DIW BERLIN

1359

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2014

Assessing the Sustainability of Government Debt

On the Different States of Debt/GDP Process

Anton Velinov

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

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

January 2014

Anton Velinov
DIW Berlin and European University Institute, Florence
Mohrenstr. 58, D-10117 Berlin, Germany
email: avelinov@diw.de
tel. +493089789460

Abstract

This paper discusses the type of trajectory a country's public debt path follows. In particular, a Markov switching ADF model is used to assess 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 generalizes their methodology. The number of lags and states are in principle unrestricted and all of the parameters can be switching. Debt trajectories of 16 countries are investigated using long time series on debt/GDP obtained from Reinhart and Rogoff (2011). 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, while Greece and Japan are found to have unsustainable debt trajectories. The debt paths of the remaining countries are mainly characterized as being in a unit root state and are therefore labeled as uncertain. Robustness tests indicate that the model is robust to the sample size and the number of states used. Further, it is shown that the models used in this paper offer an improvement to existing models investigating this subject.

Key Words: Markov switching, debt trajectory, debt sustainability, unit root

JEL classification: C22 C24 H60

1 Introduction

The ramifications of the late 2000s financial crisis have been very severe. With the spectacular bankruptcy of Lehman Brothers in late 2008 and the subsequent stock market collapse, all signs were pointing to a severe recession. In order to deal with this unprecedented situation, governments around the world initiated stimulus packages to help kick-start the ailing economy. As part of these measures, massive loan guarantees were made and financial institutions received large amounts of public money in order to stay afloat. As a consequence of these measures many governments, especially in developed countries, around the world were left straddled with high debt burdens. For instance, the UK and Ireland, 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: what type of trajectory does a country's public debt path follow? This issue has been widely studied. 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 developed by Dickey and Fuller (1979) (ADF test) or Kwiatkowski et al. (1992) (KPSS test) for example. In other words public debt is deemed sustainable if it follows a stationary trajectory, thereby precluding the existence of bubbles, which could lead to a default.

The Markov switching (MS) framework seems to offer an improvement to this basic approach. It is a popular method of modelling endogenous regime shifts, which is very useful since many studies (such as Tanner and Liu (1994) and Quintos (1995)) find evidence of structural breaks and nonlinearities in the debt process. However, in the existing literature, the MS models used to assess debt sustainability are rather restricted. For instance, not all parameters in the Markov switching ADF (MS-ADF) model are allowed to switch, especially the variance parameter is held constant across states.[1] This may not be an accurate way of modelling the data since results in this paper show evidence of heteroskedasticity. Further, there is usually no clear reasoning given to justify why the remaining parameters need to switch.

[1] With the cointegration test model for instance, Gabriel and Sangduan (2011) do allow all three parameters of that model to switch.

Another limitation is that higher order autoregressions are left out,[2] which could lead to erroneous conclusions as Kremers (1988) points out. In addition, existing studies only make use of two Markov states, which may potentially fail to capture the rich data dynamics. Finally, most studies use short range data, which may not contain the very long-term debt cycles that can last for half a century or more, as claimed in Reinhart and Rogoff (2011).

This paper therefore contributes to analyzing the issue of debt sustainability by means of a very general MS-ADF model for many different countries. Generalizing on the existing work of Raybaudi et al. (2004) and Chen (2011), 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. Moreover, I use the rich data set from Reinhart and Rogoff (2011) containing observations on debt/GDP for many countries and usually for very long time periods. This can lead to a more accurate picture by narrowing parameter estimates uncertainties. Further, I bootstrap critical values to test the null hypothesis of a unit root in each regime as in Hall et al. (1999).

The empirical results indicate that the debt path of Finland, Norway, Sweden, Switzerland and the UK is sustainable. Debt trajectories are usually characterized as sustainable depending on whether the current state is stationary and whether the other (potentially non-stationary) state is not very adverse. On the other hand, it is found that Greece and Japan have unsustainable debt trajectories. Such trajectories are concluded when the current state is explosive and the other state is also non-stationary. The debt paths of the remaining countries are labeled as uncertain since in most cases they are currently in a unit root state, or only have unit root states.

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

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

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

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 public 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. 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 may 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 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 sustainability of several 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]

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. Briefly, he states that if a series is integrated of order m say, its n-period-ahead conditional expectation can at most be an mth-order polynomial of the n time horizon, and this is discounted exponentially at a rate of n. He therefore 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 levels beyond which they

[3]Related 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 cointegration 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).

[4]Intuitively, 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.

would be unwilling 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 the type of states a country's debt process has experienced and which state it finds itself in at present. This way one can better judge whether public debt is currently on a sustainable path or not. This kind of model can hence paint a clearer picture of how the debt process has and will develop. The model is presented in the next section.

3 The model

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

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

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 the real or nominal interest rate depending on how the other variables are measured (see Hakkio and Rush (1991)). In each subsequent period there is a similar borrowing constraint, for $t+1, t+2, \ldots$, etc. Hence, the present value borrowing constraint (PVBC) is obtained by solving (1) forward:

$$B_t = \sum_{s=1}^{\infty} \prod_{j=1}^{s} \beta_{t+j}(R_{t+s} - G_{t+s}) + \lim_{s\to\infty} \prod_{j=1}^{s} \beta_{t+j} B_{t+s}, \tag{2}$$

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:

$$\lim_{s\to\infty} \prod_{j=1}^{s} \beta_{t+j} B_{t+s} = 0. \tag{3}$$

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 (1) to obtain

$$E_t + (1 + r)B_{t-1} = R_t + B_t, \tag{4}$$

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

$$B_{t-1} = \sum_{s=0}^{\infty} \beta^{s+1}(R_{t+s} - E_{t+s}) + \lim_{s\to\infty} \beta^{s+1} B_{t+s}, \tag{5}$$

where $\beta = 1/(1+r)$. Again for debt sustainability, the transversality condition needs to hold, in that the second term in (5) needs to be zero. If that is the case the term on the right hand side of (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]

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 endogenously determines regime switches, the following MS-ADF model is applied to test for unit roots

$$\Delta B_t = \nu(S_t) + \phi_1(S_t)B_{t-1} + \phi_2(S_t)\Delta B_{t-1} + \phi_3(S_t)\Delta B_{t-2} + \ldots + \phi_{p+1}(S_t)\Delta B_{t-p+1} + u_t, \quad (6)$$

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 are 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 Markov states. The next section discusses the data to be analyzed with the MS-ADF model in (6).

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

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

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 debt. Again, assuming the absence of Ponzi games, the last term on the right-hand side of the equation needs to go to zero in the limit. 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.

[6] An 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.

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 of the gross central (or when unavailable general) government debt-to-GDP ratio, both series are measured in nominal terms. I extend the data from 2010 to 2013 (so as to include the most recent debt crisis) based on the latest data provided by the OECD and the Federal Reserve Economic Data (FRED) database. For a detailed description of the other data sources one is referred to Reinhart and Rogoff (2011). The countries investigated and their sample ranges are summarized in Table 1.

Table 1: Countries and the starting date of observations (all series are until 2013).

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

[1] One period is interpolated.

[2] Several periods are interpolated.

In the literature on debt sustainability - starting with Hamilton and Flavin (1986) - it is common practice to use annual frequency data.[7] Unfortunately, there is also an unavailability of higher frequency data for many countries prior to the 1990s. Unlike the majority of the literature, which typically uses data starting from after the world war years, this study uses a much longer data range. For instance, for the UK and Sweden the data start from 1692 and from 1719 respectively. This leads to a decent number of observations, which is useful for improving estimation precision and capturing long-term debt cycles, which as Reinhart and Rogoff (2011) point out, can last for half a century or more

For the analysis I choose several countries as representative of a certain fiscal policy. For 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

[7] This is evident in Table 1 of Afonso (2005).

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

at the onset of the debt crisis. Another set of stable countries are the Nordic states, represented by Finland, Norway and Sweden, these are small and open economies. The remaining countries investigated are Argentina, France and Japan.

5 Diagnostic Tests and Model Selection

Having discussed the data and the countries of investigation, the next step involves selecting an appropriate specification of the MS-ADF model in (6) (repeated below for convenience) for each data series.

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

$u_t \sim \text{Nid}(0, \sigma^2(S_t))$. In particular, the task is to select the appropriate lag length, p, the number of states, M and to determine 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] Model diagnostic tests are useful in determining which other parameters could be subject to regime switches.

5.1 Diagnostic Tests

Table 2 summarizes the results of common model diagnostic tests for all series investigated. These tests are based on the residuals of a simple autoregressive model fitted to the data and are performed using the JMulTi software developed by Lütkepohl and Krätzig (2004).

The lag lengths are chosen using the Akaike Inforamtion 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 model lag order and hence for some countries several lag lengths are reported, albeit those that are supported by at least two criteria.

As is common in this literature, I use ADF and KPSS unit root tests to determine stationarity of the data. According to the literature review discussed in section 2, a stationary debt/GDP series means that the PVBC holds, and is hence a sufficient condition to conclude a sustainable public debt policy. Table 2 shows that for only 2 of the 16 countries investigated, both the ADF and the KPSS test statistics indicate a stationary debt/GDP process.[10] Norway, Sweden and the UK only have a stationary series according to the ADF test, while for Italy and

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

[10] These are for Argentina with a 4-lag model (at the 1% level for the KPSS test) and Finland.

Table 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^{A}_{12} [2]	LM_5 [3]	LMF_5 [4]	$ARCH_{LM}(12)$ [5]	$ARCH_{LMF}(12)$ [6]
Argentina	1	-3.70	0.78	0.26	0.23	0.22	0.23	5×10^{-4}	1×10^{-4}
	4	-3.31	0.56	0.25	0.20	0.88	0.89	2×10^{-3}	3×10^{-4}
Finland	2	-3.07	0.32	0.35	0.30	0.15	0.16	0.01	9×10^{-3}
France	2	1.50	2.66	0.98	0.98	0.55	0.59	0.25	0.09
Germany	2	0.39	3.00	0.88	0.84	0.30	0.34	0.96	0.93
Greece	1	1.99	2.99	0.57	0.46	0.23	0.25	0.82	0.71
Iceland	2	-0.47	2.54	0.89	0.86	0.66	0.68	0.99	0.99
Ireland	2	-1.83	1.82	0.24	0.19	0.14	0.15	0.53	0.40
Italy	2	-2.40	0.72	0.44	0.40	0.30	0.31	0.06	0.03
Japan	2	1.18	1.87	0.48	0.44	0.33	0.35	0.04	0.02
	4	0.20	0.99	0.52	0.47	0.23	0.24	0.07	0.03
Norway	2	-4.01	1.06	0.66	0.57	0.84	0.86	0.71	0.56
Portugal	5	-0.64	0.51	0.42	0.37	0.91	0.92	0.84	0.81
Spain	2	-1.94	4.73	0.67	0.65	0.42	0.43	7×10^{-4}	3×10^{-4}
	3	-2.22	3.20	0.75	0.71	0.73	0.74	2×10^{-4}	1×10^{-4}
Sweden	5	-2.58	1.22	0.09	0.07	0.22	0.23	3×10^{-3}	1×10^{-3}
Switzerland	2	-1.92	1.26	0.71	0.65	0.54	0.57	8×10^{-4}	1×10^{-4}
UK	2	-2.54	1.62	0.19	0.18	0.07	0.07	5×10^{-4}	2×10^{-4}
	3	-2.75	1.09	0.28	0.26	0.24	0.25	3×10^{-4}	5×10^{-4}
US	2	-1.08	4.31	0.69	0.66	0.71	0.72	2×10^{-4}	1×10^{-4}

† 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 a with 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.

Portugal only the KPSS test indicates stationarity. This inconclusiveness of the tests is a key driver for both Raybaudi et al. (2004) and Chen (2011) to use a MS model. In particular, Chen (2011) argues that due to the nonlinear nature 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 something adverse, 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.[11] 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 2 present p-values of residual Portmanteau and LM autocorrelation tests, both with adjusted test statistics more suited to small samples. All p-values are quite high and hence the null of no residual autocorrelation cannot be rejected.

The final two columns of Table 2 show the p-values of heteroskedasticity tests. In particular, ARCH-LM tests with a χ^2 and an F distribution. For most short range series it appears that there is no conditional heteroskedasticity present due to the high p-values. However, this conclusion may be unreliable due to the small sample size; 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. These tests are carried out for a range of possible break dates and indicate that, for all data series, there is evidence of a structural break at certain time points in the relevant data range. This therefore leads to a rejection of the null hypothesis of stable parameters and lends evidence in favor of a non-constant variance over time, as most ARCH tests indicate. Recursive tests for the AR coefficients and the intercept term reinforce the conclusion of breaks in these param-

[11] This is barely true for Greece, the null hypothesis of stationarity of the KPSS test is accepted at the 1% level.

eters. For some models however, the intercept parameter is stable according to these tests. Finally, the CUSUM-SQ test rejects parameter stability in almost all cases.

Overall, no single model is indicated as having stable parameters across time by all of the tests. This argues in support of a switching parameter model and it may mean that the unit root tests conducted thus far are not too reliable as Afonso (2005) and Chen (2011) point out. The exact type of MS model per data series is discussed in the following.

5.2 Model Selection

Thus far we have seen that the data is subject to nonlinearities and structural breaks. Hence, a regime switching model, such as a Markov switching (MS) model is warranted. Such a model allows for endogenous regime switches, without having to impose break dates. To determine the appropriate model specifications I rely on the information criteria developed by Psaradakis and Spagnolo (2006). In particular, I use their AIC, BIC and HQC criteria, which help to simultaneously determine the number of lags and states of the MS model. These criteria all tend to opt for the most parsimonious configuration of 1 lag and 2 states. However, this setting is not always optimal since Portmanteau tests (based on the residuals of the MS-ADF model) sometimes show significant residual autocorrelation when using a single lag. Further, sometimes the smoothed probabilities of 3-state models capture more meaningful periods than those of 2-state models. Hence, I choose the number of lags and states according to the information criteria proposed by Psaradakis and Spagnolo (2006), Portmanteau tests to ensure no residual autocorrelation and whether the smoothed probabilities convey meaningful results.

It is further necessary to decide which parameters need to switch. As noted earlier, crucial for the analysis is a switching ϕ_1 coefficient. For higher order autoregressions, I allow all autoregressive coefficients to switch. Parameter stability tests indicate that the variance is usually 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 keep it constant.[12]

Table 3 summarizes the models used for each country (based on (6)). The most general

[12] 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.

Table 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(4)AH
Finland	MS(2)-ADF(2)AH	Italy	MS(3)-ADF(2)AH	Switzerland	MS(2)-ADF(1)AH
France	MS(2)-ADF(1)A	Japan	MS(2)-ADF(1)AH	UK	MS(2)-ADF(2)AH
Germany	MS(2)-ADF(1)A	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.

model syntax is a MS(M)-ADF(p)IAH model, where 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. Note that a single lag model in Table 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.

The models used in this paper generalize earlier studies using the MS methodology by firstly allowing for higher lag orders. For instance, the US, which is one of the most studied countries in this literature, needs have a model with more than one lag so as to avoid the critique by Kremers (1988) of having residual autocorrelation, which could lead to erroneous conclusions. Such residual autocorrelation is present when using a single lag model for the US. Secondly, the variance is allowed to switch between states. As is discussed later, this provides more meaningful and accurate results. A further generalization is the use of more than two Markov states for some series. While two states can potentially distinguish between "stable" and "unstable" periods, a better picture can be obtained from a larger number of states. No reasonable results are found for models with more than three states. In section 8 on model robustness, it is shown that the 3-state models used here tend to provide more insight in distinguishing between different historical time periods.

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, which is elaborated in Chapter 21 of Hamilton (1994). Standard deviations of 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 in (6) has a null hypothesis of $\phi_1(S_t) = 0$, for $S_t = 1, \ldots, M$. 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 also used here. 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 their 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. However, due to the highly nonlinear nature of the models, bootstrapping may produce a diverse range of critical values since there are many local optima that the estimator could converge on. 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 models with more than than two states this issue could potentially become even more severe. Of course, one could increase the number of bootstrap replications in the hope of alleviating this problem, 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, this would be a notoriously time consuming exercise. Therefore, I rely on both testing criteria to determine parameter significance.

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 this section provides an overall summary.

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.

Parameter estimates of these models are given in Table 4. So as to save space, the autoregressive parameter estimates of models with more than one lag order are not reported. Further, it is indicated which coefficients are stationary, $\widehat{\phi}_1(m) < 0$, and which coefficients are explosive, $\widehat{\phi}_1(m) > 0, m = 1, \ldots M$. The criteria used to test this are discussed in the previous section, namely the standard deviations and the bootstrapped critical values. 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 test criterion. In some cases both criteria reach the same conclusion, which makes the classification the given state fairly straightforward. To obtain a better picture of each state it is also very helpful to observe the smoothed probabilities; these are shown in Figure 1. Note that, for a 2-state model, the smoothed probabilities of one state are the mirror images of the other state.

Starting with Argentina, the model indicates that its debt-to-GDP process has a stationary and an explosive state. According to bootstrapped critical values (not reported here) the coefficient on the explosive state is highly significant, well above the 1% critical value. Argentina has had 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% over the past four years. Argentina's economy has managed to pull through and its debt is now at around 40% of GDP. In spite of this rather chaotic debt history, Argentina is one of the only countries in Table 2 indicated

Table 4: Parameter estimates for 2-state models, standard deviations in parentheses.

	$\ln(\ell)$	$\widehat{p}_{11}$	$\widehat{p}_{22}$	$\widehat{\nu}$	$\widehat{\phi}_1(1)$	$\widehat{\phi}_1(2)$	$\widehat{\sigma}^2(1)$	$\widehat{\sigma}^2(2)$
Argentina	456.650	0.923 (0.027)	0.632 (0.113)	2.557 (0.499)	-0.137 (0.014)†¶	0.214 (0.157)*	8.496 (1.181)	701.888 (205.782)
Finland	258.943	0.901 (0.057)	0.694 (0.207)	0.485 (0.343)	-0.044 (0.017)†¶	-0.041 (0.058)†	2.759 (0.752)	145.061 (54.688)
France	151.120	0.845 (0.278)	0.441 (0.930)	-0.749 (0.954)	0.029 (0.021)¶	0.118 (0.042)*	4.804 (2.643)	
Germany	126.246	0.888 (0.070)	0.698 (0.184)	-0.151 (0.503)	0.009 (0.010)¶	0.087 (0.031)*	2.302 (0.299)	
Greece	174.766	0.957 (0.030)	0.896 (0.106)	0.889 (0.619)	4×10^{-4} (0.009)	0.073 (0.028)**¶	7.171 (1.666)	90.961 (43.480)
Iceland	301.831	0.972 (0.020)	0.637 (0.253)	0.537 (0.578)	-0.015 (0.021)¶	0.199 (0.135)*	12.224 (2.031)	290.769 (173.991)
Japan	441.681	0.900 (0.058)	0.964 (0.022)	-0.120 (0.574)	-0.058 (0.046)†	0.058 (0.010)**¶	159.934 (39.827)	13.162 (2.238)
Norway	168.807	0.832 (0.092)	0.779 (0.150)	5.892 (0.880)	-0.262 (0.031)‡¶	-0.230 (0.047)‡	2.417 (0.732)	23.314 (7.776)
Portugal	415.614	0.857 (0.054)	0.658 (0.126)	0.509 (0.391)	-0.013 (0.012)	0.001 (0.022)¶	3.086 (0.601)	47.497 (11.327)
Sweden	617.011	0.947 (0.025)	0.928 (0.035)	0.320 (0.170)	-0.024 (0.010)‡	-0.021 (0.010)‡¶	1.004 (0.145)	15.050 (2.057)
Switzerland	188.025	0.984 (0.016)	0.923 (0.075)	0.911 (0.473)	-0.059 (0.023)†¶	-0.016 (0.039)	2.664 (0.497)	56.083 (22.152)
UK	923.884	0.919 (0.024)	0.857 (0.042)	0.857 (0.292)	-0.020 (0.005)‡¶	0.001 (0.007)	4.161 (0.723)	98.184 (15.795)
US	458.330	0.939 (0.024)	0.817 (0.064)	-0.106 (0.120)	-0.003 (0.004)¶	0.002 (0.016)	0.949 (0.168)	39.151 (8.013)

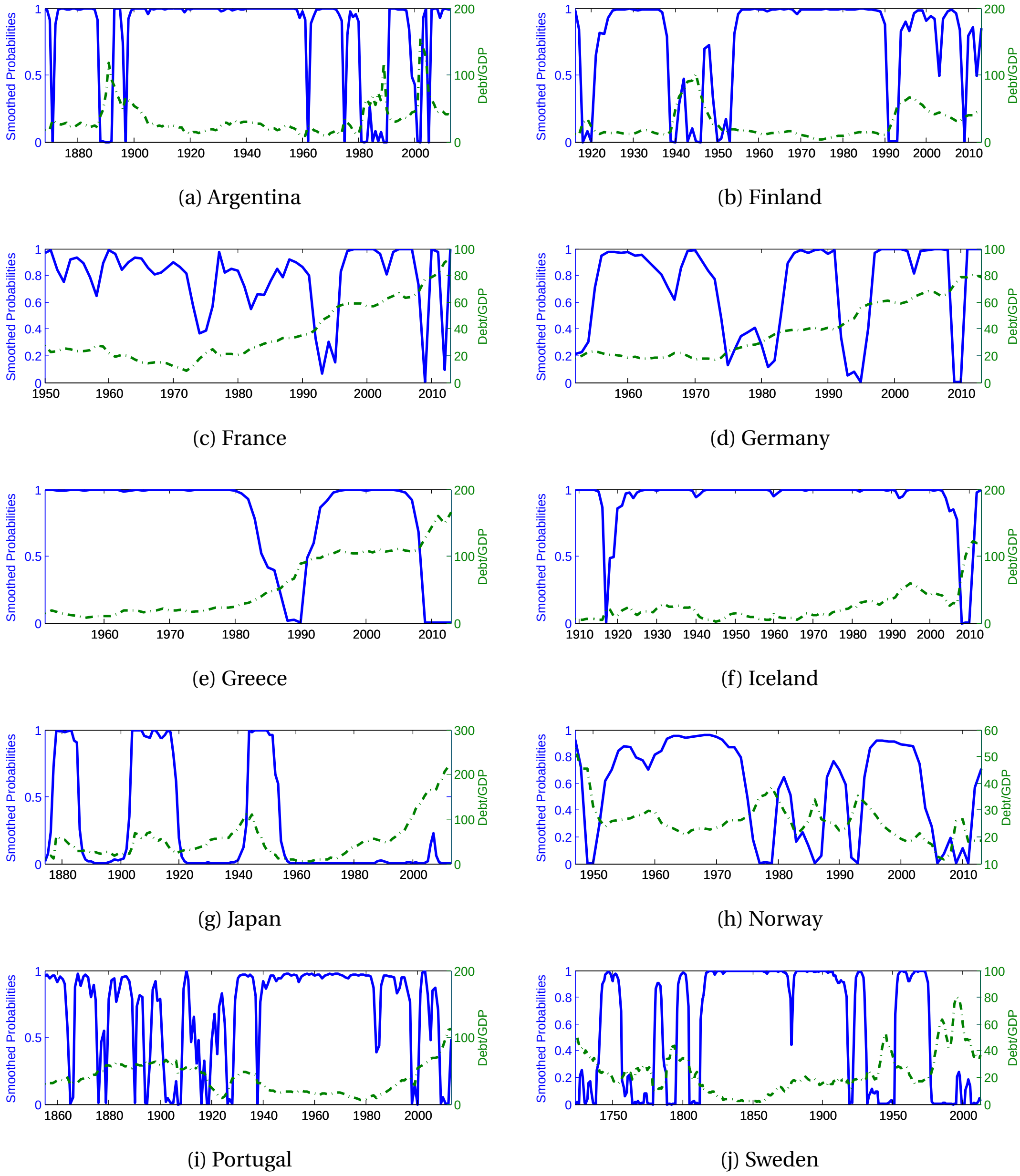
¶ The current state.

† Stationary according to one criterion.

‡ Stationary according to both criteria.

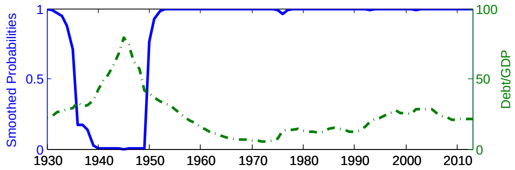
* Explosive according to one criterion.

** Explosive according to both criteria.

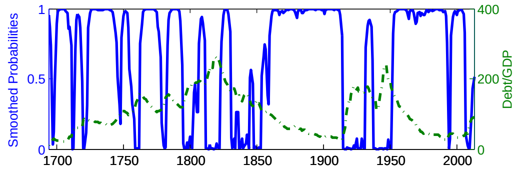


(a) Argentina (b) Finland (c) France (d) Germany (e) Greece (f) Iceland (g) Japan (h) Norway (i) Portugal (j) Sweden

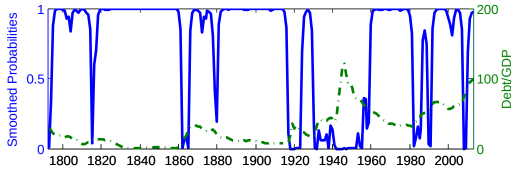
Figure 1: Smoothed probabilities of State 1, solid line (left axis) with the respective debt/GDP series, dashed line (right axis)



(k) Switzerland



(l) UK



(m) US

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

as having a stationary debt process by both unit root tests. This may not be reassuring for investors who lost money during the Argentine debt default. It further illustrates the limitations of only looking for unit roots and concluding on debt sustainability. The MS model provides an insight of the different states that the Argentine debt process finds itself in.

The smoothed probabilities in Panel (a) of Figure 1 reveal that the second state is associated with the 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 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. Currently, Argentina's debt process is in the better of the two states, however, since there is a very bad state in which the debt process can potentially switch to, I would classify the overall path of Argentine debt as uncertain. This is more than can be said by just simply looking at the ADF and KPSS test values in Table 2. It is one of the merits of using a MS model.

Moving on to Finland, Table 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 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. Both the Argentine and Finish debt processes are classified as stationary according to both unit root tests in Table 2. However, the MS model further indicates that Finish debt only has stationary states, hence, overall it seems to be on a sustainable path, unlike Agentina's debt process.

France appears to have a less stable debt trajectory than Finland, the estimated ϕ_1 coefficient for the first state indicates a unit root, while the one for the second state can be classified as explosive according to standard errors. Smoothed probabilities show that the second state follows shortly after the 1973 oil crisis and its last occurrence is during the financial crisis. It seems to be a transient state with its smoothed probabilities almost never reaching unity. Currently French debt is in the first state and in light of these results I would argue that this state is by far the most important one. It is a unit root state, though it can be labeled as quasi-stable as defined in Chen (2011). This indicates that France's debt trajectory need not be unsustainable, however its sustainability is uncertain. Finally, note that the MS model for French debt does not make use of a switching variance parameter since there is little evidence in support of it.

The same MS-ADF model used for French debt is applied to German debt. As with French debt the parameter estimate of the ϕ_1 coefficient for state 1 is insignificantly different from zero, while the same coefficient for state 2 is characterized as explosive according to standard errors. Smoothed probabilities (Panel (d) of Figure 1) for German debt are comparable to the ones for French debt (Panel (c)), however for the former, the explosive state seems to be more persistent. This is evidenced by the higher values of the transition probabilities. Currently German debt is in its first state, and even though its second state is slightly more prominent than is the case with French debt, I would still argue that its debt trajectory is best characterized as uncertain; the first state can be seen as quasi-stable, and since it dominates Germany's debt process its debt need not be headed towards an unsustainable path, though rather an uncertain one.

The model indicates that Greek debt also has a unit root and an explosive state. In fact both test criteria confirm that the second state is explosive. The smoothed probabilities in Panel (e) of Figure 1, show that the first time the Greek debt process was in that state was during the 1980s, a period characterized by high inflation and weak economic growth.[13] 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. Cur-

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

rently Greek debt is in the second, explosive and more volatile state, which seems to justify the present concerns regarding its debt trajectory. Given these results the Greek debt trajectory is deemed as unsustainable.

Similar to Greece, ever since the financial crisis, Iceland has been marred by debt problems. The model indicates state 1 as being a unit root state, while state 2 is deemed as explosive according to bootstrapped critical values. The smoothed probabilities in Panel (f) of Figure 1 show that state 2 is indeed associated with more turbulent times. It has only occurred for a brief time during WWI and for a longer time during the financial crisis. This would explain the much higher variance of that state compared with the first state. Overall, the first state could be characterized as quasi-stable, while the second state depicts an unsustainable debt path. Currently Iceland's debt process is in the more favorable of its two states and due to the high persistence of that state I would label its debt trajectory as uncertain rather than unsustainable.

In this analysis Japan is the country with the highest debt/GDP ratio of currently over 220%. It is therefore probably not surprising that its second state is characterized as explosive by both criteria - well beyond the 1% bootstrapped critical value. Its first state even appears to be stationary and can be characterized as quasi-stable. Smoothed probabilities show that the debt process during the first state either remained stable or declined. This state was present during the war years and it is therefore the more volatile of the two states according to variance parameter estimates. Japan's debt trajectory is however dominated by the second state, which is the one it is currently in. Since this state is explosive and very persistent, it seems that Japan's debt trajectory is unsustainable.

In contrast with Japan's debt, this analysis characterizes Norway's debt process as being highly sustainable since both its Markov states are significantly stationary as indicated by both testing criteria. The only difference among the states is that the second state captures more volatile periods according to the variance estimates. Norwegian debt is currently in the less volatile of the two stationary states. Note that Norway's debt process has the lowest ADF test value in Table 2 and hence the original conclusion of stationarity would remain unchanged.

Both states of Portuguese debt are unit root states since their ϕ_1 parameter estimates are insignificantly different from zero. Panel (i) of Figure 1 shows that the states have switched in a rather erratic fashion during the years up to and including WWI. The variance estimates show that state 1 is associated with less volatile periods, such as the high growth period of

1950 - 1973. The military coup of 1974 is in this state as well, indicating that this event did not destabilize government debt in any severe way. Currently, Portugal's debt is in state 2 with a smoothed probability of 52%. Since Portuguese debt only has unit root states its debt trajectory is best labeled as uncertain.

The observations on Swedish debt begin in 1719, shortly before the end of the Great Northern War (1700 - 1721). Since that war Sweden has been involved in numerous conflicts with Russia lasting until the early 1800s. Even these times of conflict did not increase Sweden's debt burden to much above 40% of its GDP. In fact only WWII and the crisis of the 1990s have caused more noticeable increases in Sweden's 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. The model unambiguously indicates both states of Swedish debt as stationary. The only difference between the two states is that the first one is associated with less volatile periods than the second one - as shown by the variance parameter estimates and the model smoothed probabilities. Currently, Sweden's debt process is in the second sate. This state is also associated with large downward movements in the debt series, which would explain its current presence. In light of these results, this analysis suggests that Sweden has a sustainable government debt. Finally, note that in Table 2 only the ADF test would have judged Sweden's debt path as stationary and that barely at the 10% level. The MS model takes into account structural breaks and offers a better understanding Sweden's debt process.

Unlike some other countries, the states of Switzerland's debt are not very exciting. Its debt process is mainly governed by the first state, which can be labeled as quasi-stable or even stationary according to standard errors. The second state clearly captures the WWII period and is a unit root state. During this state Swiss debt tended to fluctuate more, rather than head towards ever increasing levels, and hence it is also the more volatile of the two states. It could well be a state in which debt is sustainable, though it is clearly an exceptional state. Overall, since the quasi-stable state dominates Switzerland's debt process, its debt is likely to be on a sustainable path. This is a stronger conclusion than could have been reached by only looking at the unit root tests in Table 2.

The longest debt series in this analysis is that of the UK, with the first observation starting in 1692. The series is depicted in Figure 1, (l) by the dashed line. 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) it reached more than 200% of GDP. Since that spike it gradually fell over the years to a mere 25% of GDP just prior to the outbreak of WWI. This 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 at around 90% of GDP.[14] This illustrates the claim made by Reinhart and Rogoff (2011) that government debt can be subject to very long cycles persisting for half a century or more.

The model indicates that the UK's debt process has a stationary and a unit root state. The stationary state is most often associated with periods in which debt/GDP is declining, according to the smoothed probabilities. The unit root state on the other hand tends to be present during many of the war years and other such turbulent times. Currently, the UK's debt is marginally in its stationary state with a smoothed probability for this state of 52%. Since this is unambiguously a stationary state according to both test criteria, and since the unit root state has not destabilized the debt process thus far, I would conclude that the UK's debt process is on a sustainable path.

The final country with a two Markov state model is the US. This is also one of the most studied countries in this line of literature with the seminal paper by Hamilton and Flavin (1986) analyzing US debt. The present study makes use of a considerably longer data span, starting from 1790, slightly after the American War of Independence. The US debt/GDP series shown in the dashed line in Figure 1, (m) appears to be very low most of the time. This is with the exception of some turbulent events, such as the War of 1812, the American Civil War of the 1860s, the World Wars, the Great Depression and the latest financial crisis, where government debt has seen significant increases.[15] The model classifies both states of American debt as unit root states. Smoothed probabilities show that state 2 captures the more politically and economically unstable periods mentioned above. Hence, it is the one with the higher variance. Currently, American debt is in the less volatile of its unit root states, which is a positive sign. However, since there are only unit root states, the sustainability of US debt is characterized as uncertain.

7.2 Results of the 3-state models

Ireland, Italy, and Spain's debt processes are modeled with three Markov states. The relevant parameter estimates and the smoothed probabilities are reported in Table 5 and in Figure 2

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

[15] Remarkably, debt/GDP declined throughout the whole period of the Vietnam War.

Table 5: Parameter estimates for 3-state models, standard deviations in parentheses.

Parameter	Ireland	Italy	Spain
$\ln(\ell)$	260.358	489.514	491.160
$\widehat{p}_{11}$	0.900 (0.041)	0.929 (0.050)	0.880 (0.042)
$\widehat{p}_{22}$	0.918 (0.048)	0.967 (0.012)	0.893 (0.248)
$\widehat{p}_{33}$	0.696 (0.177)	0.958 (0.027)	0.893 (0.095)
$\widehat{v}$	1.087 (1.270)	3.482 (1.343)	1.535 (0.366)
$\widehat{\phi}_1(1)$	-0.084 (0.021)†	-0.078 (0.026)†	-0.100 (0.010)†
$\widehat{\phi}_1(2)$	0.018 (0.021)¶	-0.052 (0.028)	-0.014 (0.028)
$\widehat{\phi}_1(3)$	0.190 (0.075)**	-0.015 (0.023)¶	-0.005 (0.014)¶
$\widehat{\sigma}^2(1)$	8.830 (2.481)	7.303 (2.035)	4.480 (0.927)
$\widehat{\sigma}^2(2)$	12.598 (4.139)	200.124 (49.226)	268.035 (46.975)
$\widehat{\sigma}^2(3)$	35.731 (17.772)	30.728 (5.770)	18.574 (4.869)

¶ The current state.

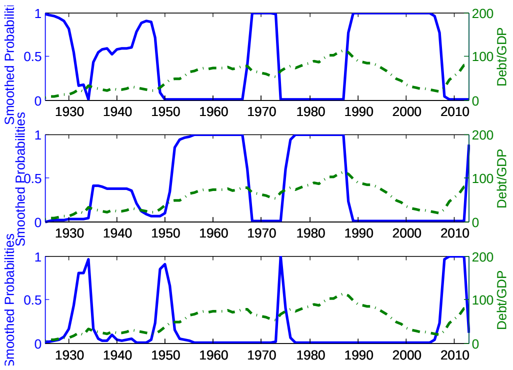
† Stationary according to one criterion.

** Explosive according to both criteria.

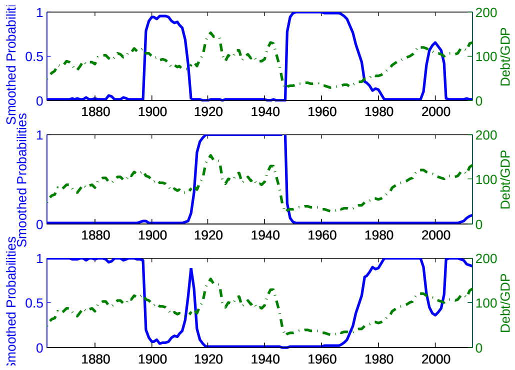
respectively.

Beginning with Ireland's debt process, the estimated ϕ_1 coefficients indicate that state 1 is stationary (or quasi-stable), state 2 follows a unit root process, while state 3 is explosive as indicated by both test criteria. Smoothed probabilities show that the first state is present during periods in which debt/GDP is stable or slightly declining, as can be seen by the dashed line representing the series. The second state depicts periods of moderately increasing debt/GDP, while the third state is present, among others, during the first oil crisis and most notably during the financial crisis, hence, depicting explosive debt trajectories. Although there is such an explosive debt state present, this need not cause sustainability problems since it tends to be rather transient compared with the other states. Hence, I would conclude that Irish debt is on an uncertain trajectory rather than an unsustainable one. Note that it is currently in the second state, which is a unit root state, and hence, labeled as uncertain.

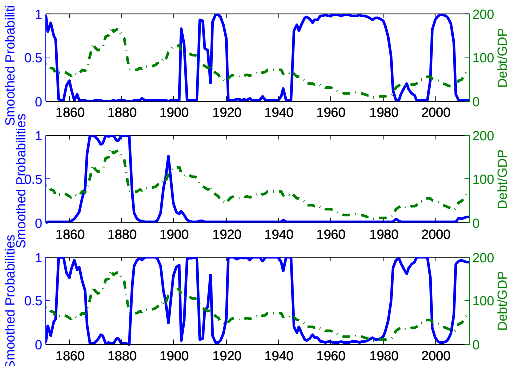
All three values of the ϕ_1 coefficients for Italian debt are negative, however, only the value for state 1 is significantly so according to its standard deviation. The other two states seem to be governed by a unit root process. The smoothed probabilities in Panel (b) of Figure 2 show



(a) Ireland



(b) Italy



(c) Spain

Figure 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)

that the first state is associated with stable and high-growth periods. From 1951 until 1973 the Italian economy grew at a rate of slightly higher than 5% per annum on average. This was one of the highest growth rates among European countries.[16] State 1 also depicts the period from the mid 1990s until before the financial crisis. The second state marks the exceptional period of WWI up to and including WWII, which is characterized by the highest volatility in the debt-income ratio. Italian debt is currently in the third state, the less volatile of the two unit root states. It depicts the period of political instability of the 1970s and the economic recession from that time. It is also present during the steady rise of debt throughout the 1980s and 1990s until debt reached 120% of GDP around the mid 1990s. Apparently Italy's debt path is dominated by the two unit root states and hence, it is best to characterize it as uncertain with

[16] It has been documented in Crafts and Toniolo (1996) and in Di Nolfo (1992).

regards to sustainability. Note, that this is a weaker conclusion than the KPSS test gives in Table 2, however, in light of the recent concerns regarding Italian debt, it would appear more to be a reliable one.

Finally, the model indicates that the sustainability of Spain's debt trajectory is similar to Italy's in that the first state seems stationary, while the other two follow a unit root process. The smoothed probabilities in Panel (c) of Figure 2 show that state 1 captures the economic stabilization of the mid 1950s and the "Spanish miracle" of 1959 to 1975 and ends prior to the early 1990s recession. It is present again during the Spanish property boom of 1997 - 2007. Clearly, this state depicts the high-growth and stable periods in Spain's debt history. State 2, which is the most volatile state, captures some earlier periods of drastic shifts in debt/GDP. Currently, Spanish debt is in the third state, which tends to be prevalent in its debt history. As is the case with Italian debt, the two unit root states play the dominant role, and hence, it is uncertain whether Spanish debt is on a sustainable path.

7.3 Summary

In line with Raybaudi et al. (2004) and Chen (2011) I summarize the above results by taking into account the duration of the states, $(1/(1-p_{mm})$, for $m = 1, \ldots, M)$ and the current state the debt process finds itself in. This is implicitly done above and is a convenient way of determining the overall sustainability of the given debt process. Table 6 summarizes the results.

The last column of the table shows that the debt path of most countries is characterized as uncertain. Considering the discussion in the literature review I prefer to err on the side of caution in concluding a (un)sustainable debt path. Debt trajectories are usually characterized as sustainable depending on whether the current state is stationary and whether the other (potentially non-stationary) state is not very adverse. Unsustainable debt trajectories are only labeled for a current explosive state and a non-stationary other state. In most cases debt paths currently in a unit root state, or only having unit root states, are characterized as being uncertain.

Naturally, the characteristic of a given state depends on historical factors and not only on coefficient estimates. For example, the explosive state of Argentine debt can be viewed as being more severe than the explosive state of Greek or Japanese debt. This factor is taken into account when giving an overall assessment of debt sustainability. Indeed, it is up to the researcher to determine the periods that are depicted by the endogenously estimated smoothed probabilities of the MS model, and hence, better classify the given states.

Table 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	13.0	2.7	-	Stationary	Uncertain
Finland	10.1	3.3	-	Stationary	Sustainable
France	6.5	1.8	-	Unit root	Uncertain
Germany	8.9	3.3	-	Unit root	Uncertain
Greece	23.3	9.6	-	Explosive	Unsustainable
Iceland	35.7	2.8	-	Unit root	Uncertain
Ireland	10.0	12.2	3.3	Unit root	Uncertain
Italy	14.1	30.3	23.8	Unit root	Uncertain
Japan	10.0	27.8	-	Explosive	Unsustainable
Norway	6.0	4.5	-	Stationary	Sustainable
Portugal	7.0	2.9	-	Unit root	Uncertain
Spain	8.3	9.3	9.3	Unit root	Uncertain
Sweden	18.9	13.9	-	Stationary	Sustainable
Switzerland	62.5	13.0	-	Stationary	Sustainable
UK	12.3	7.0	-	Stationary	Sustainable
US	16.4	5.5	-	Unit root	Uncertain

As is often the case, there is no straightforward comparison with the results in this paper and those in the existing literature 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, their longest data series is from 1970:I - 2002:IV, which excludes most of Japan's rampant debt growth 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, his data range is also shorter than the one used in this paper, starting from 1975:I for Finland and Portugal and from 1983:I for Spain. Further, both studies only use two Markov states and do not allow for a switching variance term, making the conclusions (especially for Spain) difficult to compare.

Finally worth noting is that according to the global stationarity conditions for univariate MS models developed by Francq and Zakoıan (2001), all models are found to be globally stationary. This seems somewhat surprising as conventional unit root tests in Table 2 reject stationarity in most cases. Whether this indicates some lack of power in the test may be an interesting issue to examine, however, it is beyond the scope and purpose of this paper.

8 Robustness Analysis

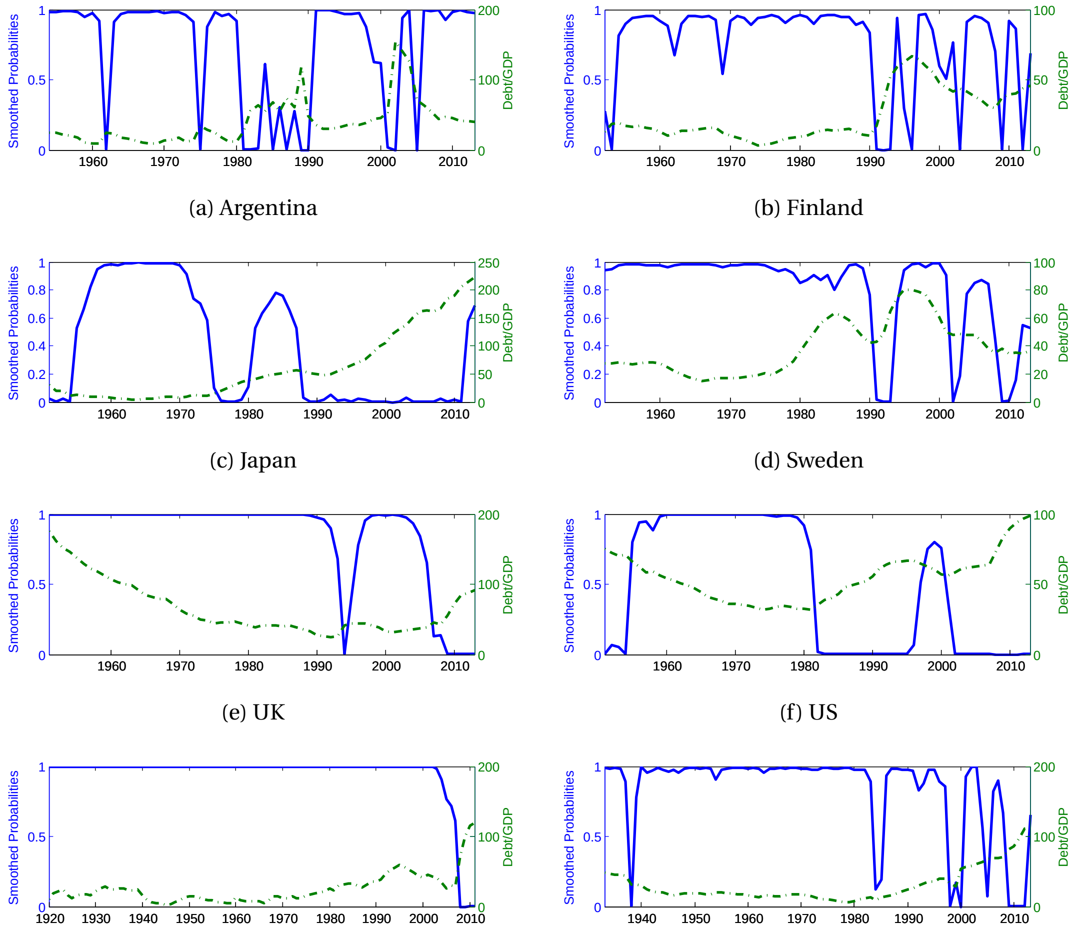
This section investigates the robustness of the findings above in several ways. Firstly, the data is shortened to examine whether the results are driven by the specific time period investigated. Secondly, it is assessed to what extent models with three Markov states can depict a clearer picture than models with two Markov states. Finally, the approach used in this paper is compared with that used in Chen (2011).

8.1 Shortening the data

Re-estimating the parameters with shortened data provides a means of verifying robustness of the results, and thus also potentially strengthening the conclusions obtained for the countries with shorter sample ranges. In particular, with many of the European countries considered, data on debt is not available during the war years and consequently only the period from the early 1950s onwards is investigated. Would examining the same period for some of the countries with longer data series change the original conclusions?

This issue is investigated for Argentina, Finland, Japan, Sweden, the UK and the US. I conduct the same analysis as before, though only using the latter part of their data starting from the early 1950s.[17] The smoothed probabilities of the debt process of these countries being in state 1 are depicted in Figure 3 Panels (a) through (f). Most often the smoothed probabilities for the specific period examined are very similar to the ones in Figure 1. Parameter estimates however, do sometimes offer different conclusions from the original ones concerning the nature of the states. For instance, state 2 for Argentina, Finland and Sweden is indicated as a unit root state. State 1, on the other hand, remains the same as before, except for Japan and the US where it is labeled as explosive and stationary respectively. The original conclusions would nevertheless be largely upheld when taking into account the duration of the states. For example, even though the UK's debt is currently indicated as being in a unit root state, this

[17] With a shorter sample size the number of lags are reduced when Portmanteau tests indicate no residual autocorrelation at lower than the original lag orders.



(a) Argentina (b) Finland (c) Japan (d) Sweden (e) UK (f) US (g) Iceland (h) Portugal

Figure 3: Smoothed probabilities of State 1 with shorter sample ranges.

appears to be a rather transitory state, while the other state is stationary according to both test criteria.

Smoothed probabilities for Iceland's debt process depicted in Panel (g) of Figure 3 show that state 2 is an absorbing state. This happens when using data from shortly after WWI until 2011. Parameter estimates indicate that the second state is an explosive state, which is characteristic of the financial crisis period. This example shows 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 endogenous switches back to the original state if the data do so indicate. Taking this into account, and the durations of both states, the original conclusion regarding Iceland's debt trajectory would be unchanged.

Finally, using data starting from 1930 for Portugal's debt excludes the sporadic shifts from the earlier period in the smoothed probabilities depicted in Figure 1, (i). One may argue that such shifts do not convey any meaningful results and may in fact give a distorted view of the true nature of the states. It is hence reassuring to observe that with the shortened data the shape of the smoothed probabilities (Figure 3, (h)) for the specific time period remains largely the same as before. Parameter estimates also reaffirm that both states are unit root states leaving the original conclusion unchanged.

Overall, shortening the data usually does not change the shape of the smoothed probabilities compared with the original one covering the same time period. Further, state 1 is usually characterized in the same way as before, while state 2 sometimes changes its characteristic. Nevertheless, the original conclusions are still largely upheld, making the model robust to the sample size used.

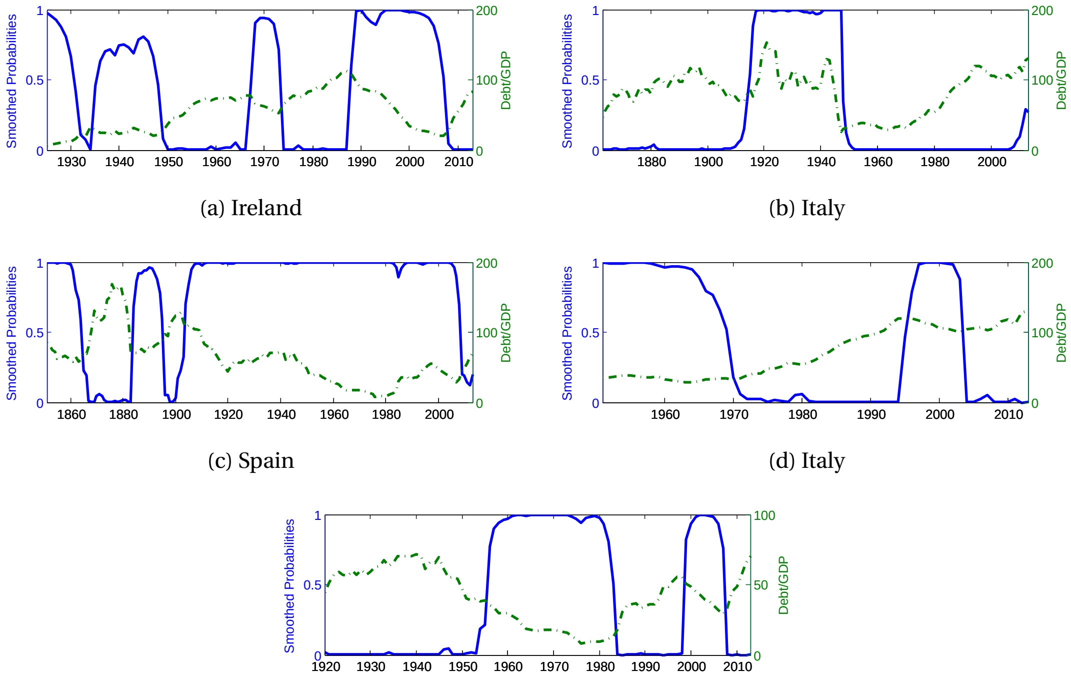
8.2 Using only two states

This paper generalizes the existing literature by using more than two Markov states for some of the debt series. To illustrate whether this genuinely improves the analysis, I use a two state model for all series previously modeled with three states. Hence, a model with two Markov states is used for Irish, Italian and Spanish debt, smoothed probabilities of which are shown in Figure 4, Panels (a) through (c).

The smoothed probabilities of Irish debt being in state 1 with a two and a three state model (depicted in Panel (a) of Figure 4 and in Panel (a) of Figure 2 respectively) appear very similar. In the model with two Markov states, the first state is stationary (according to standard errors), while the second state is a unit root state. This model combines the unit root and the (short-lasting) explosive state from the 3-state model into one state, the second state. I would still characterize the Irish debt path as uncertain, since it is currently in the unit root state, which has a longer duration than the stationary state.

When modeled with two Markov states, the first state of Italian debt is the same as the second state in a 3-state model, while state 2 combines states 1 and 3 of the 3-state model. Both states are unit root states, which means that a 2-state model cannot portray the stationary periods captured by state 1 in the 3-state model. The original conclusion that Italy's debt path is uncertain still holds.

Similarly, the second state in a 2-state model of Spain's debt is very similar to the second



(a) Ireland (b) Italy

(c) Spain (d) Italy

(e) Spain

Figure 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.

state in a 3-state model of that debt process, while states 1 and 3 from the latter model are combined to form state 1 of the former model. With the 2-state model the first state is indicated as stationary and the second state is a unit root state. This model shows that currently Spanish debt is in the unit root state, however due to the strong persistence of the stationary state, it may be tempting to classify Spain's debt path as sustainable. Since I prefer to err on the side of caution I would therefore argue that the original conclusion still remains valid, meaning that Spain's debt trajectory is uncertain.

Thus far it can confidently be said that the number of states does not influence the overall conclusion regarding sustainability of the debt process. Using more Markov states is helpful in portraying a better picture of the historical debt series. To illustrate this further I only use the observations starting from after WWII and after WWI for Italian and Spanish debt respectively so as to avoid the period captured by the second state in the 3-state model. An ADF model with two Markov states is used for these reduced samples, the smoothed probabilities of which are shown in Panels (d) and (e) of Figure 4. Clearly, the periods captured by states

1 and 2 of this 2-state model are almost identical to the ones captured by states 1 and 3 of the 3-state model for the same time period. Parameter estimates for both models label state 1 as stationary (according to standard deviations) and state 2 as a unit root state. These are exactly the same conclusions as before.

Overall, using more Markov states when applicable, enhances the analysis by capturing some of the rich data dynamics, and helps to give a more accurate evaluation with regards to debt sustainability.

8.3 Using a standard model for all countries

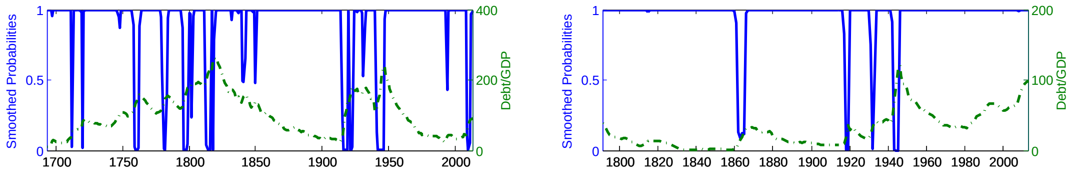
Throughout this paper an emphasis is put on selecting a proper model for the given time series investigated. In particular, Kremers (1988) shows that not using the proper lag order - to remove residual autocorrelation - could lead to erroneous conclusions. To what extent the more general models used in this paper offer an improvement to the existing ones can be seen by comparing the results obtained above with the ones one would obtain by using the model in Chen (2011). This model is a simplified version of equation (6) of the following form

$$\Delta B_t = \nu(S_t) + \phi_1(S_t)B_{t-1} + u_t, \tag{7}$$

where $u_t \sim \text{Nid}(0, \sigma^2)$. In other words, there are no higher lag orders than the first lag of the levels series and no switching variance. In addition, no more than two Markov states are considered.

The typical smoothed probabilities one would obtain using this model are shown in Figure 5. These are for the debt series of the UK and the US, two of the most investigated countries in this literature. According to this model both states of the UK's debt are unit root states, while state 2 of the US's debt is explosive. Hence, one may be lead to conclude that both countries have uncertain debt trajectories. This is a different conclusion than the one obtained above for the UK, namely that it's debt trajectory is sustainable.

There are several issues with the above results however. First, the residuals of both models show signs of significant autocorrelation, with p-values of Portmanteu tests very close to zero. Second, state 2 is rather transient and clearly depicts periods of rapidly increasing debt/GDP. For the UK's debt, this is nonetheless characterized as a unit root state, which could be due to the much higher intercept term in that state, roughly 15 times higher than in state 1. This may erroneously characterize an explosive state as a unit root state. Moreover, for other debt series modeled with (7), the magnitude of the intercept is often at logical odds with the estimate of



(a) UK

(b) US

Figure 5: Smoothed probabilities of State 1 using model (7).

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). Third, the model does not seem robust to reduction of the sample size. Trimming the sample as done in the earlier part of this section often does not give the same conclusion. Finally, diagnostic tests and model results in this paper indicate signs of residual heteroskedasticity, hence, not allowing the variance term to switch across states may not model the data well.

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 designed in such a way as to avoid any residual autocorrelation, thereby giving larger validity to their findings.

9 Conclusion

This paper addresses the issue of the type of trajectory a country's public debt path follows. In particular, a Markov switching ADF (MS-ADF) model is used 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. Hence, the model results are resilient to the critique in Kremers (1988), since it is made sure that there is no residual autocorrelation present.

Using the data set from Reinhart and Rogoff (2011) very long time series on debt/GDP are obtained for the 16 countries under investigation. This leads to more accurate parameter estimates and can capture long-term debt cycles. Diagnostic tests indicate the presence of structural breaks and nonlinearities in the parameters. This warrants the use of a regime switching model. To test the null hypothesis of a unit root in each state, I use both parameter standard deviations as in Chen (2011) and bootstrapped critical values as in Hall et al. (1999).

The empirical results indicate that the debt path of Finland, Norway, Sweden, Switzerland and the UK is sustainable. Debt trajectories are usually characterized as sustainable depending on whether the current state is stationary and whether the other (potentially non-stationary) state is not very adverse. On the other hand, it is found that Greece and Japan have unsustainable debt trajectories. Such trajectories are concluded when the current state is explosive and the other state is also non-stationary. The debt paths of the remaining countries are labeled as uncertain since in most cases they are currently in a unit root state, or only have unit root states.

Several robustness tests investigate the validity of the original findings. First, it is found that the results are not driven by the particular time period investigated. Indeed, they tend to be rather robust to reduction in the sample size. Second, it can be seen that using more (than two) Markov states when applicable, enhances the analysis by capturing some of the rich data dynamics, and helps to give a more accurate evaluation with regards to debt sustainability. Finally, it is found that 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 designed in such a way as to avoid any residual autocorrelation, thereby giving larger validity to their findings.

References

A. Afonso. Fiscal sustainability: the unpleasant European case. *FinanzArchiv/Public Finance Analysis*, pages 19–44, 2005.

H. Bohn. The behavior of U.S. 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.

S.W. Chen. Current account deficits and sustainability: Evidence from the OECD countries. *Economic Modelling*, 28:1455–1464, 2011.

N.F.R. Crafts and G. Toniolo. *Economic growth in Europe since 1945*. Cambridge University Press, 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. Zakoıan. 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.

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 University Press, 1994.

J.D. Hamilton and M. Flavin. On the limitations of government borrowing: A framework for empirical testing. *American Economic Review*, 76(4):808–819, 1986.

A.A. Haug. Cointegration and government borrowing constraints: Evidence for the United States. *Journal of Business & Economic Statistics*, pages 97–101, 1991.

J.J.M. Kremers. Long-run limits on the U.S. federal debt. *Economics Letters*, 28(3):259–262, 1988.

D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178, 1992.

H. Lütkepohl and M. Krätzig, 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.

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.

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 U.S. federal budget and current account deficits. *Journal of Money, Credit and Banking*, 23(2):206–223, 1991.

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.