464.556	157.217
λ_{32}	4.910	1.099	5.796	1.675	48.731	36.057	3.365	1.242
λ_{33}	1.361	0.322	9.632	2.682	1.778	1.150	3.609	1.381
p_{11}	0.981	0.054	0.923	0.029	0.952	0.040	0.972	0.016
p_{22}	0.906	0.033	0.950	0.034	0.870	0.046	0.921	0.038
p_{33}	0.888	0.045	0.670	0.136	0.831	0.371	0.818	0.268
$\Sigma(1)$	$\begin{bmatrix} 0.039 & - & - \\ 0.777 & 331.282 & - \\ -0.009 & 0.589 & 3.001 \end{bmatrix}$		$\begin{bmatrix} 0.069 & - & - \\ -0.322 & 199.302 & - \\ 0.010 & -0.751 & 2.203 \end{bmatrix}$		$\begin{bmatrix} 0.329 & - & - \\ -0.287 & 226.053 & - \\ 0.096 & -3.015 & 2.129 \end{bmatrix}$		$\begin{bmatrix} 0.279 & - & - \\ 0.018 & 0.046 & - \\ 0.098 & 0.023 & 2.803 \end{bmatrix}$	
$\Sigma(2)$	$\begin{bmatrix} 0.408 & - & - \\ 4.215 & 94.593 & - \\ 0.058 & 2.193 & 1.372 \end{bmatrix}$		$\begin{bmatrix} 0.038 & - & - \\ -3.771 & 2253.969 & - \\ 0.004 & -9.440 & 2.947 \end{bmatrix}$		$\begin{bmatrix} 1.872 & - & - \\ -1.483 & 1323.461 & - \\ 0.365 & -6.711 & 4.735 \end{bmatrix}$		$\begin{bmatrix} 1.100 & - & - \\ 0.147 & 0.396 & - \\ 0.130 & -0.014 & 1.937 \end{bmatrix}$	
$\Sigma(3)$	$\begin{bmatrix} 0.444 & - & - \\ 6.476 & 1625.760 & - \\ 0.029 & -13.648 & 4.293 \end{bmatrix}$		$\begin{bmatrix} 0.657 & - & - \\ -1.814 & 1155.005 & - \\ 0.200 & -4.549 & 7.563 \end{bmatrix}$		$\begin{bmatrix} 343.388 & - & - \\ 250.287 & 11024.389 & - \\ 45.811 & 29.018 & 9.863 \end{bmatrix}$		$\begin{bmatrix} 123.895 & - & - \\ 0.181 & 0.131 & - \\ 10.950 & 0.093 & 10.500 \end{bmatrix}$	
H_0 : state invariant B								
p -value	0.150		0.204		0.600		0.946	

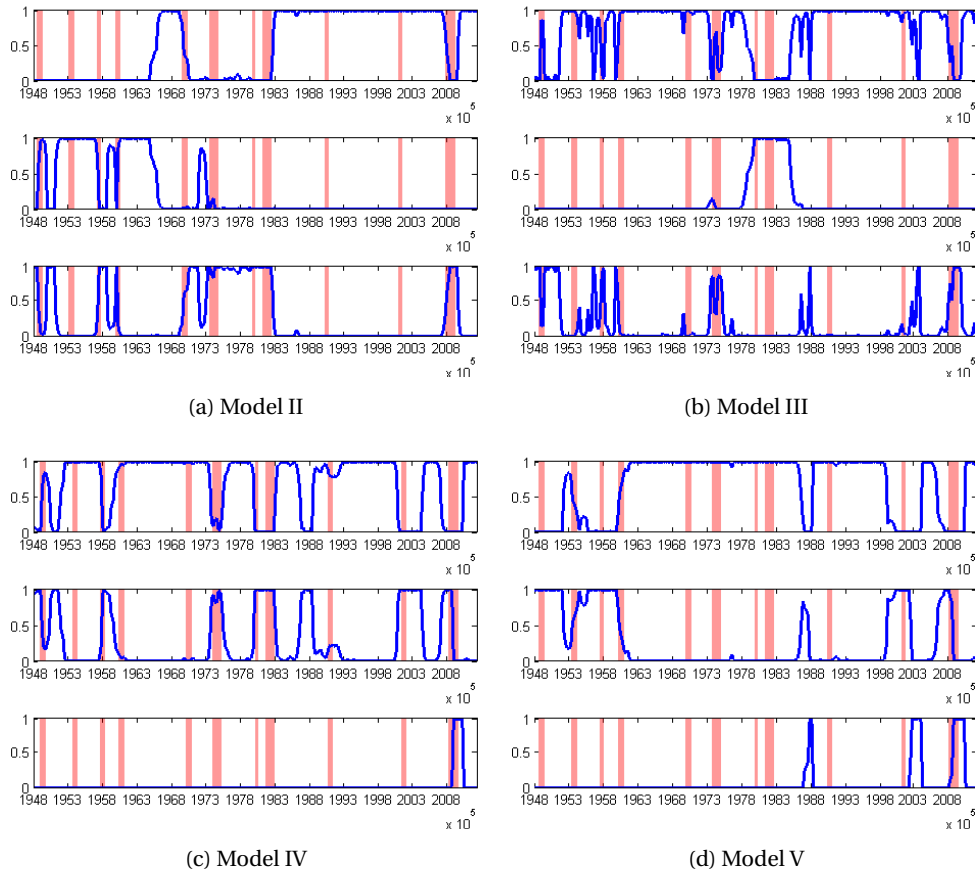


Figure 2.3: Smoothed probabilities of State 1 (top), State 2 (middle) and State 3 (bottom) along with recession dates.

to be displayed since it is no longer the case that one is the mirror image of the other.

Table 2.5 shows that the $\lambda_{ij}, i = 2, \dots, M, j = 1, \dots, K$ parameters still seem quite diverse, however in some cases their standard errors are also high. Also worth noting is that some of these parameters have rather low values, below 1, meaning that the relative variance in the given state is less than that of the first state. This can also be observed by the diagonal elements of the covariance matrices at the bottom part of the table. With the exception of model IV, the variances are not always increasing with a given state. This means that interpretation of the states is slightly more complex than with the 2-state models above.

From the smoothed probabilities in Figure 2.3 it can be seen that for all models,

state 1 is not usually associated with severe recessions. In particular, this is especially the case for model IV and, to a lesser extent for model V. State 2 largely tends to capture recession periods along with some time interval around them. For model III, state 2 is only associated with the early 1980s, where notably Binswanger (2004a,b,c), Groenewold (2004) and Jean and Eldomiaty (2010) all argue for the existence of a structural break in the relationship between stock prices and their fundamentals around that time period. However, no other model indicates any significant event around that time to warrant its own state. The great recession is always given by the third state. In model IV this is also the only occurrence of that state and in the other models state 3 is usually associated with severe recession periods. Hence, it can be interpreted as being the most volatile state.

The estimates of the transition probabilities in Table 2.5 are usually close to one, with the slight exception of p_{33} .⁵ This means that the states tend to be quite persistent as seen in Figure 2.3, in that the smoothed probabilities do not fluctuate often. The lower persistence of the third state is also something we would expect, since it is usually the case that crisis periods tend to be more transitory than economically stable periods. In this case the duration of the third state is roughly between 3 and 8 quarters, depending on the model used. This is a reasonable recession duration estimate, given the data span.

Finally, as noted in section 3, when using three or more Markov states, the assumption of a state invariant B matrix can be tested. The test distribution is asymptotically χ^2 with $(1/2)MK(K+1) - K^2 - (M-1)K$ degrees of freedom, or in this case 3. The p -values of such a test are shown in the bottom part of Table 2.5. Clearly, at conventional critical levels the null hypothesis of a state invariant B matrix cannot be rejected. Hence, one of the necessary model assumptions is justified by the data.

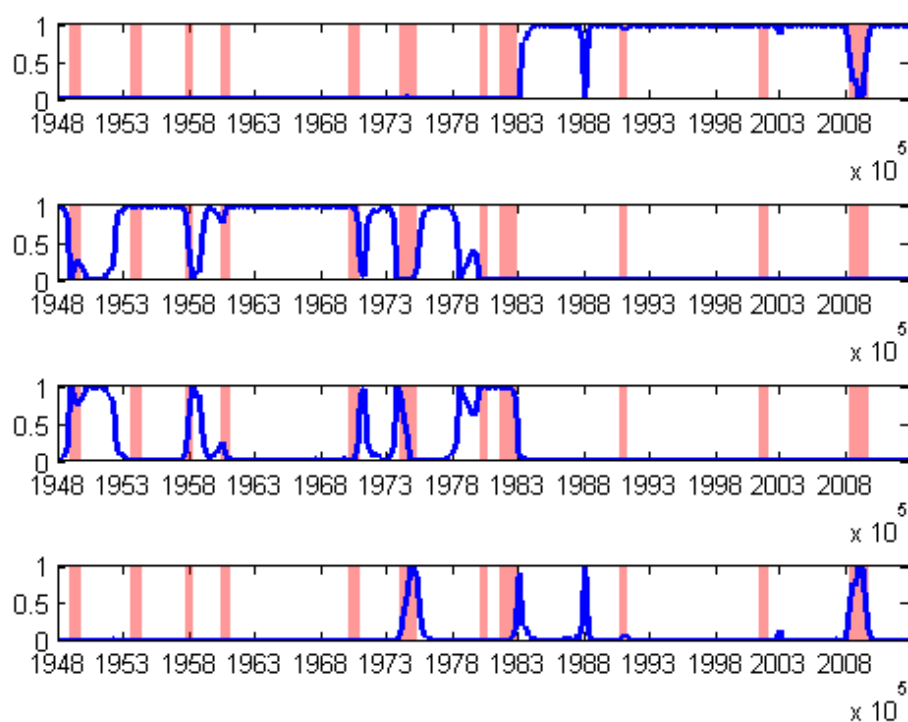
4-state models

Models I and IV are the only ones considered in 4 states. Results of their parameter estimates and smoothed probabilities are displayed in Table 2.6 and Figure 2.4 respectively. These models have the most parameters out of all the models considered

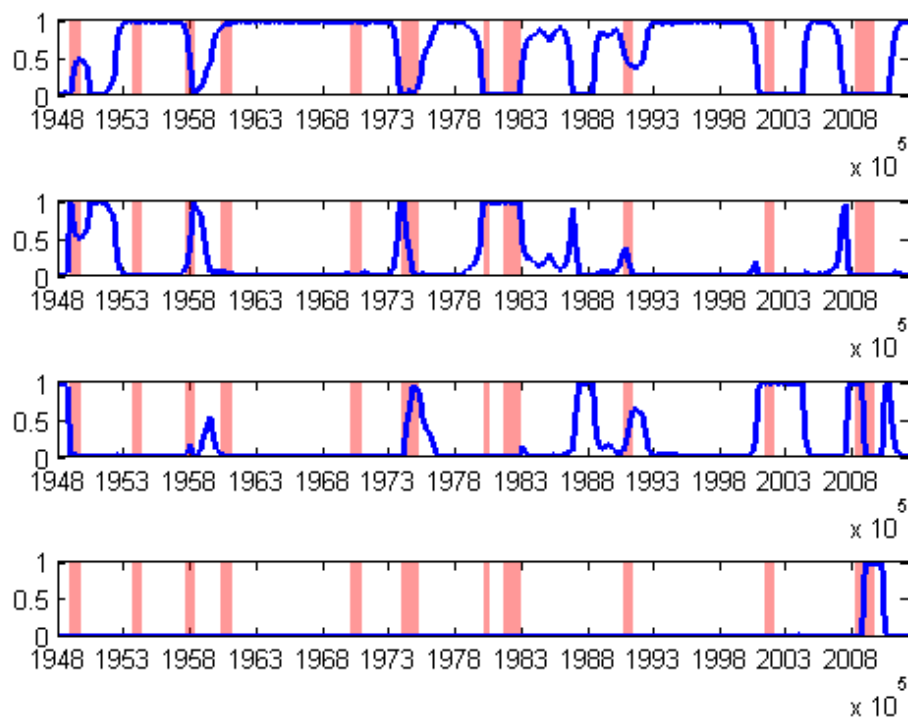
⁵These are not the only unrestricted elements of the transition probabilities, however to save space only these ones are displayed as they tend to be of most interest.

Table 2.6: Parameter estimates, standard errors and covariance matrices (scaled by 10^{-3}) for 4-state unrestricted models. Tests for a state-invariant B matrix at the bottom.

Model	I: $y_t = [Y_t, r_t, s_t]'$			IV: $y_t = [E_t, r_t, s_t]'$		
Parameter	estimate	σ	estimate	σ		
λ_{21}	3.734	1.051	1.168	0.463		
λ_{22}	0.385	0.092	9.553	2.488		
λ_{23}	0.946	0.411	1.116	0.370		
λ_{31}	2.465	1.049	9.706	2.714		
λ_{32}	0.397	0.196	1.250	0.399		
λ_{33}	5.685	1.835	2.776	0.798		
λ_{41}	9.883	3.521	1086.648	651.861		
λ_{42}	6.081	1.916	49.092	34.599		
λ_{43}	1.763	0.798	1.544	0.975		
p_{11}	0.931	0.097	0.948	0.027		
p_{22}	0.948	0.040	0.793	0.090		
p_{33}	0.668	0.169	0.816	0.068		
p_{44}	0.795	0.082	0.828	0.362		
$\Sigma(1)$	$\begin{bmatrix} 0.022 & - & - \\ 0.254 & 362.760 & - \\ 0.039 & -1.357 & 2.281 \end{bmatrix}$			$\begin{bmatrix} 0.305 & - & - \\ -0.171 & 217.329 & - \\ 0.117 & -1.879 & 2.164 \end{bmatrix}$		
$\Sigma(2)$	$\begin{bmatrix} 0.077 & - & - \\ 1.430 & 180.694 & - \\ 0.062 & -0.359 & 2.312 \end{bmatrix}$			$\begin{bmatrix} 0.358 & - & - \\ -1.881 & 2073.731 & - \\ 0.141 & -11.461 & 2.463 \end{bmatrix}$		
$\Sigma(3)$	$\begin{bmatrix} 0.214 & - & - \\ 3.287 & 2148.423 & - \\ 0.118 & -4.495 & 2.259 \end{bmatrix}$			$\begin{bmatrix} 2.938 & - & - \\ 0.236 & 272.111 & - \\ 0.580 & -3.471 & 6.030 \end{bmatrix}$		
$\Sigma(4)$	$\begin{bmatrix} 0.082 & - & - \\ 0.975 & 1261.377 & - \\ 0.265 & -7.034 & 17.680 \end{bmatrix}$			$\begin{bmatrix} 327.892 & - & - \\ 53.393 & 10667.553 & - \\ 40.652 & -47.875 & 8.602 \end{bmatrix}$		
H_0 : state invariant B						
p -value	0.375			0.082		



(a) Model I



(b) Model IV

Figure 2.4: Smoothed probabilities of state 1 to state 4 (from top to bottom) along with recession dates.

thus far. This consequently makes it more complicated to classify their states. The diagonal elements of the covariance matrices in the lower half of Table 2.6 do not always increase with the given state. Looking at the smoothed probabilities, it is however possible to classify state 1 as the least volatile state since it tends to avoid most recession periods. State 4 on the other hand tends to capture periods of severe recessions and for model IV only the great recession is present in that state. Hence, state 4 can be considered as the one with the highest volatility. States 2 and 3 are similar especially for model IV in that they are present during different recession periods.

It is again worth noting that the point of a structural break in the early 1980s can be somewhat justified when looking at the smoothed probabilities of state 1 for model I in Panel (a) of Figure 2.4. This state seems to be present mainly after the early 1980s, which could indicate a change in some fundamental relationship due to the lower volatility after that period. Also interesting is that both in 3 and 4 states, the most volatile state of model IV only captures the great recession and nothing else. This may not be that surprising however - when looking at the real earnings series in Panel (b) of Figure 2.1 a huge swing in real earnings is observed during the period of the financial crisis. Whether this particular period drives the results is investigated in the next section on model robustness.

Finally, as with the 3-state models, the assumption of a state invariant B matrix is formally tested. The resulting p -values are displayed at the bottom of Table 2.6. Note that the test statistic is asymptotically χ^2 distributed with 6 degrees of freedom. At the 5% critical level the hypothesis of a state invariant B matrix cannot be rejected. This means that for both three and four-state models, this assumption is supported by the data.

2.4.5 Testing the Identification Restrictions

We now turn to testing whether the restrictions in (2.6) and (2.7) are supported by the data. The first step is by testing whether the state invariant B matrix is identified through heteroskedasticity.

Table 2.7: Null hypotheses by state and p -values of Wald tests for all models.

States	Model	p -values		
2	H_0	$\lambda_{21} = \lambda_{22}$	$\lambda_{21} = \lambda_{23}$	$\lambda_{22} = \lambda_{23}$
	I: $[Y_t, r_t, s_t]'$	0.180	0.091	0.008
	III: $[D_t, r_t, s_t]'$	0.007	0.031	0.266
3	H_0	$\lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}$	$\lambda_{21} = \lambda_{23}, \lambda_{31} = \lambda_{33}$	$\lambda_{22} = \lambda_{23}, \lambda_{32} = \lambda_{33}$
	II: $[IP_t, r_t, s_t]'$	0.001	0.000	0.000
	III: $[D_t, r_t, s_t]'$	0.019	0.009	0.005
	IV: $[E_t, r_t, s_t]'$	0.263	0.014	0.016
	V: $[E_t, D_t, s_t]'$	0.001	0.000	0.000
4	H_0	$\lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}, \lambda_{41} = \lambda_{42}$	$\lambda_{21} = \lambda_{23}, \lambda_{31} = \lambda_{33}, \lambda_{41} = \lambda_{43}$	$\lambda_{22} = \lambda_{23}, \lambda_{32} = \lambda_{33}, \lambda_{42} = \lambda_{43}$
	I: $[Y_t, r_t, s_t]'$	0.001	0.000	0.001
	IV: $[E_t, r_t, s_t]'$	0.000	0.034	0.000

Testing for distinct lambda parameters

As discussed in section 3, in order for the B matrix to be identified through heteroskedasticity, it is necessary that all pairwise $\lambda_{ij}, i = 2, \dots, M, j = 1, \dots, K$ elements be distinct at least once in any $\Lambda_i, i = 2, \dots, M$ matrix. Since the standard errors are available, this is most easily tested by means of a Wald test. Likelihood ratio (LR) tests are also used, however, as already noted, such tests can suffer from convergence problems and in most cases converge to the same values for different hypotheses at least once. Nevertheless, when comparable, both tests yield the same results. The hypotheses and p -values of the Wald tests by models are given in Table 2.7.

The test statistic follows a χ^2 distribution with degrees of freedom equal to the number of joint hypotheses being examined. Taking a 5% or even a 10% critical level, no 2-state model can reject the null of at least one parameter pair of diagonal Λ_2 parameters being equal. For the 3-state models this is only true for model IV and all null hypotheses are rejected for the 4-state models.

It is promising that for all models the null hypotheses are rejected for a given number of states. In particular, the null hypotheses for models I, III and IV are all rejected when using a higher number of states. This means that for all models the B matrix is uniquely identified through heteroskedasticity. Hence, any restrictions now on that matrix become over-identifying and are in a position to be tested.

Table 2.8: p -values for LR tests of the long-run restrictions. The alternative hypothesis is a state invariant, unrestricted B matrix.

	model	H_0	LR test	p -value
2 states	I: $[Y_t, r_t, s_t]'$	(2.6)	0.698	0.874
	III: $[D_t, r_t, s_t]'$	(2.6)	8.735	0.033
3 states	II: $[IP_t, r_t, s_t]'$	(2.6)	56.084	4.031×10^{-12}
	III: $[D_t, r_t, s_t]'$	(2.6)	7.581	0.056
	IV: $[E_t, r_t, s_t]'$	(2.7)	19.764	1.900×10^{-4}
	V: $[E_t, D_t, s_t]'$	(2.7)	25.416	1.264×10^{-5}
4 states	I: $[Y_t, r_t, s_t]'$	(2.6)	6.601	0.086
	IV: $[E_t, r_t, s_t]'$	(2.7)	72.808	1.110×10^{-15}

Testing the restrictions

The restrictions to be tested are the lower triangular long-run identification restrictions mainly used in the literature, given by (2.6) and (2.7) for SVAR and SVEC models respectively. These restrictions are tested by comparing the log-likelihood values of the unrestricted (and identified) models with the restricted models according to (2.6) and (2.7) by means of an LR test. The results of these tests are given in Table 2.8. The distribution of the test statistic is asymptotically χ^2 with 3 degrees of freedom since all restricted models have 3 restrictions so that they are just-identified in the traditional sense. The alternative hypothesis is the model without any restrictions on the state invariant B matrix.

Starting with the 2-state models, the long-run restrictions for model I are accepted at the 10% critical value. However, that model did not have a uniquely identified B matrix as indicated in Table 2.7. Hence, any conclusions on the acceptance of the identification scheme become somewhat ambiguous. Moving on to the 3-state models, at a 5% critical value the long-run restrictions for model III are accepted. These restrictions are resoundingly rejected for all other 3-state models, given that their p -values are very close to zero. Finally, in four volatility states, the long-run restrictions for model I are again accepted, this time at the 5% critical level. Now however, the

B matrix is identified through heteroskedasticity, hence this result indeed shows support for the long-run identification scheme.

We therefore conclude that only models I and III in four and three states respectively have support from the data for the lower triangular long-run identification scheme. Such restrictions could indeed categorize shocks as fundamental and non-fundamental as the literature tends to do. With other models these restrictions do not seem to be warranted by the data, meaning that the identified shocks can probably not be interpreted as fundamental and non-fundamental.

Finally, it is worth mentioning that in most of the literature VAR models instead of VEC models are used. However, both cointegration tests indicate a strong presence of cointegration in models IV and V. Therefore, it is better to use the VEC form for such models. Note that a VAR in levels form is also possible, however this would again diverge from the literature, which mainly uses VARs in first differences.

2.5 Robustness Analysis

This section investigates whether the results obtained thus far rely to some extent on the exact model specifications used. Table 2.8 shows that the number of states do not seem to influence the final results. They only seem to matter for identifying the B matrix in (2.9) up to changes in sign and column ordering. A similar conclusion (although not reported here) can be drawn for the number of lags; though models with different lag orders may have residual autocorrelation as indicated by Portmanteau tests, the results in Table 2.8 stay similar depending on the critical level threshold chosen to evaluate them.

In order to investigate model robustness, it would be more relevant for example to try to determine whether the sample range could somehow drive the results obtained thus far. For instance, the smoothed probabilities of model IV in panels (c) and (b) of Figures 2.3 and 2.4 respectively show that there is one state that always captures the financial crisis. It would be interesting to investigate what would happen if the sample is cut to exclude the crisis years. Further, some papers for instance, Binswanger (2000), Binswanger (2004b) and Jean and Eldomiaty (2010) use data starting from 1953 to avoid having the Korean War influence their results. Hence, removing the tur-

bulent beginning and end of the sample would give a good indication of how robust the results are. Therefore, we only keep the observations from 1953:I - 2007:III for the robustness analysis.

Another factor potentially influencing the results could be the Markov switching (MS) specification itself. For example, as was already clarified, a MS model in heteroskedasticity needs to be used so that the B matrix can be identified and any restrictions on it can be tested. There is, however, little reason to assume that no other parameters can switch. When using data such as interest rates and stock price indices, it may well be the case that the intercept term is also subject to the same Markov switches as the covariance matrix. Indeed, stock prices tend to rise (fall) in periods of low (high) volatility. Allowing the intercept term to switch is another way of testing in how far the results obtained above are robust. Note that the autoregressive parameters could potentially also be switching, however the case for them to switch is harder to justify and to interpret. Further, switching autoregressive parameters may cause estimation issues; in that the number of parameters to be estimated increases and the data range may be too limited to give accurate estimates of all these parameters when using many MS states. Hence, we decide to investigate a model only with a further switching intercept term in addition to the switching covariance matrix. The VAR model, (2.1) then looks as follows

$$y_t = v(S_t) + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t, \quad (2.10)$$

where S_t follows a discrete valued first order Markov process as before and u_t still has the same distributional assumption as in (2.8). The reduced form VEC model is similar to (2.10) with the switching intercept being $v_0(S_t)$.

Finally, the Dow Jones index is not the only index followed by market participants. It consists of only 30 companies, whereas for example, the S&P 500 index consists of 500 companies as its name suggests. Even though these indices are closely correlated, the choice of index may influence the results and accept or reject some of the earlier conclusions. Further, data on the S&P 500 starts from 1957:I thereby giving a joint robustness check in terms of a different stock price index and data range compared with the original analysis.

It is worth noting that by reducing the sample range or introducing a new stock price index, the models need to again be tested for cointegration. It turns out that the

Table 2.9: p -values for tests of a state invariant B matrix.

H_0 : state invariant B					
	model	1953:I - 2007:III	intercept	S&P 500	Original
3 states	II: $[IP_t, r_t, s_t]'$	1.332×10^{-9}	0.138	0.419	0.150
	III: $[D_t, r_t, s_t]'$	0.191	0.106	0.051	0.204
	IV: $[E_t, r_t, s_t]'$	0.008	0.256	0.307	0.600
	V: $[E_t, D_t, s_t]'$	1.459×10^{-5}	0.901	0.006	0.946
4 states	I: $[Y_t, r_t, s_t]'$	3.195×10^{-8}	0.611	2.679×10^{-6}	0.375
	IV: $[E_t, r_t, s_t]'$	3.840×10^{-5}	0.481	0.999	0.082

cointegration relationships discovered earlier are all kept.⁶ Hence, models IV and V are still of the VEC form, with a cointegrating rank of 1, while models I - III are still of the VAR form. This is reassuring since cointegration is assumed to involve long-run relationships, which should not really change simply due to changes in the sample range used. Further, for all robustness specifications, the same lag lengths as in the original analysis are kept. This could in principle lead to some residual autocorrelation, however by using the same lag lengths the results from the robustness checks can best be compared to the original ones.

As in the original analysis, it is first necessary to confirm whether the assumption of a state invariant B matrix is justified. Recall, that this assumption can be tested for models with three or more Markov states. The test distribution is given as before; asymptotically χ^2 with $(1/2)MK(K+1) - K^2 - (M-1)K$ degrees of freedom. The p -values of such tests for all three robustness specifications are given in Table 2.9. The most right column of the table shows the original p -values from Tables 2.5 and 2.6 for comparison. It can be seen that the null hypothesis is usually accepted at the 5% significance level. A notable exception to this is for the shortened sample range. There, in most cases very low p -values are obtained, meaning a rejection of the assumption

⁶Only model IV with the 1953:1 - 2007:III sample range shows weak signs of cointegration, however, it is still present.

Table 2.10: p -values for LR tests of the long-run restrictions for different robustness specifications. The alternative hypothesis is a state invariant, unrestricted B matrix.

	model	H_0	1953:I - 2007:III	intercept	S&P 500	Original
2 states	I: $[Y_t, r_t, s_t]'$	(2.6)	0.911	0.644	0.699	0.874
	III: $[D_t, r_t, s_t]'$	(2.6)	0.287	0.032	0.888	0.033
3 states	II: $[IP_t, r_t, s_t]'$	(2.6)	$1.332 \times 10^{-9*}$	$2.854 \times 10^{-10*}$	0.038^*	$4.031 \times 10^{-12*}$
	III: $[D_t, r_t, s_t]'$	(2.6)	0.272^*	0.474	0.984^*	0.056^*
	IV: $[E_t, r_t, s_t]'$	(2.7)	0.277^*	2.718×10^{-4}	1.373×10^{-4}	1.900×10^{-4}
	V: $[E_t, D_t, s_t]'$	(2.7)	$1.325 \times 10^{-10*}$	$7.704 \times 10^{-13*}$	0.055^*	$1.264 \times 10^{-5*}$
4 states	I: $[Y_t, r_t, s_t]'$	(2.6)	0.451	0.023^*	0.026^*	0.081^*
	IV: $[E_t, r_t, s_t]'$	(2.7)	0.002^*	$9.525 \times 10^{-9*}$	$6.387 \times 10^{-7*}$	$1.110 \times 10^{-15*}$

* The B matrix is identified up to changes in sign.

of a state invariant B matrix. Overall however, from the results of the other robustness specifications, we can conclude that the assumption of a state invariant B matrix is a rather robust one.

Assuming that there indeed is enough justification for a state invariant B matrix, we then turn to test whether this matrix is identified through heteroskedasticity and if so whether the long-run restrictions in (2.6) and (2.7) are supported by the data. The p -values for tests of the long-run restrictions are reported in Table 2.10. Their test distributions are again asymptotically χ^2 with three degrees of freedom, since both identification schemes contain three restrictions. The most right column of the table again shows the original p -values from Table 2.8 for comparison. The stars in the table indicate when the B matrix is identified up to changes in sign and column ordering, i.e. when the null hypotheses in Table 2.7 are rejected.

Analyzing the results by model, it can be seen that there is no new conclusion for model I in two states; in all cases, the B matrix is not identified and the restricted model is accepted. When using a 1% critical value, the original conclusion also holds for that model in four states. Granted, the p -values are usually less than in the original

specification, however they are not arbitrarily close to zero, as is the case for some of the other models. Only for the short sample range is the B matrix not identified through heteroskedasticity, making any conclusion on accepting the identification scheme ambiguous.

Moving on to model II, its identification scheme is rejected over the robustness specifications, except when using the S&P 500 series and a 1% critical value. Overall, these results suggest that the identification restrictions for model II can largely be rejected. This means that when using industrial production data instead of GDP data the structural shocks may not be properly identified in the sense that the structural identifying restrictions are not supported by the data. This could be due to the stock price index and interest rates being more reflective of GDP rather than the industrial production index. It illustrates the need of being able to test a given identification scheme so as to let the data speak up about the restrictions. This shows that simply using the same identification scheme for different models, although closely related, does not necessarily lead to the identification of the same type of shocks.

The original conclusion for model III in two states does not really change since again the B matrix is not identified for any of the models. When using three states, the restricted model is accepted, as was originally the case. The p -values are in all cases higher than for the original model. However, the B matrix is not identified through heteroskedasticity when using a model with a switching intercept term. Nevertheless, it can be concluded with reasonable confidence that the identification scheme in (2.6) is robustly accepted for model III.

For model IV, the original conclusion is also accepted in all cases except when using a shorter sample range and three states. Then the identification restrictions in (2.7) are accepted. This could mainly be due to omission of the financial crisis period. As seen in Figure 2.1, panel (b), real earnings were severely affected during that time period. However, since the original conclusion rejecting the long-run identification scheme holds in almost all cases, it can be said to be quite robust.

Similarly for model V, the original conclusion is only rejected in one instance at the 5% critical level. Overall, the results in Table 2.10 lend some credibility to the original findings and show that they are rather robust over different model specifications.

To complete this analysis, a brief note on the smoothed probabilities is in order.

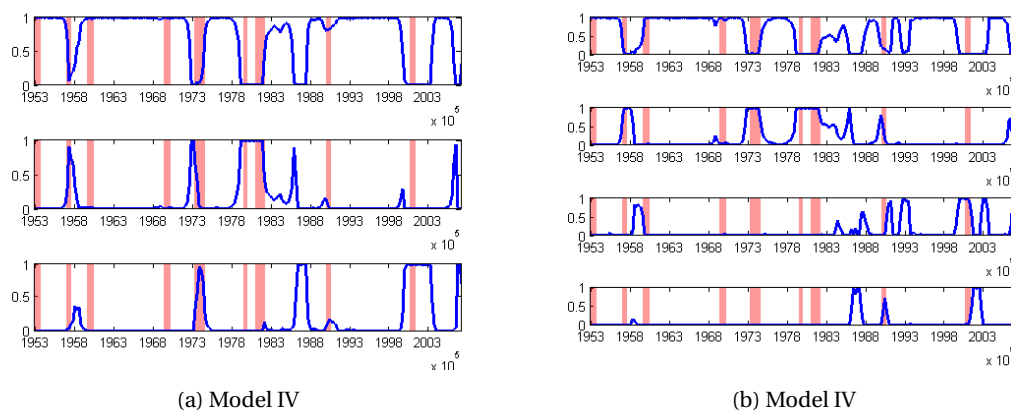


Figure 2.5: Smoothed probabilities of model IV with the 1953:I - 2007:III sample range with 3, (a) and 4, (b) states.

For the 2-state models, the periods depicted by the smoothed probabilities are very similar to the ones shown in Figure 2.2. In other words, they are not really different from the original ones.

The smoothed probabilities of the 3-state MS-SVAR models do not seem as robust to model specification however. In particular, for model II a similar picture as that in Panel (a) of Figure 2.3 is only obtained when using a model with a switching intercept term. Similarly, for model III, none of the robustness specifications show a unique event in the early 1980s as is the case in Panel (b) of Figure 2.3. Nevertheless, for all robustness models, the first state still depicts the most stable periods and recessions and crises are captured by states two and three as before.

The MS-SVEC models, on the other hand, tend to deliver rather robust smoothed probabilities over different robustness specifications when modeled in three Markov states. They largely resemble the ones in Panels (c) and (d) of Figure 2.3. Interestingly, when cutting the sample to before the financial crisis, state three no longer indicates a unique event for model IV. Rather recession periods are depicted by both states two and three as can be seen in Panel (a) of Figure 2.5. Model V with a switching intercept displays the financial crisis as a unique event in the third state.

A similar conclusion can be made for the 4-state models in the sense that the smoothed probabilities of the MS-SVEC models are more robust than those of the MS-SVAR models. It is also the case with these models that the first state depicts

the more stable economic periods, while other states capture more turbulent times. As in the 3-state case, when excluding the financial crisis, the fourth state no longer depicts a unique event for model IV, as shown in Figure 2.5, (b).

The robustness tests show that the results obtained earlier are not merely subject to chance and that there is some credible evidence either in favor or against the relevant identification scheme of a given model.

2.6 Conclusion

This analysis focuses on testing a commonly used structural parameter identification scheme that claims to identify fundamental and non-fundamental components of stock prices. In particular, five trivariate versions of the dividend discount model (DDM) are considered, which are widely used in the literature. The first variable of these models consists of different proxies of economic activity such as real GDP, the industrial production index, real dividends and real earnings; each proxy being a different model. All models are either specified in vector error correction (VEC) or in vector autoregressive (VAR) form. In this sense, the approach of this paper is similar to that of Binswanger (2004b) and Jean and Eldomiaty (2010), with the emphasis being on testing the structural parameter identification scheme. Restrictions are placed on the long-run effects matrix as in Blanchard and Quah (1989), making it lower triangular. All models are hence just-identified in the traditional sense.

A Markov switching in heteroskedasticity model as in Lanne et al. (2010) and Herwartz and Lütkepohl (2011) is used to test whether the long-run restrictions are supported by the data. It is found that for two of the models considered, the long-run identification scheme appropriately classifies shocks as being either fundamental or non-fundamental. Those are the models with real GDP and real dividends as proxies of real economic activity.

Three robustness tests are performed; one by cutting off volatile periods at the beginning and at the end of the sample. Another by allowing for a switching intercept term in addition to the switching covariance matrix; and a final robustness test uses S&P 500 data instead of DJIA 30 data. The robustness tests largely confirm the original findings.

Therefore, even though all the models are similar in the sense of being derived from the DDM, results of this paper suggest that simply using the same identification scheme for models with different variables may not be warranted by the data. Structural shocks may not be properly identified in this way, making any labeling of the shocks ambiguous. Hence, in order to ensure that economic shocks of interest are captured, it is good to test the relevant identification scheme using the MS in heteroskedasticity framework.

This paper therefore finds that models in which real GDP and real dividends are used as proxies of economic activity could potentially capture fundamental and non-fundamental shocks to stock prices. Since the findings in this paper are relatively robust, they serve as a good guideline when conducting future research in this field.

Chapter 3

Assessing the Sustainability of Government Debt - on the Different States of the Debt/GDP Process

Anton Velinov (EUI)

Abstract

This paper addresses the question of how sustainable a government's current debt path is. In particular, use is made of a Markov switching Augmented Dickey-Fuller (MS-ADF) model to determine the sustainability of public debt by testing whether a government's present value borrowing constraint holds. Building on the work of Raybaudi et al. (2004) and Chen (2011), the model in this paper is of a very general form. Using the data set from Reinhart and Rogoff (2011), it is possible to obtain long time series on debt/GDP for many different countries. In total 16 countries are investigated. Two different criteria are used to test the null hypothesis of a unit root in each state. The countries with a sustainable debt path are found to be Finland, Norway, Sweden, Switzerland and the UK. In contrast, the model indicates that France, Greece, Ireland and Japan have unsustainable debt trajectories. The remaining seven countries, (Argentina, Germany, Iceland, Italy, Portugal, Spain and the US) are all found to have uncertain debt paths. The model is robust to the sample size and number of states used. It is shown that this model is an improvement to existing models investigating this subject.

3.1 Introduction

The ramifications of the late 2000s financial crisis are ongoing. With the spectacular bankruptcy of Lehman Brothers in late 2008 and the subsequent stock market collapse, all signs were pointing to a severe recession, if not a depression. In order to deal with this unprecedented situation, governments around the world initiated stimulus packages to help kick-start the ailing economy. Alongside these measures, massive loan guarantees were made and financial institutions received large amounts of tax payer money in order to stay afloat.

The result of these market-intrusive measures left many governments (especially in developed countries) around the world straddled with high debt burdens. For instance, Ireland and even the UK, which until the mid 2000s were praised for their good budgetary housekeeping (see Afonso (2005)), saw their public debt burden skyrocket as financial institutions needed to be bailed out. For others this problem became so severe that the international community, acting through the IMF, had to step in so that contagion could be avoided. The most notable case being Greece.

This brings us to the topic of this paper: are current levels of public debt sustainable? This issue is by no means new. One of the first papers to analyze it is Hamilton and Flavin (1986). The basic idea of their paper is to set up a present value borrowing constraint (PVBC) for government spending and to test whether it is satisfied in the sense of a no-bubble condition. The test boils down to examining whether the debt and deficit series are stationary. This can be most easily accomplished by means of a unit root test as designed by Dickey and Fuller (1979) or Kwiatkowski et al. (1992) for example.

Regardless of the comprehensive amount of later literature on this issue, discussed in the next section, I argue that there is still unexplored potential. In particular, many of the earlier papers find evidence of structural breaks and nonlinearities in the debt process (such as Tanner and Liu (1994) and Quintos (1995)). This means that models capable of capturing such phenomena need to be considered. For instance, regime switching models. A popular method of obtaining endogenous regime shifts is through the use of the Markov switching (MS) technique. However, the MS models used to assess fiscal sustainability do so in a rather constricted setting. For instance, not all parameters in the Markov switching ADF (MS-ADF) model are allowed to switch, es-

pecially the variance parameter is held constant across states.¹ This may not be a correct assumption as the results of this paper show evidence of heteroskedasticity. Even for the parameters that are allowed to switch, there is no real reasoning provided as to why they should be switching. Another important limitation is that higher order autoregressions are left out,² which may make the conclusions of these studies doubtful as Kremers (1988) shows. In addition to these issues, only a handful of studies make use of long range data sets - for instance, starting earlier than the 1960s and including the latest financial and debt crises. Usually studies tend to avoid the turbulent war years.

My contribution is therefore, to analyze the issue of debt sustainability by means of a very general MS-ADF model for many different countries. Further, I determine the order of autoregression, the number of states, and which parameters should switch, based on Portmanteau tests, the information criteria in Psaradakis and Spagnolo (2006) and relevant diagnostic tests respectively. I also make use of a rich data set from Reinhart and Rogoff (2011), which includes many countries and usually a long time span (more than 100 years) of observations per country. Further, I bootstrap critical values to test the null hypothesis of a unit root in each regime. In this sense I largely expand on the work by Raybaudi et al. (2004) and Chen (2011).

The empirical results indicate that the countries with a sustainable debt path are Finland, Norway, Sweden, Switzerland and the UK. These countries either only have stationary states or their debt trajectory is currently in a stationary state. In contrast, it is found that France, Greece, Ireland and Japan have unsustainable debt trajectories. This is because their debt path is currently in an explosive state and for some countries all of the states are explosive. The remaining seven countries, (Argentina, Germany, Iceland, Italy, Portugal, Spain and the US) are all found to have uncertain debt trajectories. This is because their debt process is currently in a unit root state (with the exception of Argentina and Iceland); and in some cases both of their states are governed by a unit root process.

The next section discusses in detail the relevant literature on this topic. Section

¹For the cointegration test model, however, Gabriel and Sangduan (2011) do allow all three parameters of that model to switch.

²Both Raybaudi et al. (2004) and Chen (2011) drop the last lag term in the original ADF model by Hamilton and Flavin (1986).

3 presents the model and in section 4 the data and the countries investigated are described. Diagnostic tests and model selection are performed in section 5. The estimation and testing procedure is briefly explained in section 6 and section 7 presents the results of each model. Section 8 checks for robustness of the results and, finally, section 9 concludes.

3.2 Related Literature

To acquire an adequate overview of the extensive literature on this topic, I start with Hamilton and Flavin (1986) and work forward. Their paper uses annual US data from 1962 - 1984 on government debt and deficits and concludes, by means of an ADF test, that both series are stationary and, hence, the government is expected to balance its budget in the long-run.

Unfortunately, this conclusion is not universal. Two subsequent papers, by Kremers (1988) and Wilcox (1989), find that the US public debt series is non-stationary. They argue that Hamilton and Flavin (1986) did not specify their ADF model properly, in that higher order autocorrelation is not taken into account. A further paper by Trehan and Walsh (1991), however, accepts the original conclusions of Hamilton and Flavin (1986).

Subsequent papers, by Haug (1991) and Hakkio and Rush (1991), make use of cointegration tests to evaluate sustainability of government debt. Specifically, Hakkio and Rush (1991) argue that since government revenues and expenditures inclusive of interest payments are non-stationary, they must be cointegrated with a cointegration coefficient of around 1 for government spending to be sustainable. Their regressions find this coefficient to be below 1 in all cases and thus they conclude that the budget deficit is too large. However, Tanner and Liu (1994) conducting a very similar analysis, but including a structural break for 1981, reach an opposite conclusion. A later paper by Quintos (1995) sets out some conditions for deficit sustainability. In particular she shows that the coefficient of cointegration can be lower than 1 for the deficit to still be sustainable. Similarly, she also shows that cointegration between government revenues and expenditures inclusive of interest payments is only a sufficient condition for deficit sustainability. She finds a structural break in the 1980s after which the two

series are no longer cointegrated, however, she concludes that the deficit policy is still on a sustainable path. As can be seen, this alternative approach to testing whether the trajectory of public deficits is sustainable does not yield a universally held consensus either.

Yet another alternative approach is employed by Bohn (1998). He investigates the response of primary (non-interest) budget surpluses to changes in the debt-income ratio, claiming that a positive response provides reliable evidence for debt sustainability. In addition, he controls for wars and cyclical factors. He concludes that the current level of US debt is sustainable; although he does note that there can be some bad states of nature that can lead to excessive debt levels.

Later studies continue to investigate the issue of debt sustainability by means of the aforementioned stationarity and cointegration tests. They extend the analysis to other countries aside from the US and they continue to reach diverse conclusions. Granted, in most of the papers the data range and frequency differ. A good summary of much of the literature on this issue is provided in Table 1 of Afonso (2005) as well as in Tables 1 and 2 of Chen (2011). The former analyzes fiscal sustainability for 15 EU countries and concludes that most of them may not be on a balanced budget path; an ominous sign to the prelude of the financial and debt crisis.

Most recently, papers propose the use of regime switching models. This is a logical extension since many studies find evidence of structural breaks and also Bohn (1998) mentions that there could be different states of nature. In particular, Raybaudi et al. (2004) investigate debt for several different countries from the point of view of current account trade deficits. They use a MS-ADF type model in which one state is imposed to be non-stationary (i.e. unsustainable) and the other stationary (i.e. consistent with the PVBC). They claim that although one state would be associated with an untenable trade policy, the overall debt process may still be sustainable depending on the duration of the states and on the values of the parameter estimates. This approach is slightly generalized by Chen (2011), who does not impose a non-stationary state. Instead, it is left to the data to determine whether a state is stationary or not.³

³Related literature that uses regime switching models includes for example, Davig (2005) and Gabriel and Sangduan (2011). The former analyzes debt sustainability from a fiscal policy point of view as in Wilcox (1989). He uses a discounted debt series with a MS in intercept model. He distinguishes between a sustainable and an unsustainable state depending on whether the intercept parameter is significantly positive. The latter expand the cointe-

It is worth noting a strong critique to the whole literature on debt sustainability tests by Bohn (2007). He argues that stationarity and cointegration tests are irrelevant for assessing whether the PVBC holds. In fact, the PVBC would be satisfied after any finite number of differencing operations on the debt, revenues and interest inclusive expenditures series so that they are made stationary. Bohn only provides a mathematical intuition of this result. For instance, if a series is integrated of order m say, its n -period-ahead conditional expectation can at most be an m th-order polynomial of the n time horizon, while it will be discounted exponentially at a rate of n . He argues that since exponential growth dominates polynomial growth of any order, sustainability is still satisfied.⁴ This seems to invalidate stationarity and cointegration tests, however, they are still a sufficient condition for sustainability. Further, lenders could impose upper bounds on public debt, beyond which they would not be willing to lend so readily.

In light of this critique, I argue that the case for using a MS model is all the more potent. Such a model is able to provide information on what kind of states a country's debt process has experienced and what state it finds itself in at present. This way one could better judge whether public debt is currently on a sustainable path or not. This kind of model is not focused on a yes/no conclusion, rather it can paint a clearer picture of how the debt process is developing. The model is presented in the next section.

3.3 The model

The starting point of every analysis is the government's one-period borrowing constraint

$$G_t + (1 + r_t)B_{t-1} = R_t + B_t, \quad (3.1)$$

gration test for sustainability to include MS parameters. They subsequently classify different states as strongly and weakly sustainable depending on the coefficient of cointegration as in Quintos (1995).

⁴Intuitively, one could think of this as a country being on a seemingly highly unsustainable debt path and experiencing hyper inflation or severe exchange rate devaluation thereby making it substantially easier for it to repay its debt. Even though a country's debt-to-GDP series is currently non-stationary, there is nothing to say that at some point in the future this conclusion would not be reversed.

where G_t stands for government expenditures exclusive of interest payments, B_t is government debt, R_t government revenue and r_t can be either the real or nominal interest rate depending on how the other variables are measured (see Hakkio and Rush (1991)). In each subsequent period there will be a similar borrowing constraint, for $t+1, t+2, \dots$, etc., hence, the present value borrowing constraint (PVBC) is obtained by solving (3.1) forward:

$$B_t = \sum_{s=1}^{\infty} \prod_{j=1}^s \beta_{t+j} (R_{t+s} - G_{t+s}) + \lim_{s \rightarrow \infty} \prod_{j=1}^s \beta_{t+j} B_{t+s}, \quad (3.2)$$

where $\beta_t = 1/(1+r_t)$. For sustainability of the PVBC the last term needs to be zero, hence the following transversality condition needs to hold:

$$\lim_{s \rightarrow \infty} \prod_{j=1}^s \beta_{t+j} B_{t+s} = 0. \quad (3.3)$$

This implies that the present value of the government's debt is equal to the present value of its budget surpluses. Following Hakkio and Rush (1991), a slightly different formulation is used to derive testable implications. Assuming that interest rates are stationary with mean r , rB_{t-1} could be added and subtracted from both sides of (3.1) to obtain

$$E_t + (1+r)B_{t-1} = R_t + B_t, \quad (3.4)$$

where $E_t \equiv G_t + (r_t - r)B_{t-1}$. This formulation yields the following PVBC

$$B_{t-1} = \sum_{s=0}^{\infty} \beta^{s+1} (R_{t+s} - E_{t+s}) + \lim_{s \rightarrow \infty} \beta^{s+1} B_{t+s}, \quad (3.5)$$

where $\beta = 1/(1+r)$. Again for debt sustainability, the transversality condition needs to hold, in that the second term in (3.5) needs to be zero. If that is the case the term on the right hand side of (3.5) is expected to be stationary, which means that the left hand side, of the debt process also needs to be stationary. This is tested by means of stationarity tests on the first difference of the stock of public debt.⁵

⁵For completeness I also mention the cointegration test approach. In order to apply it, (3.5) needs to be rewritten as follows:

$$G_t + r_t B_{t-1} = R_t + \sum_{s=0}^{\infty} \beta^{s-1} (\Delta R_{t+s} - \Delta E_{t+s}) + \lim_{s \rightarrow \infty} \beta^{s+1} B_{t+s}.$$

Using the notation in Hakkio and Rush (1991), I define the left-hand side of the above equation as GG_t , hence $GG_t \equiv G_t + r_t B_{t-1}$. Meaning that the left-hand side includes government spending and interest payments on

Such stationarity tests can be extended to allow for different states of the public debt process. This implies that there may be stationary and non-stationary states of the path of government debt. Using a Markov switching (MS) framework, which captures switches in states endogenously, the following MS-ADF model is applied to test for unit roots

$$\Delta B_t = \nu(S_t) + \phi_1(S_t)B_{t-1} + \phi_2(S_t)\Delta B_{t-1} + \phi_3(S_t)\Delta B_{t-2} + \dots + \phi_{p+1}(S_t)\Delta B_{t-p+1} + u_t, \quad (3.6)$$

where the residual term u_t can also be subject to a regime switching variance. For estimation purposes this residual is also assumed to be normal, hence $u_t \sim \text{Nid}(0, \sigma^2(S_t))$.

6

It is assumed that S_t follows a first-order discrete valued Markov process with transition probabilities

$$p_{ij} = P(S_t = j | S_{t-1} = i),$$

which can be grouped in an $(M \times M)$ matrix of transition probabilities, P such that the rows add up to 1, and where M are the number of different states. The next section discusses the data to be analyzed with the MS-ADF model in (3.6).

3.4 The Data and Countries Investigated

This paper makes use of the extensive data set from Reinhart and Rogoff (2011), who provide an in-depth analysis of banking crises and public debt (defaults). The data consist of annual observations on the gross central (or when unavailable general)

debt. Again, assuming the absence of Ponzi games, the last term on the right-hand side of the equation needs to go to zero. To test whether this is the case, the estimate of the b parameter needs to be examined in the following regression

$$R_t = a + bGG_t + \epsilon_t.$$

Here ϵ_t is assumed to be stationary, while R_t and GG_t follow a unit root process. Hence, a sufficient condition for the above regression to be stationary is that R_t and GG_t are cointegrated and the estimate of b is close to 1, Bohn (2007) shows that this is not a necessary condition.

⁶An extension of the cointegration approach allowing for regime switches is given as in Gabriel and Sangduan (2011)

$$R_t = a(S_t) + b(S_t)GG_t + \epsilon_t,$$

where ϵ_t can share the same properties as u_t . This would imply that the cointegrating relationship is subject to regime changes.

Table 3.1: Countries and their data ranges.

Argentina ¹	1864 - 2011	Japan ²	1872 - 2011
Finland	1914 - 2011	Norway	1946 - 2011
France	1949 - 2011	Portugal	1851 - 2011
Germany	1951 - 2011	Spain	1850 - 2011
Greece	1950 - 2011	Sweden	1719 - 2011
Iceland	1923 - 2011	Switzerland	1929 - 2011
Ireland	1924 - 2011	UK	1692 - 2011
Italy	1861 - 2011	US	1790 - 2011

¹ One period is interpolated.

² Several periods are interpolated.

nominal government debt-to-GDP ratio. The data are until 2010 and I extend for an extra year of observations based on the original sources. For a detailed description of the data sources, see the data set accompanying Reinhart and Rogoff (2011). The countries investigated and their sample ranges are summarized in Table 3.1.

In this literature - starting with Hamilton and Flavin (1986) - it is common practice to make use of annual frequency data.⁷ This is due to the unavailability of higher frequency data for many countries prior to the 1990s, and, because of the slow changing nature of the debt/GDP process, higher frequency data is not a necessity. Unlike most of the literature, which typically uses data starting after WWII (and omitting the crisis of the late 2000s), the data series I use is far more prolific in observations. For instance, the longest series is that for UK, which is from 1692 - 2011. The second-longest is that for Sweden spanning from 1719 - 2011. This long sample range is particularly useful not only for improving estimation precision, but also since debt cycles can last for half a century or more as Reinhart and Rogoff (2011) point out.

In this study I do not make use of the full range of countries covered in Reinhart and Rogoff (2011). This is to avoid making the analysis convoluted, and also the existence of gaps in the data for some countries renders meaningful results improbable. I do however choose several countries as representative of a certain fiscal policy. For

⁷This can also be affirmed in Table 1 of Afonso (2005).

instance, Greece, Iceland, Ireland, Italy, Portugal and Spain are included as examples of budgetary-lax countries.⁸ I also include the so-called "safe haven" countries such as Germany, Switzerland, the UK, and the US, which have seen their long-term borrowing costs decrease sharply at the onset of the debt crisis. Another set of stable countries are the Nordic states, represented by Finland, Norway and Sweden; these are also small and open economies. The remaining countries investigated are Argentina, France and Japan.

3.5 Diagnostic Tests and Model Selection

Having determined the countries of investigation, the next step is to select an appropriate specification of the MS-ADF model, (3.6), for the data of each country. The model is repeated here for convenience

$$\Delta B_t = \nu(S_t) + \phi_1(S_t)B_{t-1} + \phi_2(S_t)\Delta B_{t-1} + \phi_3(S_t)\Delta B_{t-2} + \dots + \phi_{p+1}(S_t)\Delta B_{t-p+1} + u_t,$$

where $u_t \sim \text{Nid}(0, \sigma^2(S_t))$. The task is to determine the lag length, p , the number of states, M and which parameters should switch. Since this analysis tries to distinguish between periods of different fiscal regimes, (possibly a sustainable and an unsustainable one) the autoregressive parameters always need to switch.⁹ Diagnostic tests are conducted to determine whether the other parameters need to be switching.

3.5.1 Diagnostic Tests

Table 3.2 summarizes the results of some standard diagnostic tests per country data. Note that all diagnostic tests in this section are performed using the JMulTi software, developed by Lütkepohl and Krätzig (2004).

The lag lengths are chosen using the Akaike Information Criterion (AIC), the Final Prediction Error (FPE), the Hannan-Quinn Criterion (HQC) and the Schwarz Criterion (SC). These criteria do not always agree on the lag length of the model, which is why for some countries several lag lengths are reported.¹⁰

⁸Some of which are part of the recently coined term PIGS (Portugal, Italy, Greece and Spain).

⁹As will be discussed, this is supported by relevant model diagnostic tests.

¹⁰In a few cases a criterion (especially the AIC or the FPE) chooses a very high lag order, for instance, a 7-lag model for Japan. These high lag values are discarded since they do not seem very realistic when annual data is

Table 3.2: Diagnostic tests for all countries at various lag lengths

Country	Lag Length	Stationarity tests		Autocorrelation tests [†]				Heteroskedasticity tests [†]	
		ADF [*]	KPSS ^{**}	Q_{12} ¹	Q_{12}^A ²	LM_5 ³	LMF_5 ⁴	$ARCH_{LM}(12)$ ⁵	$ARCH_{LMF}(12)$ ⁶
Argentina	1	-3.70	0.78	0.28	0.25	0.22	0.23	0.00	0.00
	4	-3.27	0.45	0.28	0.23	0.86	0.87	0.00	0.00
Finland	1	-3.20	0.28	0.37	0.32	0.16	0.18	0.01	0.01
France	1	1.78	2.54	0.89	0.86	0.34	0.37	0.49	0.30
Germany	1	2.54	2.89	0.77	0.70	0.27	0.31	0.99	0.97
Greece	2	1.85	1.98	0.71	0.61	0.37	0.42	1.00	1.00
Iceland	1	0.69	2.49	0.94	0.92	0.75	0.78	0.96	0.94
Ireland	1	-0.20	1.79	0.19	0.15	0.11	0.11	0.55	0.41
Italy	1	-0.34	0.75	0.34	0.30	0.32	0.31	0.03	0.01
Japan	1	1.50	1.60	0.43	0.40	0.27	0.28	0.04	0.02
Norway	1	-1.50	0.67	0.53	0.43	0.58	0.62	0.77	0.64
Portugal	4	0.64	1.03	0.48	0.44	0.89	0.90	0.02	0.01
Spain	1	-1.08	4.82	0.68	0.65	0.42	0.43	0.00	0.00
	2	-1.17	3.25	0.75	0.71	0.73	0.74	0.00	0.00
Sweden	1	-1.63	2.87	0.01	0.01	0.02	0.02	0.00	0.00
	4	-1.40	1.20	0.16	0.14	0.29	0.30	0.00	0.00
Switzerland	1	-1.06	1.27	0.71	0.65	0.54	0.57	0.00	0.00
UK	1	-1.02	1.62	0.19	0.18	0.07	0.07	0.00	0.00
	2	-1.10	1.09	0.28	0.26	0.24	0.25	0.00	0.00
US	2	-0.08	4.20	0.69	0.66	0.62	0.64	0.00	0.00

[†] Only p -values are reported.

^{*} Critical values are -3.43 at 1%, -2.86 at 5% and -2.57 at 10%. Tests use an intercept term.

^{**} Critical values are 0.74 at 1%, 0.46 at 5% and 0.35 at 10%. Tests use an intercept term.

¹ Portmanteau test statistic using 12 lags with a χ^2 distribution.

² Adjusted Portmanteau test statistic using 12 lags with a χ^2 distribution.

³ LM test statistic using 5 lags with a χ^2 distribution.

⁴ LM test statistic using 5 lags with an F distribution.

⁵ ARCH-LM test statistic using 12 lags with a χ^2 distribution.

⁶ ARCH-LM test statistic using 12 lags with an F distribution.

As is common in most of the literature, I use ADF and KPSS unit root tests to determine stationarity of the data. From the relevant literature review discussed in section 2, it appears that a unit root or any non-stationary debt/GDP series is in violation of the PVBC, and hence, indicates an unsustainable public debt policy. From Table 3.2 it can be seen that for only 2 out of the 16 countries investigated, both the ADF and the KPSS test statistics indicate a stationary debt/GDP process at the 5% level.¹¹ Norway only has a stationary series according to the KPSS statistic; the ADF test still indicates non-stationarity. This inconclusiveness of the tests is a key motivation for both Raybaudi et al. (2004) and Chen (2011) to use a MS model. In particular, Chen (2011) argues that due to the nonlinear property of the time series involved, a conventional unit root test could have low statistical power. Afonso (2005) states that, in the presence of structural breaks, in particular the ADF test would be biased towards nonrejection of the unit root hypothesis. The reasoning in Raybaudi et al. (2004) is that non-stationarity due to large falls in the series (due to budget surpluses) is not an adverse event and therefore cannot mean that debt is on an unsustainable trajectory.

Further, as Bohn (2007) shows, stationarity is not a necessary condition for the PVBC to hold. What is required is that the series are difference-stationary of any arbitrary order, which is satisfied since all first difference series are stationary.¹² One may argue that a unit root process leads to an exploding debt-to-GDP ratio, which is clearly unrealistic. That said, factors such as (hyper)inflation or a currency devaluation can significantly reduce a government's debt burden without it having to default.

The next four columns of Table 3.2 present the p -values of residual Portmanteau and LM autocorrelation tests, both also with adjusted tests, more suited to small samples. One can notice that both statistics usually have similar p -values, unless short sample series are investigated. A high p -value cannot reject the null hypothesis of no residual autocorrelation.

The final two columns of Table 3.2 show the p -values of heteroskedasticity tests. In particular ARCH-LM tests with a χ^2 and an F distribution. For most countries concerned. Hence, the lag lengths presented in the table are the reasonable ones or the ones for which all criteria agree on, the latter is most often the case.

¹¹These are for Argentina with a 4-lag model and Finland.

¹²This is not really true for Greece since the ADF test still indicates nonstationarity, even though the KPSS test accepts the null of stationarity. Taking second differences makes both tests give a unanimous conclusion of stationarity.

with a short sample range it appears that there is no conditional heteroskedasticity present due to the high p -values. However, this conclusion may be unreliable; and, as discussed later, models that allow for a switching variance seem to convey more meaningful results than those that do not.

The next battery of diagnostic tests is preoccupied with model stability. In particular, Chow tests, recursive tests and cumulated sum of recursive residuals of squares (CUSUM-SQ) tests are used.

I use three types of Chow tests; the sample-split, break-point, and forecast test. In particular, the break-point test can provide (further) evidence of a structural break in the variance parameter, which is especially relevant for countries with positive p -values for the ARCH tests. These tests are carried out for a range of possible break dates and indicate that, for all country models, there is evidence of a structural break at certain time points in the relevant data range. This leads to a rejection of the null hypothesis of stable parameters. It lends further evidence in support of a non-constant variance over time, as most ARCH tests indicate.

The recursive tests for the AR coefficients and the intercept term reinforce the conclusion of breaks in the parameters. For some models, however, the intercept parameter could be stable according to this test. Also for a handful of data series the CUSUM-SQ test indicates potential parameter stability at the 5% level.

Overall, no single model is indicated as having stable parameters by all the tests. This lends support to using a switching parameter model and it means that the unit root test results may not be too reliable as Afonso (2005) and Chen (2011) point out. Exactly what kind of MS model is needed is discussed in the following.

3.5.2 Model Selection

The tests carried out above show that the data is subject to nonlinearities and structural breaks, hence a regime switching model, such as a Markov switching (MS) model is warranted. A MS model allows for parameters to switch endogenously, without having to impose a given break date. It is a very general model and therefore, can encompass other more restricted models. For instance, a smooth transition (ST) model is comparable to a MS model where one state is forced to be an absorbing state. As is shown later, if there is indeed an absorbing state, a MS model is able to capture it

without requiring the need of a ST model.

For univariate MS models, Psaradakis and Spagnolo (2006) develop several information criteria that can help decide simultaneously on the number of lags and states of the model. I use their AIC, BIC and HQC criteria. These criteria all give the same conclusions on the number of lags and states for all country models. They tend to opt for the most parsimonious configuration of 1 lag and 2 states. However, this configuration is not always optimal since Portmanteau tests sometimes show significant residual autocorrelation when using a single lag.¹³ Further, sometimes the smoothed probabilities of 3-state models capture more meaningful periods than those of 2-state models. Hence, when choosing the appropriate number of lags and states, I not only focus on the information criteria proposed by Psaradakis and Spagnolo (2006), but I also make sure that there is no residual autocorrelation and that the smoothed probabilities convey meaningful results.¹⁴

It is further necessary to decide which parameters need to be switching. As noted earlier, crucial for the analysis is a switching ϕ_1 coefficient. For higher order autoregressions, I let all such autoregressive coefficients switch. Parameter stability tests indicate that the variance is non-constant over time, which lends support to a switching variance. The intercept term is sometimes indicated as stable by some of the stability tests and therefore, I decide to usually keep it constant. A switching intercept term can also offer conflicting results. For instance, in many cases a switching intercept tends to capture periods with very high intercept levels *together* with very negative values of the ϕ_1 coefficient and visa versa; i.e. periods in which debt/GDP is very high (low), but supposedly very sustainable (unsustainable) as well. This is also found in Raybaudi et al. (2004) and in Chen (2011), but is not commented upon.¹⁵ Hence, I find it best to keep the intercept term non-switching in most cases, the only exception being the model for Germany.

Table 3.3 summarizes the models used for each country (based on (3.6)). The most general model syntax is a MS(M)-ADF(p)IAH model, where MS(M) stands for

¹³As Kremers (1988) points out, one should use a model without any residual autocorrelation.

¹⁴In section 8 on model robustness, two and three-state models are compared in more detail.

¹⁵It could be a consequence of allowing all the parameters to switch at the same time and may be averted if they can switch at different time points. This means that there would need to be more states in the model, which consequently may lead to inaccurate parameter estimates due to potentially very few observations per state.

Table 3.3: Model investigated by country.

Country	Model	Country	Model	Country	Model
Argentina	MS(2)-ADF(4)AH	Ireland	MS(3)-ADF(1)AH	Sweden	MS(2)-ADF(5)AH
Finland	MS(2)-ADF(2)AH	Italy	MS(3)-ADF(1)AH	Switzerland	MS(2)-ADF(1)AH
France	MS(2)-ADF(1)AH	Japan	MS(2)-ADF(1)AH	UK	MS(2)-ADF(2)AH
Germany	MS(2)-ADF(1)IAH	Norway	MS(2)-ADF(1)AH	US	MS(2)-ADF(2)AH
Greece	MS(2)-ADF(1)AH	Portugal	MS(2)-ADF(4)AH		
Iceland	MS(2)-ADF(1)AH	Spain	MS(3)-ADF(1)AH		

MS(M) stands for Markov switching with M states, ADF(p) for ADF model with p lags, I for a switching intercept term, A for switching autoregressive parameters and H for a switching variance parameter.

Markov switching with M states, ADF(p) for ADF model with p lags, I for a switching intercept term, A for switching autoregressive parameters and H for a switching variance parameter. Note that a single lag model in Table 3.3 is defined as $\Delta B_t = \nu(S_t) + \phi_1(S_t)B_{t-1} + u_t$, while a 2-lag model is $\Delta B_t = \nu(S_t) + \phi_1(S_t)B_{t-1} + \phi_2(S_t)\Delta B_{t-1} + u_t$, etc.

Ultimately, it turns out that the most popular model has only a single lag. This seems to vindicate earlier studies in which higher autoregressive orders are left out when using a MS model. However, a note of caution needs be made concerning this conclusion. Firstly, due to the small sample size of some (European) countries, for estimation purposes it is necessary to use a single lag model.¹⁶ Secondly, the US, which is one of the most studied countries in this literature, should have a model allowing for higher autoregressive orders. As Kremers (1988) shows, not accounting for the proper lag order of the ADF model could lead to residual autocorrelation and erroneous conclusions. Residual autocorrelation is present when using a single lag model for the US.

As a further generalization to the existing literature, I make use of more than two Markov states for some countries. Indeed, all the papers I am aware of that use MS

¹⁶This specification is always supported by at least one model selection criterion and Portmanteau tests indicate no residual autocorrelation present.

models to assess debt sustainability assume models with only two states. While two states can potentially distinguish between "stable" and "unstable" periods, a better picture can be obtained from a larger number of states. The reader may already anticipate that due to the limited number of observations for some countries it is not possible to use more than two states; just as it is not recommended to go beyond one lag order. This argument is correct, however, for countries with a longer data range, 3-state models not only give a higher log-likelihood, their smoothed probabilities and parameter estimates are sometimes more meaningful. No reasonable results are found for models with more than three states. The countries for which a model with three states is appropriate are Ireland, Italy and Spain. In section 8 on model robustness, it is shown that the 3-state model for these countries does indeed seem more reasonable in that the smoothed probabilities tend to be more meaningful in distinguishing between different historical time periods.

3.6 A Note on Estimation and Testing

All models in this paper are estimated in Matlab (R2011a) by means of the Expectation Maximization (EM) algorithm for univariate processes, as explained in Chapter 21 of Hamilton (1994). Standard deviations of the parameter estimates are obtained from the negative of the inverse of the Hessian matrix evaluated at the optimum.

The Markov switching ADF (MS-ADF) model given in (3.6) has a null hypothesis of $\phi_1(S_t) = 0$, for $S_t = 1, \dots, M$. This means that there is a unit root in each state according to the null. Also, unlike the conventional ADF test where the alternative hypothesis is a value of the test statistic lower than zero, in a MS framework there can be positive values of the test statistic in given states. This indicates the presence of an explosive process as argued in Hall et al. (1999).

In order to assess whether the estimated coefficients significantly differ from zero, Chen (2011) makes use of their standard deviations. He justifies this citing Gabriel et al. (2002), who come to the conclusion that testing for cointegration in a MS model can be accomplished by means of the standard errors. This approach seems to offer quite reasonable results and is used in this paper. However, as argued by Hall et al. (1999), the distribution of the test statistic under the null is unknown in a MS

framework. Hence, they parametrically bootstrap the model under the null to obtain critical values for hypothesis testing. They show through simulations that this is indeed a reliable approach. For completeness I make use of this approach as well. In particular I bootstrap the model in the vein of Psaradakis (1998).

It is usually the case that both standard deviations and bootstrapped critical values offer similar conclusions. Although, a note of caution is required when using the bootstrap technique. For instance, due to the highly nonlinear nature of the models, bootstrapping may not be a very accurate procedure. There are many local optima that the estimation could converge on. This may lead to a diverse range of critical values. Indeed, when running 2000 bootstrap replications for a given model several times, the critical values are found to diverge by a not too small amount in some cases. For more than two-state models this problem could potentially become even more severe. Of course, one could increase the number of bootstrap replications in the hope of alleviating this issue, however it is still not certain whether this would lead to an improvement in accuracy since the asymptotic properties of the bootstrap are not really known. Moreover, it would be a notoriously time consuming exercise. Therefore, it is advisable to not only rely on the bootstrapped critical values.

3.7 Empirical results

In the analysis that follows, I order the first state as being the one with the lowest value of the ϕ_1 parameter. In other words the states are classified as going from most "stationary" to least "stationary". Note that this in no way puts any restrictions on the parameter estimates since the states can be ordered in whichever way is desirable. Naturally, this is also done for the bootstrapped critical values.

For better clarity, I begin with the results of the 2-state models and subsequently present the results of the 3-state models. The final part of the section provides a summary of the results.

3.7.1 Results of the 2-state models

The most copious model is the one with two Markov states. This is always favored by the selection criteria in Psaradakis and Spagnolo (2006). It is the most parsimonious

configuration that can potentially distinguish between stationary and non-stationary periods.

The parameter estimates of these models are given in Table 3.4. So as not to take up too much space, the autoregressive parameter estimates for models with more than one lag order are not reported. Further, it is indicated which coefficients are stationary, $\hat{\phi}_1(m) < 0$, and which coefficients are explosive, $\hat{\phi}_1(m) > 0, m = 1, \dots, M$. The criteria used to test this are discussed in the previous section, namely the standard deviations and the bootstrapped critical values. Since standard deviations are reported, it is straightforward to see which criterion accepts/rejects the null hypothesis. Significance is concluded at the 10% level.

In most cases the estimate of $\phi_1(1)$ is negative, while that of $\phi_1(2)$ is positive. Usually this is indicated as significantly different from zero by at least one criterion. In some cases both criteria reach the same conclusion, which makes it easy to classify the given state. To obtain a better picture of each state it is further necessary to observe the smoothed probabilities; these are shown in Figure 3.1. Note that, for a 2-state model, the smoothed probabilities of state 2 are the mirror images of those of state 1.

Starting with the first country, Argentina, it can be seen that it has a stationary and an explosive state according to at least one test criterion. It is the country with the largest debt default in history. This happened a year after its economic collapse of 2001, at a time when its GDP had declined by 20% in four years. Argentina's economy has already recovered and its debt is now at around 40% of GDP. Regardless of this rather chaotic debt history, Argentina is one of the only countries in Table 3.2 to have a stationary debt process. This may not be very heartening for the investors who lost money during the Argentine default. It also illustrates the limitations of only looking for unit roots and concluding on debt sustainability. The MS model offers a way of obtaining information about the different states that the Argentine debt process finds itself in.

The smoothed probabilities in Panel (a) of Figure 3.1 reveal that the second state is associated with tumultuous periods in Argentina's history. The most notable being the severe stagflation of the 1980s and 1990s along with the debt default associated with the economic collapse of 2001. Clearly, state 2 follows a more volatile pattern

Table 3.4: Parameter estimates for 2-state models, standard deviations in parentheses.

	$\ln(\ell)$	\hat{p}_{11}	\hat{p}_{22}	$\hat{v}(1)$	$\hat{v}(2)$	$\hat{\phi}_1(1)$	$\hat{\phi}_1(2)$	$\hat{\sigma}^2(1)$	$\hat{\sigma}^2(2)$
Argentina	452.101	0.922 (0.027)	0.630 (0.114)	2.553 (0.502)	-	-0.137 (0.014) ^{¶¶}	0.214 (0.157) [*]	8.589 (1.204)	702.554 (206.193)
Finland	249.306	0.907 (0.050)	0.794 (0.125)	0.804 (0.279)	-	-0.073 (0.012) ^{¶¶}	-0.045 (0.043) [†]	1.748 (0.367)	115.680 (32.843)
France	144.261	0.792 (0.160)	0.717 (0.163)	-0.409 (0.596)	-	0.021 (0.015) ^{**¶}	0.056 (0.034) [*]	2.324 (1.254)	11.549 (4.245)
Germany	120.618	0.858 (0.095)	0.760 (0.138)	-0.502 (0.662)	0.389 (1.115)	0.017 (0.016)	0.062 (0.031) [¶]	1.801 (0.571)	3.159 (1.152)
Greece	162.045	0.953 (0.034)	0.876 (0.120)	0.639 (0.593)	-	0.003 (0.009) [*]	0.125 (0.024) ^{**¶}	6.819 (1.445)	28.626 (13.464)
Iceland	296.221	0.974 (0.019)	0.758 (0.291)	0.700 (0.757)	-	-0.024 (0.034)	0.178 (0.130) [¶]	12.146 (2.068)	279.742 (166.518)
Japan	432.681	0.916 (0.047)	0.951 (0.025)	-1.150 (0.536)	-	0.004 (0.024) ^{**¶}	0.091 (0.011) ^{**}	151.934 (33.827)	9.587 (1.560)
Norway	161.807	0.830 (0.089)	0.751 (0.161)	6.134 (0.905)	-	-0.268 (0.031) ^{‡¶}	-0.227 (0.049) [‡]	2.437 (0.759)	22.705 (7.776)
Portugal	411.614	0.861 (0.051)	0.638 (0.133)	0.409 (0.393)	-	-0.010 (0.012)	0.008 (0.025) [¶]	3.178 (0.607)	57.497 (14.422)
Sweden	607.011	0.934 (0.028)	0.905 (0.042)	0.235 (0.144)	-	-0.019 (0.008) [†]	-0.016 (0.011) ^{¶¶}	0.946 (0.145)	14.757 (2.067)
Switzerland	186.025	0.983 (0.017)	0.922 (0.076)	0.906 (0.480)	-	-0.059 (0.024) ^{¶¶}	-0.016 (0.039)	2.752 (0.524)	56.382 (22.438)
UK	916.884	0.917 (0.024)	0.857 (0.042)	0.857 (0.292)	-	-0.021 (0.004) [‡]	0.001 (0.007) [¶]	3.997 (0.619)	98.184 (14.889)
US	456.330	0.938 (0.025)	0.839 (0.061)	-0.067 (0.120)	-	-0.008 (0.005)	0.001 (0.015) [¶]	0.906 (0.162)	38.151 (7.669)

^{¶¶} The current state.

[†] Stationary according to one criterion.

[‡] Stationary according to both criteria.

^{*} Explosive according to one criterion.

^{**} Explosive according to both criteria.

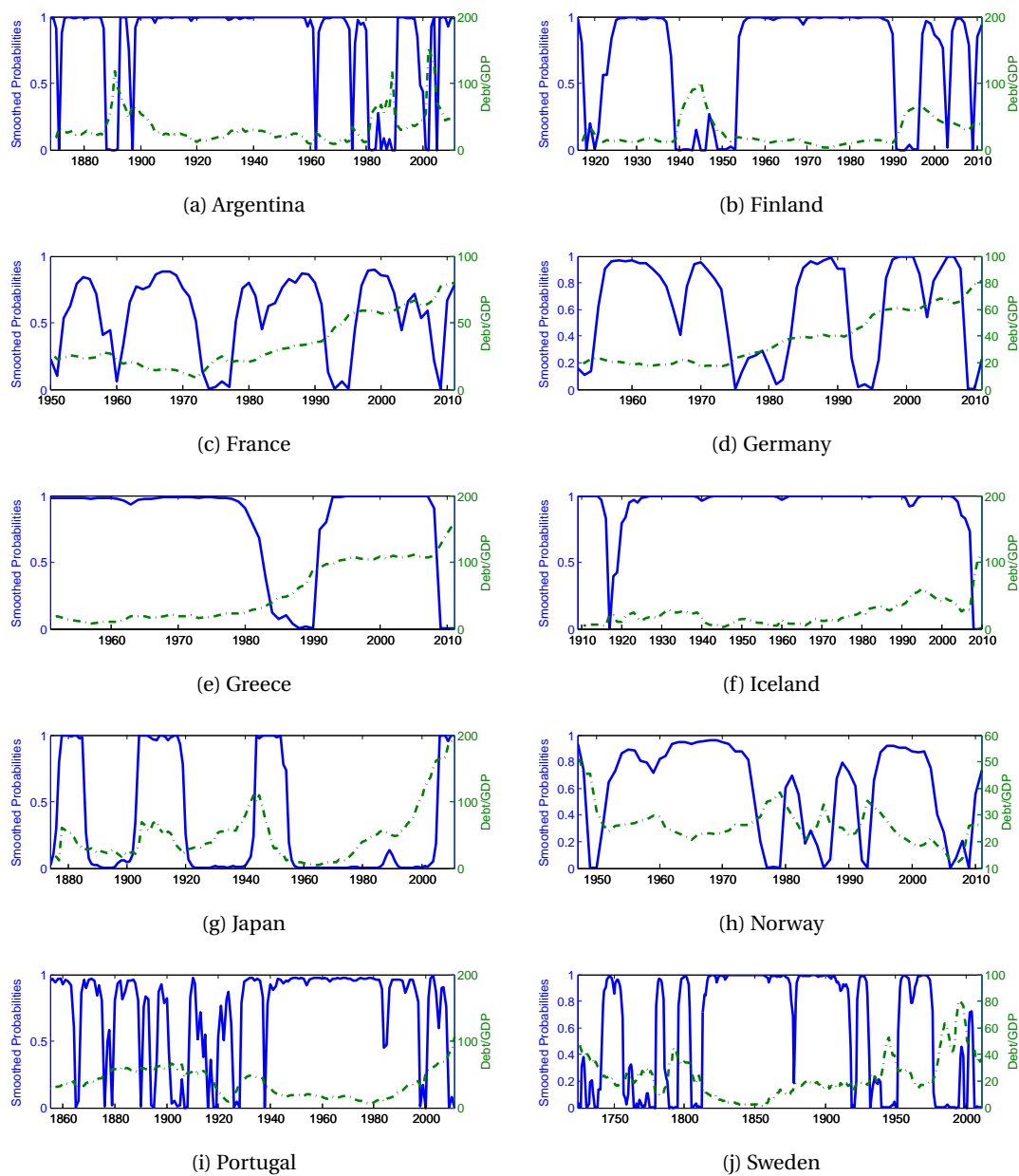


Figure 3.1: Smoothed probabilities of State 1, solid line (left axis) with the respective debt/GDP series, dashed line (right axis)

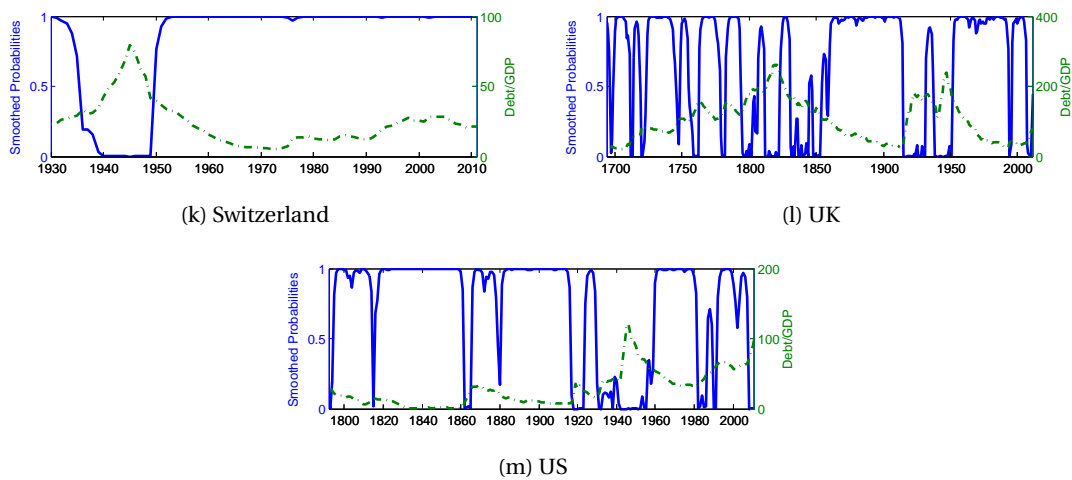


Figure 3.1: (continued) Smoothed probabilities of State 1, solid line (left axis) with the respective debt/GDP series, dashed line (right axis)

than state 1. This is reflected in the estimate of the variance of the second state, which is huge in comparison to that of state 1. Even the crisis of the late 2000s was not enough to plunge Argentina's debt into the more volatile state. In fact, the bootstrapped critical values (not reported here) indicate that the second state is highly explosive, with a coefficient of well beyond the 1% critical value.

From the point of view of an investor, it is good to know that the Argentine debt process currently finds itself in the less volatile, stationary state. However, its other state is explosive and associated with very extreme events, such as a severe stagflation and a debt default. It may still mean that, overall, the path of Argentine debt is uncertain. This is more than can be said by simply looking at the ADF and KPSS test values in Table 3.2. It is one of the merits of using a MS model.

Moving on to Finland, Table 3.4 shows that both coefficients of ϕ_1 are negative and significantly so according to at least one test criterion. From Panel (b) of Figure 3.1 it is clear that the first state depicts more stable periods. It is absent during both World Wars, the Finish depression of the early 1990s and the recent financial crisis. Currently, Finland's debt process is in the first state. With a debt-to-GDP ratio of below 50%, Finish debt seems to be on a sustainable path. It is also one of the few countries to have a stationary debt process according to both the ADF and KPSS tests

in Table 3.2.

France is the opposite of Finland, the estimated ϕ_1 coefficients are significantly positive according to at least one criterion, (the bootstrapped critical values). This means that France has two explosive or at least non-stationary unit root states. Its smoothed probabilities do not seem to convey too much additional information. The second state follows shortly after the 1973 oil crisis and its last occurrence is during the financial crisis. It seems that the second state captures periods of higher volatility, as can also be seen by the variance estimates. France is currently in the less volatile, though still non-stationary state. Such a state could be labeled as quasi-stable as in Chen (2011). However, these results seem to indicate that France's debt may not be on a sustainable path.

Germany is the country with the shortest data range, starting from 1951. Even though data starting from after WWII is standard in the literature, there is the possibility that very long-term debt cycles are not captured.¹⁷ This could potentially lead to inaccurate conclusions as is the case for all countries with short data ranges. This issue is investigated in detail in the next section on model robustness.

Although the parameter estimates of both ϕ_1 coefficients for Germany are positive, they are not significantly so according to both test criteria. Both states can hence best be characterized as having a unit root. The smoothed probabilities for Germany in Panel (d) of Figure 3.1 show that its debt process is currently in the second state. This is also the slightly more volatile of the two unit root states and has a positive intercept term. As already noted, a unit root does not necessarily mean that debt is headed toward unsustainable levels. The data show that until the mid 1970s, German debt-to-GDP was fluctuating at around the 20% level, since then it has been steadily on the rise with the latest estimate for 2011 putting it at above 80% of GDP. A unit root state means that the path of German debt is uncertain.

Greece also has both estimates of ϕ_1 above zero. However, unlike Germany, they are significantly positive. In fact, the estimate for the second state is positive according to both test criteria. Thus it can confidently be concluded that the second state is explosive. From the smoothed probabilities in Panel (e) of Figure 3.1, it can be seen

¹⁷Reinhart and Rogoff (2011) claim that these cycles can persist for half a century or more. In fact the UK debt/GDP process seems to have cycles lasting for more than a century.

that, unfortunately, Greece's debt is currently in the explosive, more volatile state. The first time its debt process was in that state was during the 1980s, which were characterized by a period of high inflation and weak economic growth.¹⁸ The first state, on the other hand, captures the Greek economic miracle from 1950 through the 1970s, as well as the stabilization and high growth that followed the turbulent 1980s. Since Greek debt is currently in the second explosive state, which is associated with adverse economic conditions, its current debt levels are probably unsustainable.

Similar to Greece, ever since the financial crisis, Iceland has been marred by debt problems. The ϕ_1 parameter estimates for Iceland are slightly negative and fairly positive for states 1 and 2, respectively. Since the estimate for state 1 is insignificantly different from zero, it could be characterized as a unit root state. State 2 seems to be explosive according to bootstrapped critical values. The smoothed probabilities in Panel (f) of Figure 3.1 show that state 2 is indeed associated with more turbulent times. The first occurrence of this state was for a brief spell during WWI, while the second one is at present. This also explains the much higher variance estimate of that state compared to the first state. Overall, the first state seems quasi-stable, as characterized in Chen (2011), while the second state seems to be one in which the debt path may be unsustainable. These results also raise the question as to what kind of government actions could throw a country into a situation that it only experienced once before; during the dark days of a world war.

Japan as well fares similarly to Greece. Its second state is characterized as explosive by both criteria - well beyond the 1% bootstrapped critical value. Its first state is also indicated as explosive according to bootstrapped critical values. Currently, Japanese debt is in state 1, although the smoothed probabilities show that this is a rather transitory state.¹⁹ It seems to capture periods during which debt is either increasing and decreasing or only decreasing. State 1 is present during the world wars and, hence, it is the more volatile of the two states according to variance estimates. What is remarkable is the rampant growth of Japanese debt since the Lost Decade of the 1990s. It currently stands at above 200% of GDP. Since Japanese debt only seems to have explosive states - one of which is explosive according to both criteria and per-

¹⁸Inflation throughout the 1980s was at an average of 19% while average GDP growth rate was 0.7%.

¹⁹State persistence is given by $1/(1 - p_{mm})$, for $m = 1, \dots, M$, in this case the first state is said to persist for about 12 years.

sists for long time periods - it is not likely that its debt path is sustainable. Again there is a curiosity as to how long investors will continue to be attracted to the ever growing debt of a country with little to negative economic growth.

Both testing criteria indicate that both states of Norwegian debt are significantly stationary. The second state seems to capture more volatile periods as indicated by the variance estimates. Norwegian debt is now in its less volatile state and can accordingly be classified as being on a sustainable path. Norway is a large oil and natural gas exporter, which is clearly helpful in creating current account surpluses to stabilize excess deficits.

Both states of Portuguese debt can be characterized as having a unit root since their ϕ_1 parameter estimates are very close to zero. Panel (i) of Figure 3.1 shows that the states have changed in a rather erratic way during the years up to and including WWI. Clearly, as the variance estimates also show, state 1 is associated with less volatile periods, for instance the high growth period of 1950 - 1973. Note that the military coup of 1974 is in this state as well, indicating that this event did not destabilize government debt too much. Currently, Portugal's debt is in state 2, the unit root state with the higher variance. A unit root state means that it is uncertain as to whether debt levels are likely to be sustainable.

Sweden has the second longest sample after that of the UK, beginning in 1719, shortly before the end of the Great Northern War (1700 - 1721). Since that war there were numerous conflicts with Russia, which occurred until the early 1800s. This can be observed in the debt-GDP series, the dashed line in Panel (j) of Figure 3.1. Even these times of conflict did not push Swedish debt to much above 40% of its GDP. The 19th century brought about a period of industrialization and modernization to Sweden. This kept its debt in relation to GDP at stable and low levels. After that only WWII and the crisis of the 1990s seem to have caused large rises in debt. Government debt reached its highest level of almost 80% of GDP in 1994. From 1998 onwards the Swedish government has run budget surpluses, except for 2003 and 2004, and has reduced its debt burden to around 35% of GDP. Both states of Swedish debt are therefore, indicated as stationary according to one of the testing criteria.²⁰ The first state

²⁰Even though parameter estimates of the ϕ_1 coefficients are close to zero, the long data range leads to more precise estimation results thereby making it easier to determine significance.

captures less volatile periods than the second state - as can be seen by the variance parameter estimates and the model smoothed probabilities. Currently, the Swedish debt process is in the second state, the more volatile stationary state. This state is also associated with large downward movements in the debt series, which could explain why Swedish debt is currently in it. Overall, this analysis suggests that Sweden has a sustainable government debt.

Switzerland appears to have a stationary and a unit root state, where the stationary state is the less volatile one. The smoothed probabilities clearly depict state 2 as being the WWII state, in which Swiss debt spiked to close to 80% of GDP in 1945. All remaining time periods have been in state 1. Accordingly, state 1 can be labeled as stationary, or at least quasi-stable as in Chen (2011), which means that Swiss debt is currently on a sustainable path.

The UK is the country with the longest data range, with the first observation starting in 1692.²¹ The series is depicted in Figure 3.1, (l) by the dashed line. Ever since it began, it has seen a persistent growth throughout the 18th century, the result of numerous conflicts such as the American War of Independence and the Napoleonic Wars. After the Battle of Waterloo (1815) this figure reached to more than 200% of GDP. Since that spike government debt gradually fell over the years to a mere 25% of GDP just prior to the outbreak of WWI.²² The above-mentioned pattern repeated itself in the years following WWI. Government debt soared again to over 200% of GDP by the end of WWII, only to "slowly" drop to a low of 25 % of GDP in 1992. It is currently estimated to be at about 80% in 2011.²³ This illustrates the claim made by Reinhart and Rogoff (2011) that government debt can be subject to cycles persisting for half a century or more.

The relevant coefficient estimates for the UK are both close to zero with small

²¹This was before the establishment of the Bank of England, (1694) and before the treaty of Great Britain, (1707) that resulted in the political union of the Kingdom of England and the Kingdom of Scotland.

²²This is a good example of the critique by Bohn (2007). If we were to observe the UK debt-to-GDP series from its initial period until the early 1800s we would conclude that the PVBC is violated, since there seems to be a unit root in the debt trajectory - this is indeed confirmed by the ADF and KPSS test, the latter giving a test statistic value of 5.57, far above the 1% critical value of 0.74. However, this conclusion would in the end be an erroneous one, since in the subsequent years that followed, the UK has managed to reduce its debt burden. Hence, a unit root in the debt/GDP process does not necessarily mean that government debt is on an unsustainable path.

²³A good summary on the UK public debt series can be found at http://www.ukpublicspending.co.uk/debt_brief.php

standard errors - the large number of observations improves estimation precision. State 1 can be characterized as stationary by both criteria, while state 2 is a unit root state associated with higher volatility. Judging by the smoothed probabilities, state 1 is more often associated with periods in which the debt/GDP ratio has been declining. State 2 tends to capture many of the war years and other such turbulent times. At present it appears that the UK's debt is in such a situation as characterized by the second state.²⁴ Since this is a unit root state, it is not certain how sustainable the UK's debt really is. Fortunately, the other state is stationary, which bodes well for British debt.

The last country with a 2-state model is the US. It too has a long data series, starting from 1790, slightly after the American War of Independence. Although comparable in length, its series is by no means as colorful as that of the UK. From the dashed line in Figure 3.1, (m) one can notice that most of the time American debt/GDP has been very low. This is with the exception of some major events, such as the War of 1812, the American Civil War of the 1860s, the World Wars and the Great Depression, where government debt has seen significant rises with respect to GDP. Note that debt/GDP was declining throughout the whole period of the Vietnam War. Both ϕ_1 parameter estimates are very close to zero and indeed indicate that both states are unit root states. State 2 has a higher variance than state 1. From the smoothed probabilities it can be seen that state 2 captures the more politically and economically unstable periods already mentioned. Currently, American debt is in the more volatile unit root state making the sustainability of its debt path uncertain. One must also keep in mind that American debt/GDP is close to its record highs.

3.7.2 Results of the 3-state models

All 3-state models for the three countries investigated are the same, namely the MS(3)-ADF(1)AH model. The single lag implies that all of the parameter estimates for each country are reported in Table 3.5.²⁵ The relevant smoothed probabilities are depicted in Figure 3.2. Note that, unlike with the 2-state models, no probabilities can be mirror images of each other and hence, all are shown.

²⁴The current smoothed probability for this state is slightly above 50%.

²⁵The only exception is three unrestricted transition probabilities parameters.

Table 3.5: Parameter estimates for 3-state models, standard deviations in parentheses.

Parameter	Ireland	Italy	Spain
$\ln(\ell)$	253.358	484.514	484.160
\hat{p}_{11}	0.902 (0.037)	0.929 (0.206)	0.870 (0.057)
\hat{p}_{22}	0.922 (0.050)	0.969 (0.221)	0.887 (0.303)
\hat{p}_{33}	0.709 (0.174)	0.976 (0.011)	0.882 (0.054)
\hat{v}	1.210 (1.024)	2.028 (0.700)	1.586 (0.595)
$\hat{\phi}_1(1)$	-0.086 (0.018) [†]	-0.043 (0.012) [†]	-0.103 (0.024) [†]
$\hat{\phi}_1(2)$	0.017 (0.016)	-0.037 (0.026)	-0.013 (0.027)
$\hat{\phi}_1(3)$	0.188 (0.075) ^{**¶}	-0.008 (0.011) [¶]	-0.010 (0.024) [¶]
$\hat{\sigma}^2(1)$	8.921 (2.266)	3.198 (1.177)	4.480 (2.723)
$\hat{\sigma}^2(2)$	12.798 (3.996)	242.124 (59.226)	255.035 (102.623)
$\hat{\sigma}^2(3)$	36.580 (16.853)	29.728 (4.770)	16.574 (3.869)

[¶] The current state.

[†] Stationary according to one criterion.

^{**} Explosive according to both criteria.

Starting with Ireland, from the estimated ϕ_1 coefficients, it seems that state 1 is stationary, state 2 follows a unit root, while state 3 is clearly explosive as indicated by both testing criteria.²⁶ The smoothed probabilities show that Irish debt is currently in the third state, which is explosive and is also the most volatile state. The first state captures periods in which debt/GDP is stable or slightly declining, as can be seen by the dashed line representing the series. The second state shows periods in which debt/GDP is moderately increasing. Since Irish debt is currently in the explosive state, it is not likely to be sustainable.

For Italy, all three values of the ϕ_1 coefficients are negative. The one for state 1 significantly so according to the standard deviation, while the other two states seem to be governed by a unit root process. The smoothed probabilities in Panel (b) of Figure 3.2 show that the first state is associated with stable and high-growth periods.

²⁶A word of caution needs to be noted about the bootstrapped critical values; since these are 3-state models, the bootstrapped estimates may converge to many different local optima thereby giving a very diverse range of critical values. This could mean that they may potentially be insensitive to detect coefficient significance.

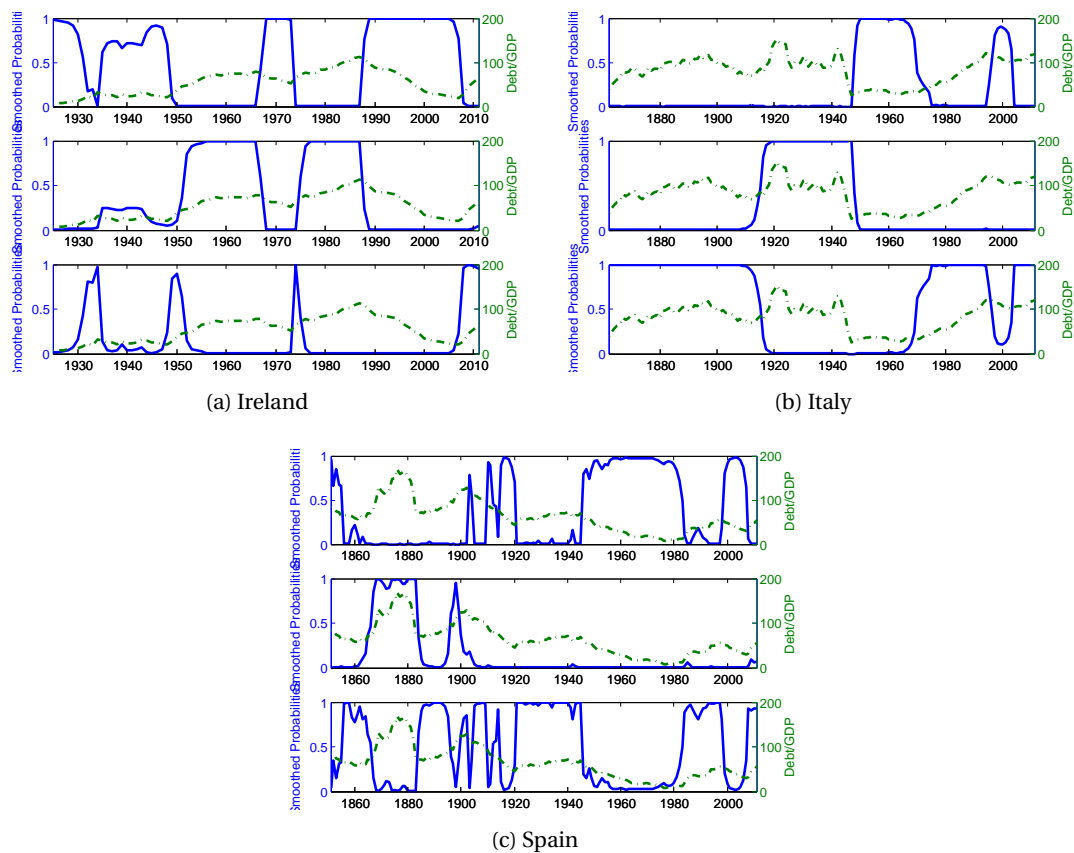


Figure 3.2: Smoothed probabilities of all State 1 (top), State 2 (middle) and State 3 (bottom) solid lines (left axis) with the respective debt/GDP series, dashed lines (right axis)

From 1951 until 1973 the Italian economy grew at a rate slightly higher than 5% per annum on average. In fact this was one of the highest growth rates among European countries.²⁷ The second occurrence of state 1 is from the mid 1990s until before the financial crisis. Although the economy did not grow at such vigorous rates during that period, debt/GDP was apparently stable. The second state captures the period from WWI up to and including WWII. Clearly an exceptional time frame, in which the debt-income ratio experienced erratic shifts, as shown by the high value of $\hat{\sigma}^2(2)$. Currently, Italian debt is in the third state, the less volatile of the two unit root states. It depicts the period of political instability from the 1970s and the economic recession of that time. Also, it is a period in which debt started to increase to reach 120% of GDP

²⁷This has been documented in Crafts and Toniolo (1996) and in Di Nolfo (1992).

by the mid 1990s. It is unclear whether this state of Italian debt could be sustainable and hence, its debt path is at best characterized as uncertain.

The final country investigated in this category is Spain. It has very similar results as Italy in that the first state seems stationary, while the other two follow a unit root process. The smoothed probabilities in Panel (c) of Figure 3.2 show that state 1 captures the economic stabilization of the mid 1950s and the "Spanish miracle" of 1959 to 1975. It ends prior to the early 1990s recession and is present again during the Spanish property boom of 1997 - 2007. Clearly, state 1 depicts the high-growth, rather stable periods in Spain's history. State 2, which is also the most volatile state, captures some earlier periods in which debt/GDP has had some drastic shifts. Spanish debt is currently in the third state, which has been quite prevalent in its history. Though since it is a unit root state - as is the case with Italy - it is uncertain whether Spanish debt is sustainable at the moment.

3.7.3 Summary

Due to the many country models investigated, it is convenient to have a short summary of the above results. In particular, Raybaudi et al. (2004) and Chen (2011) in their analyses took notice of the duration of the states, $(1/(1 - p_{mm}))$, for $m = 1, \dots, M$ and the current state the debt process finds itself in. This is also a good way to summarize the results obtained here and is presented in Table 3.6.

Naturally, there are some limitations to the general model used in this paper. Most prominent would be the comparison of states that are characterized in the same way, across countries. For instance, an explosive state could be more or less severe for one country than for another country. The same holds true for a unit root state. A further issue, as discussed before, is the unavailability of long data series for some countries. This is a problem of empirical work in general. Shorter time series may mean that very long term debt cycles as noted in Reinhart and Rogoff (2011) are not captured. Therefore, the conclusions in the last column of Table 3.6 need not be interpreted too literally, though they can provide a good guidance.

From the table one can see that the debt path of only few countries is characterized as sustainable. This is because it is preferred to err on the side of caution in concluding a sustainable debt path. Such a conclusion is usually reached if the

Table 3.6: Durations of regimes and current state of regime by country.

Country	$(1 - p_{11})^{-1}$	$(1 - p_{22})^{-1}$	$(1 - p_{33})^{-1}$	Current state	Debt path
Argentina	12.8	2.7	-	Stationary	Uncertain
Finland	10.8	4.9	-	Stationary	Sustainable
France	4.8	3.5	-	Explosive	Unsustainable
Germany	7.0	4.2	-	Unit root	Uncertain
Greece	21.3	8.1	-	Explosive	Unsustainable
Iceland	38.5	4.1	-	Explosive	Uncertain
Ireland	10.2	12.8	3.4	Explosive	Unsustainable
Italy	14.1	32.2	41.7	Unit root	Uncertain
Japan	11.9	20.4	-	Explosive	Unsustainable
Norway	5.9	4.0	-	Stationary	Sustainable
Portugal	7.2	2.8	-	Unit root	Uncertain
Spain	7.7	8.8	8.5	Unit root	Uncertain
Sweden	15.1	10.5	-	Stationary	Sustainable
Switzerland	58.8	12.8	-	Stationary	Sustainable
UK	12.0	7.0	-	Unit root	Sustainable
US	16.1	6.2	-	Unit root	Uncertain

current state is stationary and if it has a longer duration than the other (potentially non-stationary) state. This is the case for Switzerland and for those countries which only have stationary states, such as Finland, Norway and Sweden. The UK's debt is also labeled as sustainable since the smoothed probability of it currently being in the unit root state is barely above 50% and its other state is stationary according to both test criteria. Argentina is absent from this list, since its explosive state is clearly very severe and were it to occur, it may mean a debt default. Its stationary state only persists for about 13 years, which is not a reasonable enough time span to conclude on long-term debt sustainability and explains why Argentine debt is mainly issued in short-term maturities of three to six years.

Most countries with a current explosive state are characterized as having an unsustainable debt path. These countries are France, Greece, Ireland and Japan. Per-

haps it is surprising that France is in this category. This is because both of its states are deemed as explosive according to at least one test criterion. Since France has a short time series, it could be that a long term debt cycle has not been captured and that, in the long run, French debt is indeed sustainable. However, given the current data used, the model indicates that France's debt is on an unsustainable path. The next section on model robustness considers whether the sample size indeed affects the results in a significant way.

For the countries with unsustainable debt paths, (especially Japan) the explosive state has a relatively long duration compared to the other (in some cases also explosive) state. One can notice that this is not the case for Iceland. Even though its debt is in an explosive state at the moment, this state is rather transitory. Its other - unit root - state lasts for a much longer time period and hence its debt path is at best labeled as uncertain.

The remaining seven countries, (Argentina, Germany, Iceland, Italy, Portugal, Spain and the US) all have uncertain debt trajectories. This is because they are currently in a unit root state, (with the exception of Argentina and Iceland); and in some cases both of their states are governed by a unit root process, such as Germany, Portugal and the US. Italy and Spain each have a stationary state, however their other two states are unit root states. Especially in the case of Italy, its unit root states last for much longer time periods than its stationary state. Iceland, fares probably the worst out of this group, since it has an explosive and a unit root state.

Comparing these results with those in the existing literature is not a straightforward task since the models and data used differ. For instance, Raybaudi et al. (2004) using current account data also analyze Argentina, Japan, the UK and the US. They reach a different conclusion only for Japan. However, no real comparison can be drawn, since the longest data range they use is from 1970:I - 2002:IV. This excludes, for example, Japan's rampant debt increase in the 2000s. Chen (2011), who also uses current account data, though expanded until 2009:III, reaches the same conclusion for Portugal and Spain and a different conclusion for Finland. However, again the data range is much shorter than the one investigated here, starting from 1975:I for Finland and Portugal and from 1983:I for Spain. Further, only two states are used throughout the analysis and there is no switching variance term, making the conclu-

sions - especially for Spain - hard to compare.

Finally, it is worth mentioning that, according to the global stationarity conditions for univariate MS models by Francq and Zakoian (2001), all models are found to be globally stationary. This seems somewhat surprising as conventional unit root tests in Table 3.2 reject stationarity in most cases. Whether this indicates some lack of power in the test itself or is indeed a reliable result is an interesting issue to examine. However, it is beyond the scope and purpose of this paper.

Other potential issues worth considering are refinements that can be made to the MS-ADF model considered here. For instance, it may be a good avenue to explore separate parameter switches. There is no reason to assume that all parameters need to switch simultaneously. Perhaps some interesting dynamics can be captured by letting them switch separately, according to their own regime. Although, with this approach there may be some estimation issues when using small samples since the number of states also needs to increase.

3.8 Robustness Analysis

One potential problem with this analysis is that some countries have a short sample range and hence, very long term debt cycles may not be captured. Even though it is standard practice in the literature to use data starting from after the war years, this may provide erroneous results. Other issues to consider are about the model specifications used in this paper. Would similar results be obtained by just using two-state models, and how can the results be compared with the model used in Chen (2011)? These issues are investigated in this section.

3.8.1 Shortening the data

An easy way to determine whether the results are subject to the data range used is to cut the sample, re-estimate the parameters and observe if any of the conclusions change. This approach is naturally, only feasible for the countries with longer time series. Further, when using a shorter sample size, in most cases the lags are reduced since Portmanteau tests indicate no residual autocorrelation at lower lag orders.

For example, starting from the beginning of the 1950s for Argentina, Finland, Japan,

Sweden, the UK and the US, the new smoothed probabilities of being in state 1 are depicted in Figure 3.3. One can see that, with the exception of Japan, the smoothed probabilities for the period examined are almost identical to the ones in Figure 3.1. However, parameter estimates indicate different conclusions about the nature of the states. For instance, state 2 for Argentina, Finland and Sweden is now a unit root state, while for the UK it is a stationary state. State 1 remains the same as before in all cases except for the US, where it is indicated as a stationary state. The above conclusions would nevertheless, probably be unchanged for these countries, given the duration of the states and current debt levels. For example, even though Swedish debt is currently in a unit root state, it is a transitory state, which also captures large downward movements in the debt/GDP ratio. The only exception is perhaps Argentina, where the explosive state is now attenuated to a unit root state.

As already noted, the smoothed probabilities for Japan with the short data sample²⁸, (Figure 3.3, (d)) look very different to the ones for the same period when using the full data range, (Figure 3.1, (g)). Parameter estimates however, indicate that the same conclusion as before still holds, since now both states are indicated as explosive by both test criteria.

For the remaining countries in Figure 3.3 the same analysis is conducted, except with different starting dates. For instance, taking data from shortly after WWI excludes the first occurrence of state 2 for Iceland. This leads to the same smoothed probabilities as before, however, the model now indicates that state 2 is an absorbing state. In other words, going back to the 1920s, the Icelandic debt process has never experienced a similar state as it is currently in. Parameter estimates confirm, as before, that state 1 is a unit root state, while state 2 is an explosive state. From this example it is also clear that a MS model is more general than say a smooth transition (ST) model. If there happens to be an absorbing state it is captured; however, unlike a ST model, a MS model allows for switches back to the original state if the data do so indicate.

Finally, trimming the data series of Portugal to start from 1930 removes the sporadic shifts from the earlier period in the smoothed probabilities, (Figure 3.1, (i)) and, hence, shows a higher persistence of state 1, (Figure 3.3, (e)). Parameter estimates reaffirm that both states are still unit root states.

²⁸The sample starts from 1955 in order to avoid interpolation issues.

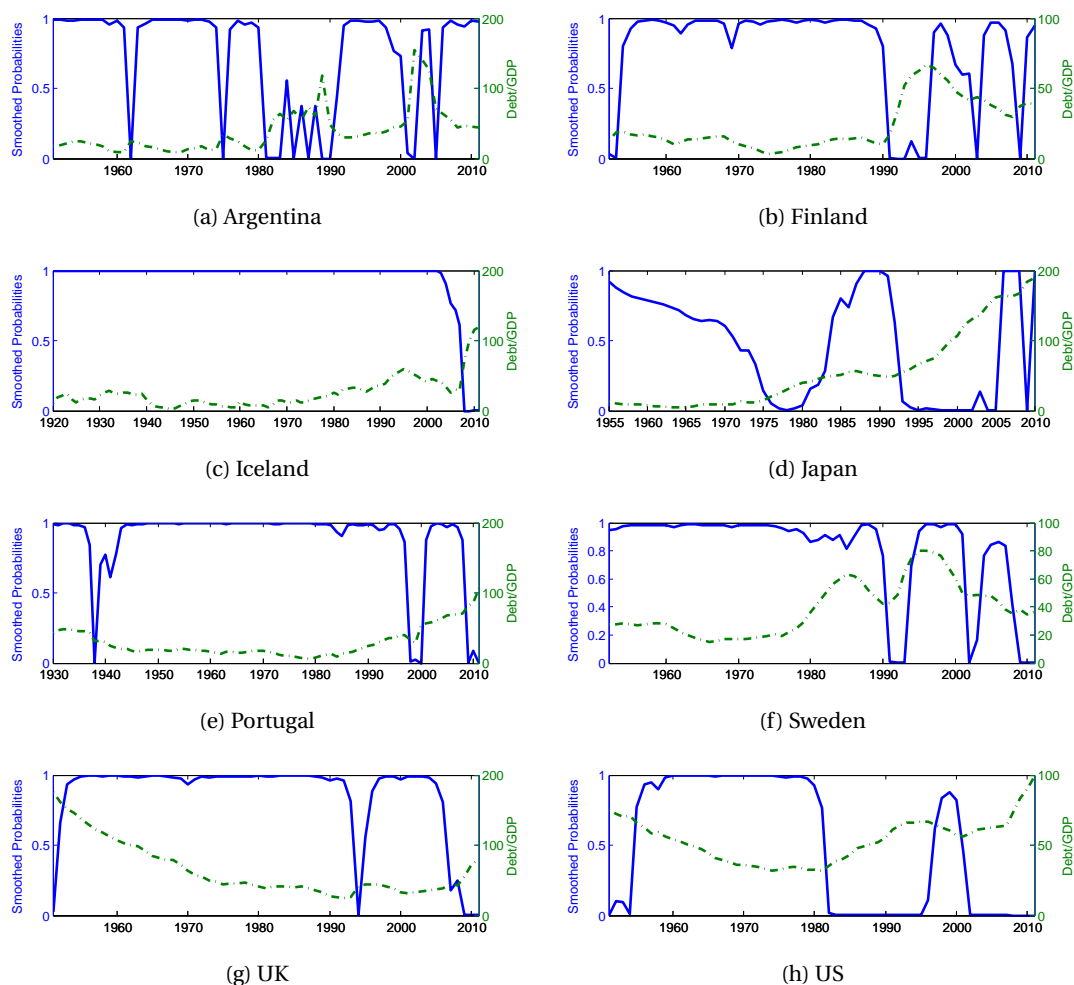


Figure 3.3: Smoothed probabilities of State 1 with shorter sample ranges.

Removing periods of economic and social upheaval from the data usually does not change the smoothed probabilities compared to the original ones over the same time period. State 1 is almost always the same as before. However, state 2 sometimes changes its characteristic. Nevertheless, in most cases the original conclusions are still upheld, making the model rather robust to the sample size used.

3.8.2 Using only two states

Only Ireland, Italy and Spain have models with three Markov states. It is argued that such models fit their data well and portray a better picture than a 2-state model can.

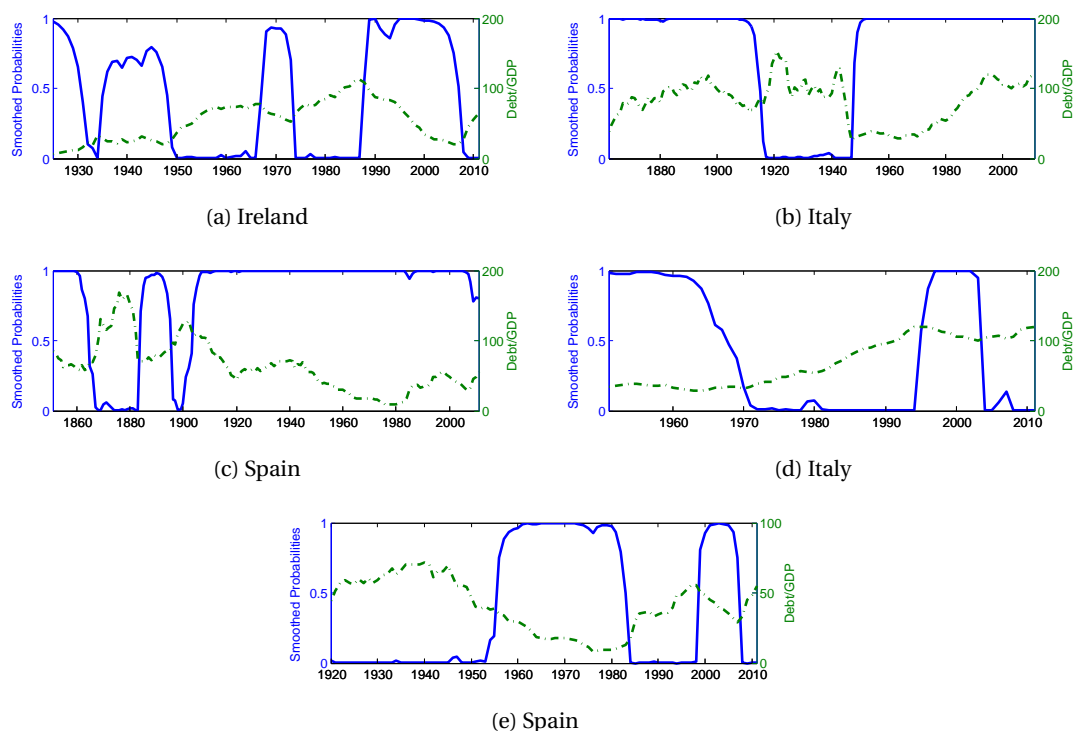


Figure 3.4: Smoothed probabilities of State 1 for Ireland, Italy and Spain using a two-state model and with shorter sample ranges for Italy and Spain.

In particular, there are periods for Italy and Spain, which have caused violent shifts in their debt/GDP series - that are captured by one of the states in a 3-state model. Nevertheless, using only a 2-state model for their data, the smoothed probabilities of being in state 1 are shown in Panels (a) through (c) of Figure 3.4.

When comparing Panel (a) of Figure 3.4 with Panel (a) of figure 3.2, one can see that the smoothed probabilities of state 1 for Ireland appear very similar in both models. In the 2-state model, state 1 is deemed as stationary according to standard errors, while state 2 is a unit root state. The 2-state model has mixed the unit root and the (short-lasting) explosive state into one. At best Irish debt can now be characterized as uncertain, since it is currently in the unit root state and that state has a longer duration than the stable state.

For Italy and Spain the results are even more extreme when using only two states. In both cases, state 1 in the 2-state model is a mixture of states 1 and 3 of the three-

state model. The fact that state 2 of the 3-state model stands out on its own could be because it captures periods of very high volatility. For Italy, both states are unit root states, and the more stable first state of the 3-state model can no longer be distinguished. For Spain, state 1 is stationary according to standard deviations, (although its coefficient point estimate is less negative than before) while state 2 is a unit root state.

Starting the analysis after WWII and WWI for Italy and Spain respectively excludes the very volatile periods from the sample. The smoothed probabilities with this shorter data range are shown in Panels (d) and (e) of Figure 3.4. Now the periods captured by states 1 and 2 in the 2-state model are almost identical to the ones captured by states 1 and 3 in the 3-state model. Parameter estimates reveal that state 1 is stationary (according to standard deviations) and state 2 is a unit root state. These are exactly the same conclusions as before.

Using fewer states is more parsimonious, however, it also means that some rich data dynamics may not be captured. In this case, even when using fewer states, the original conclusions for these countries' debt trajectories remain largely unchanged.

3.8.3 Using a standard model for all countries

This paper advocates in favor of selecting a proper model for a given time series. In particular, Kremers (1988) shows that not using the proper lag order - to remove residual autocorrelation - could lead to erroneous conclusions. Whether the more general models in this paper indeed offer an improvement can be seen by comparing their results with the results one would obtain by using the model in Chen (2011). This model is a simplified version of equation (3.6) of the form

$$\Delta B_t = v(S_t) + \phi_1(S_t)B_{t-1} + u_t, \quad (3.7)$$

where $u_t \sim \text{Nid}(0, \sigma^2)$. In other words, there are no further lag orders involved and no switching variance. Further, no more than two states are considered.

Figure 3.5, (a) shows the typical smoothed probabilities of being in state 1 when using model (3.7). For all data series starting from before the war years, the smoothed probabilities resemble those of the UK, the only exception being Japan. For the UK, the first state is still deemed as stationary, while the second one is a unit root state 1.

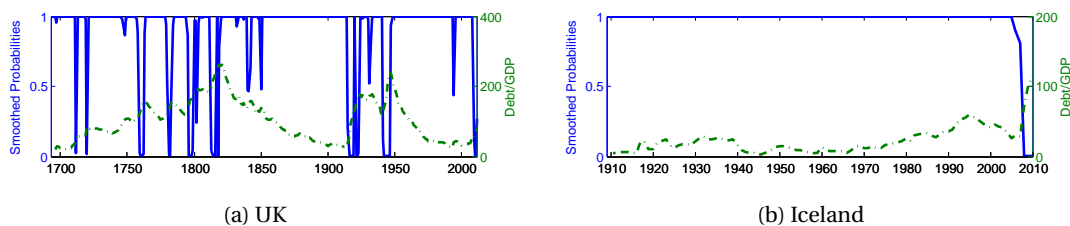


Figure 3.5: Smoothed probabilities of State 1 using model (3.7).

This model is mainly capable of capturing sudden changes in the series investigated. Furthermore, the magnitude of the intercept is often at logical odds with the estimate of the autoregressive parameter. In other words, a state may depict periods that are supposedly more (less) stable, but have a higher (lower) debt/GDP level. This is also found in Raybaudi et al. (2004) and Chen (2011), although it is not commented upon. In that vein it is worth noting that the smoothed probabilities for Iceland when using model (3.7), depicted in Panel (b) of Figure 3.5, show the financial crisis period as a unique event in Iceland's debt history. However, the parameter estimates of ϕ_1 are very similar in both states and are insignificantly different from zero. This is because the intercept parameter in the absorbing state is roughly 50 times higher than it is in state 1. Hence, the previously explosive state, has now turned into a unit root state with a much higher intercept term. Such a result ultimately characterizes both states as unit root states making them therefore indistinguishable from each other. In other words the explosive nature of the second state is no longer apparent when using model (3.7). Although, the original conclusions for both Iceland and the UK would probably be unchanged, clearly some interesting dynamics cannot be captured by model (3.7).

Further, in most cases, when using long-range data, the residuals from model (3.7) show significant signs of autocorrelation. This is the case for instance with Sweden, the UK and the US among others. This renders the conclusions of this model subject to the critique in Kremers (1988), thereby making its results ambiguous. Finally, this model is not robust in terms of reducing the sample size. Unlike the models used in this paper, when using (3.7) the sample length matters for the results one would obtain. Trimming the sample as is done in the earlier part of this section does not give similar conclusions in most cases.

Overall, the models used in this paper do offer an improvement to existing models. They help to better classify different states and are more robust to the sample range used. They are also designed in such a way as to avoid any residual autocorrelation, thereby giving some validity to their findings.

3.9 Conclusion

This paper makes use of a Markov switching ADF (MS-ADF) model to assess the sustainability of public debt by testing whether a government's present value borrowing constraint (PVBC) holds. Building on the work of Raybaudi et al. (2004) and Chen (2011), the model in this paper is of a very general form. The number of lags and states are in principle unrestricted and all of the parameters can be switching. This makes the model resilient to the critique in Kremers (1988), who shows that excluding higher lag orders may mean that there is still residual autocorrelation present, which could lead to erroneous conclusions.

Using the data set from Reinhart and Rogoff (2011), it is possible to obtain long time series on debt/GDP for many different countries. This is in contrast to most of the literature, which uses data starting from after WWII. In total 16 countries are investigated. Several diagnostic tests indicate the presence of structural breaks and nonlinearities in the parameters. This warrants the need of using a MS model. Such a model is appropriate since it is very general and can therefore encompass other models. For instance, a smooth transition model can be thought of as a MS model with an absorbing state. To test the null hypothesis of a unit root in each state, I make use of parameter standard deviations as in Chen (2011) and bootstrapped critical values as in Hall et al. (1999).

The countries with a sustainable debt path are found to be Finland, Norway, Sweden, Switzerland and the UK. These countries either only have stationary states or their debt is currently in a stationary state. In contrast, the model indicates that France, Greece, Ireland and Japan have unsustainable debt trajectories. This is because their debt is currently in an explosive state and in some cases both states are explosive. The remaining seven countries, (Argentina, Germany, Iceland, Italy, Portugal, Spain and the US) are all found to have uncertain debt trajectories. This is be-

cause their debt is currently in a unit root state (with the exception of Argentina and Iceland); and in some cases both of their states are governed by a unit root process.

Robustness tests are performed to investigate the validity of the original findings. First, the influence of the sample size on the results is determined by shortening the original data and observing whether any of the original conclusions change. It is found that the smoothed probabilities are in most cases almost identical for the short and long sample ranges over the same time period. State 1 is usually the same as before. However, state 2 sometimes changes its characterization and is less persistent than before. Most of the conclusions are nevertheless, unchanged, making the model largely robust to the sample size used. Second, two Markov states are used for all models originally in three states. The original conclusions are unchanged when using fewer states, however some of the rich dynamics in the data are not captured with only two states. Finally, the more general MS-ADF model in this paper is compared with the one in Chen (2011). Both the smoothed probabilities and the parameter estimates are more meaningful for the model proposed in this paper. Residuals of the model advocated in this paper are also not subject to autocorrelation, thereby warranting the conclusions as Kremers (1988) observes. Further, unlike the model in this paper, the model in Chen (2011) is not robust to the sample range used. Overall, it is found that this model is an improvement to existing models investigating debt sustainability.

Bibliography

- A. Afonso. Fiscal sustainability: the unpleasant european case. *FinanzArchiv/Public Finance Analysis*, pages 19–44, 2005.
- DE Allen and W. Yang. Do uk stock prices deviate from fundamentals? *Mathematics and computers in simulation*, 64(3):373–383, 2004.
- G. Amisano and C. Giannini. Topics in structural VAR econometrics. *Berlin/Heidelberg*, 1997.
- M. Binswanger. Stock market booms and real economic activity: Is this time different? *International Review of Economics & Finance*, 9(4):387–415, 2000.
- M. Binswanger. How important are fundamentals? Evidence from a structural VAR model for the stock markets in the US, Japan and Europe. *Journal of International Financial Markets, Institutions & Money*, 14(2):185–201, 2004a.
- M. Binswanger. How do stock prices respond to fundamental shocks? *Finance Research Letters*, 1(2):90–99, 2004b.
- M. Binswanger. Stock returns and real activity in the g-7 countries: Did the relationship change during the 1980s? *The Quarterly Review of Economics and Finance*, 44(2):237–252, 2004c.
- O.J. Blanchard and D. Quah. The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review*, 79(4):655–673, 1989.
- H. Bohn. The behavior of us public debt and deficits. *The Quarterly Journal of Economics*, 113(3):949–963, 1998.

- H. Bohn. Are stationarity and cointegration restrictions really necessary for the intertemporal budget constraint? *Journal of Monetary Economics*, 54(7):1837–1847, 2007.
- J. Burbidge and A. Harrison. An historical decomposition of the great depression to determine the role of money. *Journal of Monetary Economics*, 16(1):45–54, 1985.
- J.Y. Campbell and R.J. Shiller. Stock prices, earnings, and expected dividends. *Journal of Finance*, 43(3):661–676, 1988. ISSN 0022-1082.
- N.F. Chen, R. Roll, and S.A. Ross. Economic forces and the stock market. *Journal of Business*, 59(3):383–403, 1986. ISSN 0021-9398.
- S.W. Chen. Current account deficits and sustainability: Evidence from the oecd countries. *Economic Modelling*, 2011.
- Y.W. Cheung and L.K. Ng. International evidence on the stock market and aggregate economic activity. *Journal of Empirical Finance*, 5(3):281–296, 1998. ISSN 0927-5398.
- H. Chung and B.S. Lee. Fundamental and nonfundamental components in stock prices of pacific-rim countries. *Pacific-Basin Finance Journal*, 6(3):321–346, 1998.
- N.E.R. Crafts and G. Toniolo. *Economic growth in Europe since 1945*. Cambridge Univ Pr, 1996.
- T. Davig. Periodically expanding discounted debt: a threat to fiscal policy sustainability? *Journal of Applied Econometrics*, 20(7):829–840, 2005.
- E. Di Nolfo. *Power in Europe? II: Great Britain, France, Germany, and Italy, and the origins of the EEC, 1952-1957*, volume 2. Walter de Gruyter, 1992.
- D.A. Dickey and W.A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, pages 427–431, 1979.
- C. Francq and J.M. Zakoiian. Stationarity of multivariate markov-switching arma models. *Journal of Econometrics*, 102(2):339–364, 2001.

- V.J. Gabriel and P. Sangduan. Assessing fiscal sustainability subject to policy changes: a markov switching cointegration approach. *Empirical Economics*, pages 1–15, 2011.
- V.J. Gabriel, Z. Psaradakis, and M. Sola. A simple method of testing for cointegration subject to multiple regime changes. *Economics Letters*, 76(2):213–221, 2002.
- Ø. Gjerde and F. Sættem. Causal relations among stock returns and macroeconomic variables in a small, open economy. *Journal of International Financial Markets, Institutions and Money*, 9(1):61–74, 1999. ISSN 1042-4431.
- S. Goncalves and L. Kilian. Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, 123(1):89–120, 2004.
- N. Groenewold. Fundamental share prices and aggregate real output. *Applied Financial Economics*, 14(9):651–661, 2004.
- C.S. Hakkio and M. Rush. Is the budget deficit “too large”? *Economic inquiry*, 29(3):429–445, 1991.
- S.G. Hall, Z. Psaradakis, and M. Sola. Detecting periodically collapsing bubbles: a markov-switching unit root test. *Journal of Applied Econometrics*, 14(2):143–154, 1999.
- J.D. Hamilton. *Time series analysis*. Princeton Univ Pr, 1994.
- J.D. Hamilton and M. Flavin. On the limitations of government borrowing: A framework for empirical testing. *American Economic Review*, 1986.
- J.D. Hamilton and G. Lin. Stock market volatility and the business cycle. *Journal of Applied Econometrics*, 11(5):573–593, 1996.
- A.A. Haug. Cointegration and government borrowing constraints: Evidence for the united states. *Journal of Business & Economic Statistics*, pages 97–101, 1991.
- H. Herwartz and H. Lütkepohl. Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks. 2011.

- C. James, S. Koreisha, and M. Partch. A VARMA analysis of the causal relations among stock returns, real output, and nominal interest rates. *Journal of Finance*, 40(5): 1375–1384, 1985. ISSN 0022-1082.
- R. Jean and T. Eldomiaty. How do stock prices respond to fundamental shocks in the case of the United States? Evidence from NASDAQ and DJIA. *The Quarterly Review of Economics and Finance*, 50(3):310–322, 2010. ISSN 1062-9769.
- S. Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, USA, 1995. ISBN 0198774508.
- R. King, C.I. Plosser, J.H. Stock, and M.W. Watson. Stochastic trends and economic fluctuations, 1992.
- J.J.M. Kremers. Long-run limits on the us federal debt. *Economics Letters*, 28(3):259–262, 1988.
- HM Krolzig. Markov-Switching Vector Autoregressions–Modelling. *Statistical Inference and Application to Business Cycle Analysis*, 1997.
- D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root* 1:: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178, 1992.
- M. Lanne and H. Lütkepohl. Structural vector autoregressions with nonnormal residuals. *Journal of Business & Economic Statistics*, 28(1):159–168, 2010.
- M. Lanne, H. Lütkepohl, and K. Maciejowska. Structural vector autoregressions with markov switching. *Journal of Economic Dynamics and Control*, 34(2):121–131, 2010.
- N.T. Laopodis. Are fundamentals still relevant for european economies in the post-euro period? *Economic Modelling*, 26(5):835–850, 2009.
- WD Lastrapes. International evidence on equity prices, interest rates and money. *Journal of International Money and Finance*, 17(3):377–406, 1998.
- B.S. Lee. Fundamentals and bubbles in asset prices: Evidence from us and japanese asset prices. *Asia-Pacific Financial Markets*, 2(2):89–122, 1995a.

- B.S. Lee. The response of stock prices to permanent and temporary shocks to dividends. *Journal of Financial and Quantitative Analysis*, 30(1), 1995b.
- B.S. Lee. Permanent, temporary, and non-fundamental components of stock prices. *Journal of Financial and Quantitative Analysis*, 33:1–32, 1998.
- H. Lütkepohl. *New introduction to multiple time series analysis*. Springer, 2005. ISBN 3540401725.
- H. Lütkepohl and M. Kräzig, editors. *Applied Time Series Econometrics*. Cambridge University Press, Cambridge, 2004.
- Z. Psaradakis. Bootstrap-based evaluation of markov-switching time series models. *Econometric Reviews*, 17(3):275–288, 1998.
- Z. Psaradakis and N. Spagnolo. Joint determination of the state dimension and autoregressive order for models with markov regime switching. *Journal of Time Series Analysis*, 27(5):753–766, 2006.
- C.E. Quintos. Sustainability of the deficit process with structural shifts. *Journal of Business & Economic Statistics*, pages 409–417, 1995.
- D.E. Rapach. Macro shocks and real stock prices. *Journal of Economics and Business*, 53(1):5–26, 2001.
- M. Raybaudi, M. Sola, and F. Spagnolo. Red signals: current account deficits and sustainability. *Economics Letters*, 84(2):217–223, 2004.
- C. M. Reinhart and K. S. Rogoff. From financial crash to debt crisis. *American economic review*, 101(5):1676–1706, 2011.
- P. Saikkonen and H. Lütkepohl. Testing for the cointegrating rank of a var process with structural shifts. *Journal of business & economic statistics*, 18(4):451–464, 2000.
- G.W. Schwert. Why does stock market volatility change over time? *Journal of finance*, 44(5):1115–1153, 1989.
- N.R. Swanson and C.W.J. Granger. Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92(437):357–367, 1997.

- E. Tanner and P. Liu. Is the budget deficit "too large"?: Some further evidence. *Economic Inquiry*, 32(3):511–518, 1994.
- B. Trehan and C.E. Walsh. Testing intertemporal budget constraints: theory and applications to us federal budget and current account deficits. *Journal of Money, Credit and Banking*, 23(2):206–223, 1991.
- P.J.G. Vlaar. On the asymptotic distribution of impulse response functions with long-run restrictions. *Econometric Theory*, 20(05):891–903, 2004. ISSN 1469-4360.
- D.W. Wilcox. The sustainability of government deficits: implications of the present-value borrowing constraint. *Journal of Money, Credit and Banking*, 21(3):291–306, 1989.