Fiscal policy asymmetries and the sustainability of US government debt revisited*

Steven P. Cassou† Kansas State University

Hedieh Shadmani‡ Fairfield University

Jesús Vázquez§ Universidad del País Vasco (UPV/EHU)

April 4, 2015

Abstract

This paper empirically investigates US fiscal policy sustainability and cyclicality in an empirical structure that allows fiscal policy responses to exhibit asymmetric behavior. We investigate this over two quarterly intervals of data, both of which begin in 1955. The short sample, which ends in the second quarter of 1995, was investigated to demonstrate that the financial crisis and Great Recession are sufficiently different from this earlier period, which is most analogous to the one used by Bohn (1998), that the asymmetric empirical methods used in this paper are important and that the sustainability of U.S. government debt topic needed to be revisited. For the full sample, US fiscal policy is asymmetrical in regard to both sustainability and cyclicality. Regarding fiscal policy sustainability, the best fitting models show evidence of fiscal policy sustainability for the short sample. However, the fiscal sustainability question does become less clear for the full sample that includes the recent financial crisis and the Great Recession. Regarding fiscal policy cyclicality, we find that during times of distress, policy is strongly countercyclical, but during good times the results are more mixed with some models showing that fiscal policy can be procyclical.

JEL Classification: E31, E52, E61

Keywords: Asymmetric fiscal policy, fiscal policy sustainability, cyclicality

---

*We are grateful for helpful comments from Máximo Camacho, Josep Lluís Carrion-i-Silvestre and seminar participants at the 89th Annual Conference of the Western Economic Association International (Denver, USA), XXXIX Symposium of the Spanish Economic Association (Palma de Mallorca, Spain) and the IV Workshop in Time Series Econometrics (Zaragoza, Spain). Some of this research was supported by the Spanish Ministry of Science and Innovation through projects ECO2010-16970 and ECO2013-43773-P and the Basque Government through project IT793-13.

† Department of Economics, 327 Waters Hall, Kansas State University, Manhattan, KS, 66506 (USA), (785) 532-6342, Fax:(785) 532-6919, email: scassou@k-state.edu.

‡ Department of Economics, 1073 North Benson Road, Fairfield, CT 06824 (USA), email: hshadman@k-state.edu.

§Depto. Fundamentos del Análisis Económico II, Facultad de Ciencias Económicas y Empresariales, Universidad del País Vasco (UPV/EHU), Av. Lehendakari Aguirre 83, 48015 Bilbao (SPAIN), Phone: +34 946013779, Fax: +34 946017123, email: jesus.vazquez@ehu.es.

1 Introduction

Interest in the sustainability of the US fiscal policy stance toward public debt became a concern to both policy makers and economists during the 1980s when the US started to incur large government budget deficits. Many investigations, including Hamilton and Flavin (1986), Trehan and Walsh (1988), Kremers (1989), Hakkio and Rush (1991), Quintos (1995) and Davig (2005), used unit root and cointegration empirical methods to provide answers to this question.[1] On the other hand, Bohn (1998) used a method that had a policy reaction interpretation. He showed that the US government historically responded to increases in the debt-GDP ratio by increasing the primary surplus-GDP ratio. That is, the government has historically reacted to debt accumulation by taking corrective measures when the debt-GDP ratio starts to rise. In a later paper, Bohn (2007) further defended the policy reaction function approach and built on the criticism of the unit root and cointegration methods, by showing that the work which used such methods could not properly evaluate sustainability because they implicitly ruled out higher-order integration which would still be consistent with the intertemporal government budget constraint. In short, Bohn (2007) showed that the intertemporal budget constraint alone imposes very weak econometric restrictions on debt and budget deficits and thus the policy reaction function approach is an effective way to evaluate the fiscal policy sustainability question.

Although some level of comfort regarding the sustainability of fiscal policy may have been established from this early body of work, the rapidly growing US government debt since the start of the Great Recession in late 2007 has thrust the sustainability question back into the public consciousness, once again raising doubts about fiscal policy sustainability. This paper follows Bohn's (1998, 2007) approach based on the analysis of fiscal reaction functions to revisit the issue of US fiscal pol-

[1] Hamilton and Flavin (1986) Trehan and Walsh (1988), Quintos (1995) and Davig (2005) found evidence in support of sustainability, while Kremers (1989), Hakkio and Rush (1991) were more pessimistic.

icy sustainability by not only extending the data series, but by also extending the empirical model suggested by Bohn (1998) to account for, (i) the possibility of asymmetries in the fiscal reaction function to business cycle conditions, and (ii) changes in the size of shocks during the post-World War II period. More precisely, we investigate three related empirical models. The simplest model is a basic linear regression model which is close to the empirical model used in Bohn (1998) for the US and Collignon (2012) for a sample of European Union countries, while the others are popular switching regression formulations which can address policy asymmetries. One of these non-linear formulations is piecewise linear and is such that the basic linear model is nested as a special case. This type of non-linear model has been used by Golinelli and Momigliano (2008) and Balassone, Francese and Zotteri (2010) to analyze asymmetries in fiscal policy for several European countries. The third formulation is a Markov switching model of the form described by Hamilton (1989). The advantage of a Markov switching model over a piece-wise linear model is that the former allows us to handle changes in shock volatility in a rather straightforward manner.

To investigate these issues, we estimate models over two different quarterly intervals of time, one running from 1955:1 until 1995:2 and the other from 1955:1 to 2013:3. The former, which we call the short sample, is intended to provide a quarterly data series that can roughly match the series in Bohn (1998), while the later, which we call the long sample, uses all available data. By investigating these two samples, we are able to demonstrate that the financial crisis and Great Recession are sufficiently different from the earlier period that the asymmetric empirical methods used in this paper are important and that the sustainablity of U.S. government debt question needed to be revisited. We find several important results regarding fiscal policy modeling and the empirical analysis of fiscal policy for the United States. First, and perhaps most important, we find that recent economic data, that include the financial crisis and Great Recession, are sufficiently different from earlier data that simple linear models are not appropriate for modelling fiscal policy. Second, we find that for the long sample, the fiscal response to lagged debt and lagged output gap are asymmetric, however, over the shorter sample, the results for lagged debt do not exhibit asymmetric behavior. Third, regarding fiscal policy sustainability, we again find that the sample period used is important. For the short sample, fiscal policy is sustainable, confirming results by Bohn (1998). Over the full sample period, fiscal policy is only sustainable during good economic times and is unsustainable during times of distress. We interpret the unsustainable behavior during times of distress as evidence that policy makers, by running larger budget deficits, are more concerned about getting the economy back on track and perhaps temporarily ignore sustainability. This finding that fiscal policy is not always sustainable is consistent with the recent doubts expressed by some policy makers and political pundits about sustainability of the US fiscal policy. Finally, regarding the cyclicality of fiscal policy, we find robust evidence that fiscal policy is countercyclical during bad economic times and becomes less countercyclical during good times. A few of our models even indicate that during good economic times fiscal policy may become marginally procyclical. This procyclical finding is consistent with results found in Balassone et al. (2010) using European data and are interpreted to show that during good times, the government budget deficit-GDP ratio grows with the rest of the economy.

The rest of the paper is organized as follows. Section 2 describes the empirical models analyzed in this paper. Section 3 discusses the empirical findings. In Section 4 we study an alternative sample period that is similar to the one analyzed by Bohn (1998), in order to see if his conclusion about debt sustainability holds up using our richer asymmetric fiscal policy empirical models and thus reconcile our empirical results with his findings. Finally, Section 5 concludes.

2 Empirical equations

We investigate three related empirical models. The simplest model is a basic linear regression model, while the others are popular switching regression formulations. One of the more general formulations is piecewise linear and is such that the basic linear model is nested as a special case while the third formulation is nonlinear and, as such, the other two are not nested in general.

Some authors, such as Golinelli and Momigliano (2008), have considered other models that are further nested in a few of the switching regression formulations considered here. They referred to these different formulations as two parameter models and referred to the piecewise linear formulations considered in this paper as two sample models. The two sample models are more general and nest the two parameter models as restricted cases. In some earlier work, we investigated the two parameter models and tested the restrictions implied between them and the two sample models. We always rejected the restrictions, so here we only present the more general two sample models.

Following Bohn (1998), Balassone et al. (2010) and others, we focus on the primary balance as our measure for the fiscal policy stance.[2] Our models exhibit richer dynamics than the one used in Bohn (1998) and some are similar to the one used by Balassone et al. (2010). In particular, our models include the lagged primary balance as an explanatory variable to allow for fiscal policy persistence.[3] A significant role of the lagged primary balance may reflect, for instance, the existence of an optimal fiscal policy inertia where the fiscal authority aims at reaching the optimal primary balance target in small steps due to economic uncertainty. Also, all our models only include lagged variables as explanatory variables. This feature allows us to interpret our models as fiscal policy rules where the fiscal authority only reacts to variables included in its information set. This feature further helps to overcome endogeneity issues which may show up when current explanatory variables are used in the empirical model. Finally, two of the formulations studied in this paper are nonlinear. We choose this set up as it allows us to not only investigate debt sustainability, but it also allows us to investigate fiscal policy cyclicality. Further, these nonlinear features help to assess the presence of policy persistence since, as noted by Perron (1989), evidence

[2] Bohn (1998) actually used the primary surplus which is the primary balance multiplied by -1.

[3] Afonso, Agnello and Furceri (2010) provided empirical evidence of fiscal policy persistence in a cross-country study.

of a persistent process may arise simply as a result of ignoring important sources of nonlinearity.

The basic regression model can be written as

$$b_t = \alpha + \beta_1 b_{t-1} + \beta_2 d_{t-1} + \beta_3 w_{t-1} + \varepsilon_t, \tag{1}$$

where b_t is the ratio of the primary balance (spending, net of interest expenses, minus taxes) to gross domestic product (GDP) at date t, d_{t-1} is the ratio of federal debt to GDP at date $t-1$, and w_{t-1} is the output gap at date $t-1$.[4] In the tables and discussion below, we will often refer to this model as the linear model or LM for short.

The various anticipated signs for the parameters have been discussed in various places in the literature. For instance, following the arguments in Bohn (1998, 2007), sustainability of fiscal policy should be marked by negative values of β_2.[5] Intuitively a negative value for β_2 indicates that as lagged debt increases, policy makers take action to reduce b_t either by reducing spending or raising taxes. Also, as noted in Balassone et al. (2010), the sign of β_3 depends on whether policy is pro-cyclical or counter-cyclical. A positive value, corresponds to a pro-cyclical policy, where higher values of the lagged output gap are accompanied by higher primary balances either through more government spending or lower taxes, while a negative value corresponds to a counter-cyclical policy.

Part of our empirical focus will be on a popular asymmetry hypothesis in which the policy variable, b_t, responds to the lagged output gap, w_{t-1}, differently, depending on whether the economy is strong or weak. To evaluate asymmetry, one approach is to replace the various variables with two interaction terms in which a dummy indicating the strength of the economy at period $t-1$, I_{t-1}, is multiplied by each of

[4] There are alternative, but related, formulations used in the literature. For instance, Golinelli and Momigliano (2008) focus on the cyclically-adjusted primary balance and prefer to use differenced levels as the dependent variable.

[5] To be more precise, Bohn (1998) used the primary surplus as his dependent variable rather than the primary balance (or deficit) we use here. For his model, Bohn (1998) argued that sustainability is associated with a positive surplus coefficient which would imply that our deficit coefficient should be negative.

the variables to get

$$\begin{aligned} b_t &= \alpha I_{t-1} + \beta_1 I_{t-1} b_{t-1} + \beta_2 I_{t-1} d_{t-1} + \beta_3 I_{t-1} w_{t-1} + \alpha'(1 - I_{t-1}) \\ &\quad + \beta_1'(1 - I_{t-1}) b_{t-1} + \beta_2'(1 - I_{t-1}) d_{t-1} + \beta_3'(1 - I_{t-1}) w_{t-1} + \varepsilon_t, \end{aligned} \tag{2}$$

where I_{t-1} is given by

$$I_{t-1} = \begin{cases} 0 & \text{for} \quad w_{t-1} \leq w^T, \\ 1 & \text{for} \quad w_{t-1} > w^T, \end{cases} \tag{3}$$

and w^T is referred to as the threshold value for the lagged output gap. This type of model is known as a Threshold Regression model and we will use the abbreviation TR to reference it from now on.

The parameters of (2) can provide evidence for a number of interesting policy questions. For instance, asymmetry in the response of the policy variable to economic conditions can be noted when $\beta_2 \neq \beta_2'$ or $\beta_3 \neq \beta_3'$. If $\beta_2 \neq \beta_2'$, then there is evidence that policy responds to lagged debt differently when the economy is doing well than when it is not, while if $\beta_3 \neq \beta_3'$, then there is evidence that policy responds to the lagged output gap differently when the economy is doing well than when it is not. It is also useful to note that (1) is nested in (2). So one can test whether the TR model fits better than the basic regression model by performing an F test on the null that $\alpha = \alpha'$, $\beta_1 = \beta_1'$, $\beta_2 = \beta_2'$, and $\beta_3 = \beta_3'$.

We also investigate two versions of the two state Markov Switching (MS) empirical model. These MS models use a nonlinear relationship between the primary balance and the explanatory variables given by

$$b_t = \alpha(s_t) + \beta_1(s_t) b_{t-1} + \beta_2(s_t) d_{t-1} + \beta_3(s_t) w_{t-1} + \sigma(s_t) u_t, \tag{4}$$

where u_t is a standard normal random variable and s_t denotes the unobservable regime or state variable featuring the reaction of the fiscal authority to both observables entering in its information set (b_{t-1}, d_{t-1} and w_{t-1}) and the fiscal shocks (u_t). This state variable has values of either 1 or 2 and follows a first-order two-state Markov process with transition matrix given by

$$P = \begin{pmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{pmatrix},$$

where the row j, column i element of P is the transition probability p_{ij}, which is the probability that state i will be followed by state j. Since the transition probabilities beginning in each state sum to 1, the off diagonal terms are given by $p_{12} = 1 - p_{11}$ and $p_{21} = 1 - p_{22}$.

A somewhat more flexible alternative two-state MS model allows the transition probabilities to depend on an observable variable of the economy. This is more analogous to the various TR models which switch between the two linear terms depending on the value of the output gap. We investigated this connection in one of our models. In particular, we also considered a two-state MS model that allowed the transition probabilities to depend on the lagged output gap and used a logistic functional form for the probabilities as follows

$$p_{11}(w_{t-1}) = \frac{\exp(\theta_{10} + \theta_{11} w_{t-1})}{1 + \exp(\theta_{10} + \theta_{11} w_{t-1})},$$

$$p_{22}(w_{t-1}) = \frac{\exp(\theta_{20} + \theta_{21} w_{t-1})}{1 + \exp(\theta_{20} + \theta_{21} w_{t-1})}.$$

This type of formulation is described in Filardo (1994). Because the transition probabilities are time varying, we refer to this model as a time varying transition probability Markov Switching model, or TVTP-MS for short. One thing to notice is that when the lagged output gap does not impact the transition probabilities (i.e. when $\theta_{11} = \theta_{21} = 0$) the TVTP-MS model collapses into the MS model with constant probabilities. In this case, there is a one-to-one map between p_{11} and θ_{10} and between p_{22} and θ_{20}.

The advantage of these MS models over the TR models is twofold. First, they allow changes in the size of shocks during the sample period. Second, they allow the interaction of fiscal policy asymmetries with changes in shock volatility. However, it is important to point out that fiscal policy sustainability issues in a MS framework does require some extra consideration. As shown by Francq and Zakoïan (2001), local stationarity conditions, i.e. stationarity within each regime, are neither sufficient nor necessary to achieve the global stationarity of a MS system, and thus, for our application, to rule out explosive behavior that results in the violation of the intertemporal budget constraint. But, as also pointed out by Francq and Zakoïan (2001), this situation occurs when the probability of staying in the same regime is small. As a consequence, we argue that the likelihood of explosive behavior largely decreases when local stationary regimes are largely persistent (i.e. when the probability of staying in the same stationary regime is close to one).

3 Empirical results

Our empirical analysis mostly uses quarterly data for the US economy obtained from the Federal Reserve Bank of St. Louis data set. There are two possible government perspectives one could use. One could use the total of the federal, state and local government budgets, which would be more closely connected to the data used by Balassone et al. (2010) in their study of debt sustainability in European Union countries or one could use just the federal government budget. We used both for our analysis, but report only the results for federal government budget perspective below.[6]

The primary balance series was obtained by dividing minus 100 times the sum of the net federal saving series (FGDEF) and federal government interest payments (A091RC1Q027SBEA) by the nominal GDP series (GDP).[7] For the debt variable, we used the total public debt series (GFDEBTN) and divided it by the nominal GDP (GDP) series.[8] For the output gap, we computed the difference between the observed annual growth rate and the average annual growth rate. In particular, we computed

[6] We choose the federal government data because the annual data we used in a robustness investigation, not shown for the sake of brevity, was more consistent to the quarterly data used in this section. In other words, the annual series for the total federal, state and local government budgets had quite different numerical values from the quarterly data series making the comparison between the two data series less clear.

[7] Note, this computation for the primary balance results in the proper calculation. The reason that interest payments are added to the federal savings series (before multiplying by a negative number) is that the federal saving series is the difference between current receipts and current expenditures, where current expenditures includes, among other things, interest payments. By adding the interest payment series, we cancel the interest payment portion of the current expenditures.

[8] The Federal Reserve Bank of St. Louis only has quarterly debt data from 1966:01 to the present, so to get the additional 11 years back to 1955:01, we used data from the US treasury which can be found at http://www.treasurydirect.gov/govt/reports/pd/mspd/mspd.htm.

the growth rate in percentage terms by multiplying 100 times the log difference between the current value of real GDP (GDPC1) and the value four quarters earlier. Next these growth rates were averaged and then the average was subtracted from the annual growth rate series to give a series that has positive values when the current growth rate exceeds the average and negative values when the growth rate is below the average.[9]

Our main empirical studies use the data interval from the first quarter of 1955 to the third quarter of 2013. However, in the further investigations, we investigate a restricted data interval with the end date of the second quarter of 1995 to conform with Bohn (1998) and investigate comparability to his results and to show the importance of the switching regression structures which are needed over the full sample.

Table 1 summarizes the results for the various empirical models. The first column describes various parameter possibilities that the different models may take as well as a row for the threshold value for the TR models. As is common in such empirical results tables, many of the parameters are specific to particular models that were estimated. Also, to save space, results for the constant terms are not reported.[10] We use the "AT" and "BT" shorthand when describing TR model estimates to indicate parameter estimates when the threshold variable is above the threshold and below the threshold, respectively. We also list the parameters for the MS models in the same rows as the TR models even though the models, and thus the model parameters, have quite different interpretations. We made this choice to save space in reporting our results, but would like to take a moment to clearly spell out the different interpretations for the parameters to avoid confusion. Table 1 shows that

[9] In a preliminary draft of the paper, we defined the output gap as the cyclical component of real GDP obtained with the Hodrick-Prescott (HP) filter. The main qualitative conclusions about fiscal asymmetries in regard to both sustainability and cyclicality are robust to this alternative definition of the output gap. However, we choose the output gap definition described above over the output gap based on the HP filter for two main reasons. First, the use of the HP output gap introduces an endogeneity issue in the estimation of our empirical fiscal reaction function since the HP filter is a two-sided filter. Second, the interpretation of the empirical fiscal equation as a fiscal reaction function become dubious when the HP output gap is a regressor since the HP output gap uses, by construction, much more information on GDP than the one available for the policy maker at the time of fiscal policy implementation.

[10] Complete tables with all the constant terms can be obtained upon request.

we grouped the above threshold parameters for the TR models with the state 1, s_1, parameters for the MS models and the below threshold parameters for the TR models with the state 2, s_2, parameters for the MS models. Although there is some logic to grouping these parameters as we have, there are also important differences. For the TR models, the AT parameters are the values for the linear portion of the model when the output gap is *observed* to be above the threshold, while for the MS model, the s_1 parameters are the values for the linear portion of the model when the *unobserved* state variable is in regime 1, which is shown below to be mostly associated with recent expansionary periods. Similarly, for the TR models, the BT parameters are the values for the linear portion of the model when the output gap is *observed* to be below the threshold, while for the MS model, the s_2 parameters are the values for the linear portion of the model when the *unobserved* state variable is in regime 2, which is somewhat related to economic downturns.

The second column presents the results of the basic regression model given by (1), which was run using OLS. This model has no threshold, so the first row is left blank as are many of the later rows which correspond to parameters from alternative models. This regression finds the lagged dependent variable is somewhat less than 1 and is highly significant, showing evidence of a highly persistent fiscal rule, the lagged debt coefficient is negative and insignificant and the lagged output gap coefficient is negative and significant. Even though it is insignificant, the negative sign on the lagged debt coefficient is useful to note because, as emphasized by Bohn (1998, 2007), Collignon (2012) and others, a negative value can be viewed as being consistent with a sustainable fiscal policy. In particular, high values of debt imply lower primary deficit values, which result in a decrease in the debt level the following period. Next note that the negative coefficient on the lagged output gap indicates that fiscal policy is countercyclical. Toward the bottom of the column are various measures of fit, which are used for comparison purposes with the alternative models.

The remaining columns of Table 1 present various nonlinear models in which the dependent variable is a nonlinear function of the various economic variables.[11] Part of the motivation for these models is to explore asymmetric policy and, furthermore, to see how fiscal policy sustainability results are impacted when the model is flexible enough to handle asymmetric policy responses. However, these models can also be motivated by statistical testing. One type of statistical test is a procedure suggested by Teräsvirta (1994). For our application, the test begins by first running the basic LM model (1) and obtaining the residuals, which we denote by e_t. Next an unrestricted regression given by

$$\begin{aligned} e_t &= \alpha' + \beta_1' b_{t-1} + \beta_2' d_{t-1} + \beta_3' w_{t-1} \\ &\quad + \gamma_{11}\left(b_{t-1} w_{t-1}\right) + \gamma_{12}\left(d_{t-1} w_{t-1}\right) + \gamma_{13}\left(w_{t-1} w_{t-1}\right) \\ &\quad + \gamma_{21}\left(b_{t-1} w_{t-1}^2\right) + \gamma_{22}\left(d_{t-1} w_{t-1}^2\right) + \gamma_{23}\left(w_{t-1} w_{t-1}^2\right) \\ &\quad + \gamma_{31}\left(b_{t-1} w_{t-1}^3\right) + \gamma_{32}\left(d_{t-1} w_{t-1}^3\right) + \gamma_{33}\left(w_{t-1} w_{t-1}^3\right) + \varepsilon_t', \end{aligned}$$

is run. Testing for the presence of nonlinearity amounts to testing whether the added terms are significant. In particular, this amounts to testing

$$H_0 : \left\{\gamma_{ij} = 0 \text{ for } i = 1,2,3 \text{ and } j = 1,2,3\right\}.$$

We implemented this and found an F-statistic value of 2.831. The 5% critical value for this statistic can be found in a standard table of the F distribution with 9 degrees of freedom in the numerator and infinity in the denominator and is equal to 1.880. We see that linearity is rejected and conclude there is nonlinearity.

Since nonlinear models are appropriate for exploration, consider the estimates for the TR model presented in column 3.[12] This model uses the lagged output gap as the threshold variable and is analogous to the Balassone et al. (2010) model in that the threshold value is arbitrarily set to 0, hence the TR-0 notation. Estimates for the coefficients are mostly consistent with the LM regression. To be more specific,

[11] It is useful to emphasize that although the TR models are linear in the interaction terms denoted in equation (2), they are nonlinear in the economic variables.

[12] In our preliminary empirical analysis, we also investigated logistic smooth TR and exponential smooth TR type nonlinear models. These models did not fit as well as the MS models, so they have been left out of the final draft of the paper.

looking down column 3, we see that when the economy was both above and below potential output in the previous period (i.e. both above and below the threshold value of zero), the lagged balance coefficients are highly significant and close to one. We see that the lagged debt coefficient is negative and insignificant when the lagged output was above potential and positive and insignificant when the economy was below potential. Furthermore, the lagged output gap coefficients are negative when the economy was both above and below potential in the previous period, but only the below potential coefficient is significant.

The coefficients on the lagged debt and output gap variables provide insight into how fiscal policy makers behave. In particular, the negative lagged debt coefficient during above threshold values for the lagged output gap shows a sustainability feature to fiscal policy when the economy is doing well, while the positive lagged debt coefficent during below threshold values could be interpreted as showing that policy makers are more concerned with getting the economy back on track than fiscal policy sustainability during downturns. In other words, the positive coefficient indicates that during times of distress policy makers temporarily forgo fiscal policy sustainability (i.e. a long-term goal) in favor of trying to revive the economy back to a situation with higher-than-average output growth rates (i.e. a short-term priority), at which time they then follow a sustainable fiscal policy program. To some extent, this interpretation is consistent with ideas developed in Collignon (2012), where he noted that various European stability agreements allow for temporary deviations from various sustainability rules. However, we should qualify this conclusion by noting that the insignificant value for both debt coefficients does weaken this interpretation. A similar asymmetric contingent interpretation can be made for the lagged output gap coefficients. In particular, the larger (in absolute value) and significant coefficient during the below threshold case shows that policy makers are more focused on getting the economy on track during downturns than during good times where the coefficient is somewhat smaller (in absolute value) and is insignificant. This interpretation is consistent with the asymmetric fiscal policy result found in Balassone et al. (2010)

for a sample of fourteen European Union countries and shows a strong countercyclical response when economic conditions are poor. The various fit parameters toward the bottom of the table show that the TR-0 model fits better than the simple LM model and the null hypothesis that there is no improvement in the fit, as indicated by the row labeled $F - Stat$, shows an F-statistic of 2.513, which is greater than the 5% critical value of 2.371, and thus can be rejected.

The fourth column of Table 1 presents the estimates for a TR model in which the threshold is endogenously chosen so as to obtain the best fit. We distinguish it by using the TR-E notation. The best fitting threshold for this model occurs at a lagged output gap value of -1.870 which is somewhat lower than the zero threshold used in the TR-0 model. The parameter estimates for this model are largely the same as those for the zero threshold model and can be interpreted as showing the same countercyclical behavior for fiscal policy makers. In particular, the lagged debt coefficients provide some weak evidence for fiscal policy sustainability during good times, but show that during bad times concerns toward improving the economy overwhelm sustainability concerns. Similarly, the negative coefficients on the lagged output gap show that fiscal policy is countercyclical. Furthermore, this policy is asymmetric in that there is a greater level of intervention during economic downturns. One difference between the endogenous TR model and the zero threshold TR model is that two of the coefficients which were insignificant in the zero threshold TR model are significantly different from zero in the endogenous TR. Finally, looking at the various measures of fit at the bottom of the fourth column we see that the endogenous threshold TR model fits somewhat better than the zero threshold TR model. The test that this model fits no better than the LM model, as indicated by the row labeled $F - Stat$, has an F-statistic of 4.483, which is larger than the 5% critical value of 4.122, and thus is easily rejected.[13]

[13] The critical value of 4.122 does not come from a conventional F distribution table. We computed this number by using a bootstrap simulation procedure described in Hansen (1997) which showed that the F-statistic in TR models do not have F distributions and that proper critical values can be found using a bootstrap procedure.

The next two columns of Table 1 show the two-state MS model formulation parameter estimates. The fifth column shows the parameter estimates for the two-state MS formulation in which the switching probabilities are constant, hence the added notation below the MS notation, while the sixth column is a two-state MS model in which the transition probabilities are time varying, hence the notation TVTP. Before describing the estimation results, it is useful to study Figure 1 in order to understand how to interpret the two-state conditions. Figure 1 shows the smoothed state 1, s_1, probability for the constant probability MS model.[14] Focusing on the boom economic period of the 1990s, one can see that the probability of being in state 1 is very high. Similarly, during the boom period of the middle 2000s, Figure 1 shows a high probability of being in state 1. These high probabilities for state 1 during good economic times show that we can say that state 1 is generally associated with good economic times and correspondingly, state 2 is generally associated with poor economic times. Moreover, state 1 is also strongly associated with the Great Moderation period, which is characterized by low volatility.[15]

[14] The smoothed probabilities are computed using the information over the whole sample of size T (i.e., $prob[s_t = 1|I_T]$) as discussed in Hamilton (1994, p. 694). The plot for the MS model in which the transition probabilities are time varying is very similar. Moreover, Figure 1 and Figure 2 below show the smoothed probabilities along with various business cycle turning points which have been dated by the NBER.

[15] The precise dating for the Great Moderation is the subject of debate, but many studies, such as Stock and Watson (2002), have suggested this low volatility period starts around 1984. Furthermore, although the debate is still open, the end of this low volatility period likely ended with the financial crisis in late 2007.

Table 1: Federal Government Primary balance (1955:1 - 2013:3)

	LM	TR-0	TR-E	MS const prob.	MS TVTP
Threshold		0	-1.870		
Lagged Balance	0.924**				
	(0.021)				
Lagged Debt	-0.001				
	(0.003)				
Lagged Output Gap	-0.104**				
	(0.018)				
Lagged Balance - AT or s_1		0.937**	0.942**	0.936**	0.936**
		(0.036)	(0.022)	(0.011)	(0.011)
Lagged Balance - BT or s_2		0.877**	0.779**	0.861**	0.863**
		(0.030)	(0.052)	(0.040)	(0.040)
Lagged Debt - AT or s_1		-0.003	-0.003	-0.012**	-0.012**
		(0.005)	(0.003)	(0.001)	(0.001)
Lagged Debt - BT or s_2		0.005	0.018**	0.006	0.006
		(0.004)	(0.006)	(0.006)	(0.006)
Lagged Output Gap - AT or s_1		-0.023	-0.061**	0.015*	0.015*
		(0.045)	(0.030)	(0.008)	(0.009)
Laggad Output Gap - BT or s_2		-0.177**	-0.260**	-0.173**	-0.173**
		(0.036)	(0.082)	(0.037)	(0.037)
1st-regime volatility $\sigma(s_1)$				0.129**	0.129**
				(0.016)	(0.016)
2nd-regime volatility $\sigma(s_2)$				0.760**	0.759**
				(0.079)	(0.079)
p_{11} or θ_{10}				0.982**	4.053**
				(0.014)	(0.763)
p_{22} or θ_{20}				0.884**	2.028**
				(0.016)	(0.158)
θ_{11}					-0.107
					(0.154)
θ_{21}					-0.101*
					(0.060)
RSS	90.383	86.550	83.765		
AIC	1066.453	1064.269	1056.585		
SBC	1080.292	1091.946	1084.262		
log-likelihood				-126.676	-126.110
$F-Stat$		2.513	4.483		

Notes: Standard errors in parentheses. One asterisk indicates significance at the 10% level, while two asterisks indicate significant at the 5% level.

Now focusing on the parameter estimates we see that, like the two TR models, the two MS models have very similar estimation results to each other and even to the TR models. Both show that the lagged balance coefficients are near 1 and significant in both states, the response to lagged debt is asymmetric, showing a negative and significant coefficient during state s_1 but an insignificant coefficient during state s_2, and the response to the lagged output gap is asymmetric, showing a negative and significant coefficient during state s_2 but a positive and marginally significant coefficient during state s_1. Moreover, the volatility of innovations in state 1, $\sigma(s_1)$, is roughly five times lower than the volatility of innovations in state 2, $\sigma(s_2)$. It is important to keep in mind this changing volatility feature uncovered by the MS formulation when comparing the estimation results obtained from the two approaches because the TR formulation assumes, in contrast to the MS formulation, an identical level of innovation volatility both above and below the threshold level.

Since state s_1 is generally associated with good economic times and s_2 is generally associated with poor economic times, the interpretations for these coefficients are largely the same as described for the TR models. In particular, during good economic times they show sustainability of fiscal policy and less concern with countercyclical policy, while during weak economic times they show no concern with sustainability of fiscal policy and a countercyclical fiscal policy. However, it is useful to point out that the fiscal policy sustainability result during good times is considerably stronger with the MS models than with the TR models as the good economic time coefficients are larger (in absolute value) and highly significant. Put differently, one can say that the MS models show stronger evidence that during the good economic times, which feature low volatility, policy makers are focused on fiscal policy sustainability.

Finally let us comment a bit on the parameters for the transition probabilities. Focusing on the constant probability model, we see that probability of staying in state 1 if one begins in state 1, p_{11}, is very high at 0.982, while the probability of staying in state 2 if one begins in state 2, p_{22}, is smaller at 0.884. These estimates show that there is high persistence for both states, with the persistence during good times being larger. As noted above, the high persistence of the fiscal sustainablity state given by a value of p_{11} that is close to one reduces the likelihood of a situation described by Francq and Zakoïan (2001) and increases the likelihood that there is global fiscal policy sustainability. To understand the results for the time varying transition probability MS model, it is sufficient to look at the two parameters associated with the lagged output gap, θ_{11} and θ_{21}. Both are insignificant, which means that the transition probabilities for these models are largely the same as the constant probability MS model. This result is consistent with the fact that the parameter estimates for the two MS models are very close.

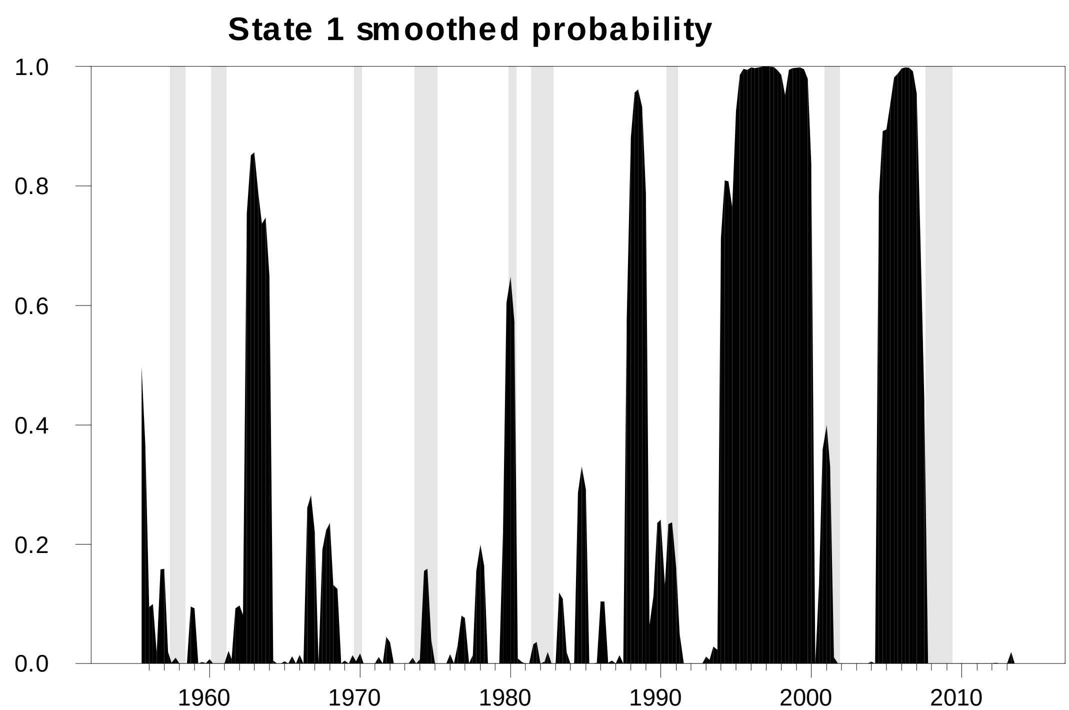

Figure 1. Two-State MS model

Let us conclude this section with a brief summary of the key findings. In general, we found that there was a sequence of improvements in the fit and performance of the models presented. While the linear model exhibited desirable properties, such as debt sustainability and countercyclical behavior, it did not fit as well as both versions of the TR models which allowed asymmetric behavior on the part of policy makers. However the TR models showed their own weaknesses. For instance, they did not show debt sustainability during weak economic times. This behavior is interpreted as reasonable as it shows that during weak economic times policy makers are more concerned with getting the economy back on track. The MS models also showed some debt sustainability issues during weak economic times, but the coefficients were smaller than the TR models and insignificant. In addition, the good economic times showed larger (in absolute value) lagged debt coefficients indicating a stronger response towards fiscal policy sustainability.

4 An alternative sample period

In this section we undertake an investigation which more directly compares our results with those in Bohn (1998). Using annual US data from 1916 to 1995, Bohn (1998) argued that fiscal policy is consistent with sustainability in a simple static linear regression model. Here, we show that a similar, but dynamic, linear regression methodology as the one used by Bohn over this shorter sample is appropriate, but we also argue that when using the longer sample used in Section 3, the nonlinear models are more appropriate. Before discussing these results, let us first describe the data.

Since our quarterly data set begins in 1955:1, we cannot match the annual data starting date used in Bohn (1998), but we can match the end date of 1995 which we do with 1995:2 for our sample.[16] For ease of exposition, we will often refer to this data interval as the short sample while the longer data period from Section 3 will be referred to as the full sample.

Table 2 displays the results for the same models as in Table 1 using this short

[16] Moreover, Bohn (1998) also studied several subsamples, finding similar results to those found in his reference sample. One of his sample robustness exercises covered the period 1948-1995, which is rather similar to the sample period analyzed in this section.

sample data. Table 2 shows many consistencies with the results from Section 3, but it also shows a few differences. One key difference between the short and long samples concerns the lagged debt parameter. The short sample shows a somewhat symmetric and sustainable fiscal policy. Indeed, the lagged debt coefficient in the LM is negative, significant and close to the corresponding estimated values under the two regimes associated with the TR and MS models. This contrasts with the long sample results in Section 3, which only showed the negative debt sustainability coefficients during above threshold or state 1 periods. These short sample results are in line with the results in Bohn (1998) who only used a simple linear model and show that fiscal policy does exhibit fiscal policy sustainability for this interval of time. It is also useful to point out that one could interpret, the difference between these short and long sample results as showing that the addition of 18 years of data makes the question of fiscal policy sustainability less clear. This may in part account for the growing concern for the size of the US government debt and deficits seen in more recent political debates.

The results for the lagged output gap coefficient are largely the same between the two sample periods for both the TR and MS models. In particular, we see an asymmetric response to the lagged output gap in both samples with stronger countercyclical policy applied during economic weaknesses. Furthermore, as with the long sample, the short sample shows some differences between the TR models and the MS models, in that the TR models show countercyclical policy both above and below the threshold, while the MS models only show countercyclical policy during state 2.

Table 2: Federal Government Primary balance (1955:1 - 1995:2)

	LM	TR-0	TR-E	MS const prob.	MS TVTP
Threshold		0	-2.410		
Lagged Balance	0.796**				
	(0.041)				
Lagged debt	-0.010**				
	(0.004)				
Lagged Output Gap	-0.081**				
	(0.019)				
Lagged Balance - AT or s_1		0.828**	0.824**	0.935**	0.934**
		(0.059)	(0.044)	(0.013)	(0.013)
Lagged Balance - BT or s_2		0.754**	0.565**	0.715**	0.716**
		(0.057)	(0.097)	(0.084)	(0.084)
Lagged debt - AT or s_1		-0.007	-0.009*	-0.009**	-0.009**
		(0.006)	(0.005)	(0.001)	(0.001)
Lagged debt - BT or s_2		-0.011*	-0.015	-0.011*	-0.011*
		(0.006)	(0.010)	(0.006)	(0.006)
Lagged Output Gap - AT or s_1		-0.035	-0.047*	0.005	0.004
		(0.044)	(0.028)	(0.007)	(0.007)
Lagged Output Gap - BT or s_2		-0.141**	-0.386**	-0.132**	-0.132**
		(0.043)	(0.116)	(0.034)	(0.034)
1st-regime volatility $\sigma(s_1)$				0.080**	0.081**
				(0.009)	(0.009)
2nd-regime volatility $\sigma(s_2)$				0.711**	0.712**
				(0.091)	(0.090)
p_{11} or θ_{10}				0.999**	15.337**
				(0.002)	(0.828)
p_{22} or θ_{20}				0.870**	1.900**
				(0.016)	(0.141)
θ_{11}					0.241
					(0.561)
θ_{21}					-0.098
					(0.063)
RSS	53.769	52.616	50.185		
AIC	665.519	658.009	650.346		
SBC	665.870	682.710	675.047		
log-likelihood				-72.980	-72.498
$F-Stat$		0.843	2.749		

Finally, we consider whether the nonlinear models are important for this data interval. First consider the nonlinear model test suggested by Teräsvirta (1994) which was described earlier. Running the same test as before, we found an F-

statistic of 1.653 which is smaller than the 5% critical value of 1.880. Thus we reject the null that there is a nonlinearity in the data which is in contrast to the test over the full sample. Second, consider the tests that the threshold TR models fits no better than the LM model. As indicated by the row labeled $F - Stat$ in Table 2, the F-statistic of 0.843 for the zero threshold TR model is lower than the critical value of 2.371 from the regular F-distribution and the F-statistic of 2.749 for the endogenous threshold TR model is lower than the critical value of 4.147 computed using Hansen (1997) bootstrap procedure for the short sample. These tests show that the LM is not rejected for the short sample which is in contrast to the test over the full sample. Next consider the probability diagram associated with the constant probability MS model. Figure 2 displays the analogous short sample diagram to the long sample diagram given in Figure 1. This figure shows that the probability of being in state 1 is very low throughout the sample period. Interestingly, toward the end of the short sample, we do see a rising probability of being in state 1 and this period corresponds with some of the data periods in the long sample that also have a high probability of being in state 1. However, overall, state 1 does not appear to be useful for fitting the model and this indicates that the simple linear model similar to the one used by Bohn (1998) is sufficient to fit the data.

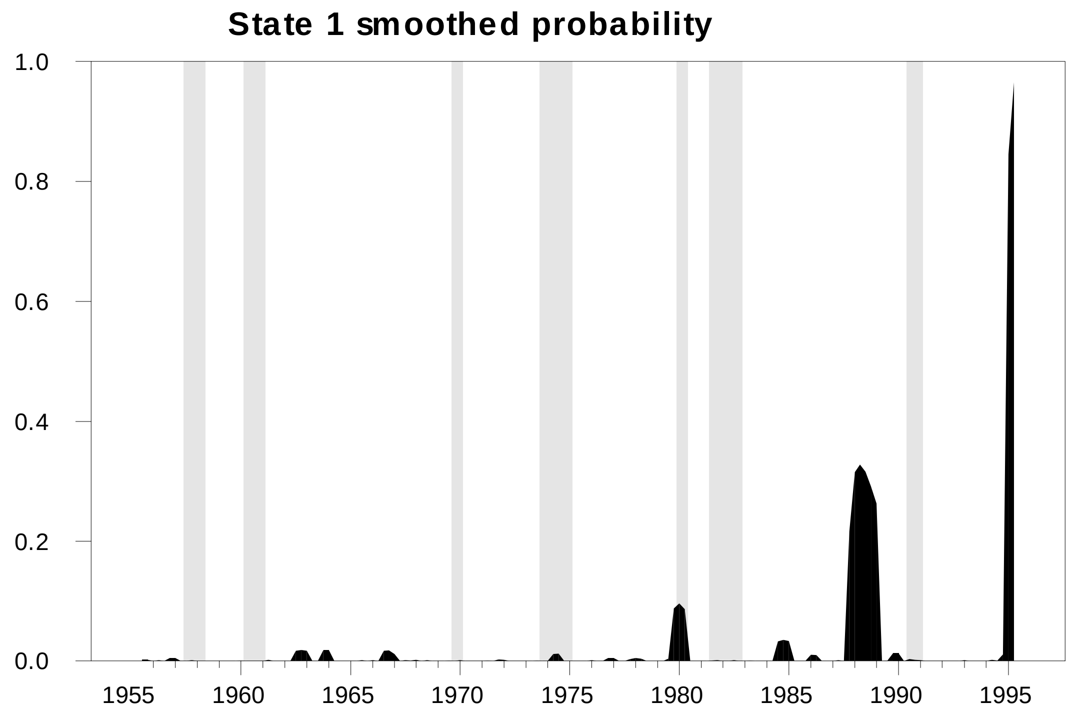

Figure 2. Two-State MS model (Sample 1955-1995)

5 Conclusion

This paper has found several important results regarding fiscal policy modelling and the empirical analysis of fiscal policy for the United States. First, and perhaps most importantly, we showed that recent economic data, that include the financial crisis and Great Recession, are sufficiently different from earlier data that simple linear models are not appropriate for modelling fiscal policy.

Regarding our policy findings, we found that over a sample that includes the financial crisis and the Great Recession, the fiscal responses to lagged debt and lagged output gap are asymmetric, however, over the shorter sample, which ended in the second quarter of 1995, the results for lagged debt did not exhibit asymmetric behavior. Regarding fiscal policy sustainability, we also find that the sample period used is important. For the short sample, fiscal policy is sustainable, confirming results by Bohn (1998), while over the full sample period, fiscal policy is only sustainable during good economic times. While fiscal policy does not appear to be sustainable during times of distress over the full sample, we interpret this as evidence that policy makers are more concerned about getting the economy back on track and perhaps temporarily ignore sustainability. Overall these empirical findings suggests that the fiscal sustainability question does become less clear when using data that includes the recent financial crisis and the Great Recession.

Regarding the cyclicality of fiscal policy, our analysis found robust evidence that fiscal policy is countercyclical during bad economic times and becomes less countercyclical during good times. A few of our models even indicated that during good economic times fiscal policy may become marginally procyclical. These latter findings are consistent with results found in Balassone et al. (2010) using European data and are interpreted to show that during good times, the government budget deficit-GDP ratio grows with the rest of the economy.

References

[1] Afonso, A., Agnello, L., and Furceri, D. 2010. Fiscal policy responsiveness, persistence, and discretion. *Public Choice* 145, 503-530.

[2] Balassone, F., Francese, M., and Zotteri, S. 2010. Cyclical asymmetry in fiscal variables in the EU. *Empirica* 37, 381-402.

[3] Bohn, H. 1998. The behavior of US public debt and deficits. *The Quarterly Journal of Economics* 113, 949-963.

[4] Bohn, H. 2007. Are stationarity and cointegration restrictions really necessary for the intertemporal budget constraint? *Journal of Monetary Economics* 54, 1837-1847.

[5] Collignon, S. 2012. Fiscal policy rules and the sustainability of public debt in Europe. *International Economic Review* 53, 539-567.

[6] Davig, T. 2005. Periodically expanding discounted debt: A threat to fiscal policy sustainability? *Journal of Applied Econometrics* 20, 829-840.

[7] Golinelli, R., and Momigliano, S. 2008. The cyclical response of fiscal policies in the Euro area. Why results of empirical research differ so strongly? Banca d'Italia Temi di Discussione 654.

[8] Enders, W. 2010. *Applied Econometrics Time Series.* Wiley (3rd Edition).

[9] Filardo, A.J. 1994. Business-cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics* 12, 299-308.

[10] Francq, C., and Zakoïan, J.-M. 2001. Stationarity of multivariate Markov-switching ARMA models. *Journal of Econometrics* 102, 339-364.

[11] Hakkio, C.S. and Rush, M. 1991. Is the budget deficit too large? *Economic Inquiry* 24, 429-445.

[12] Hamilton, J.D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357-384.

[13] Hamilton, J.D. 1994. *Time Series Analysis.* New Jersey: Princeton University Press.

[14] Hamilton, J.D., and Flavin, M.A. 1986. On the limitations of government borrowing: A framework for empirical testing. *American Economic Review* 76, 808-819.

[15] Hansen, B. 1997. Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics* 2, 1-14.

[16] Kremers, J.M. 1989. US Federal indebtedness and the conduct of fiscal policy. *Journal of Monetary Economics* 23, 219-238.

[17] Quintos, C.E. 1995. Sustainability of the deficit process with structural shifts. *Journal of Business and Economic Statistics* 13, 409-417.

[18] Perron, P. 1989. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57, 1361-1401.

[19] Stock, J., and Watson, M.W. 2002. Has the business cycle changed and why? in Gertler, M., and K.S. Rogoff (eds.) *NBER Macroeconomics Annual*, Cambridge, MA: MIT Press, pp. 159–218.

[20] Teräsvirta, T. 1994. Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* 89, 208-218.

[21] Trehan, B., and Walsh, C. 1988. Common trends, the government budget constraint, and revenue smoothing. *Journal of Economic Dynamics and Control* 12, 425-444.