mathematics

MDPI

Article

A Commentary on US Sovereign Debt Persistence and Nonlinear Fiscal Adjustment

Vladimir Andric *, Dusko Bodroza and Mihajlo Djukic

Institute of Economic Sciences, 11000 Belgrade, Serbia; dusko.bodroza@ien.bg.ac.rs (D.B.); mihajlo.djukic@ien.bg.ac.rs (M.D.)

* Correspondence: vladimir.andric@ien.bg.ac.rs; Tel.: +381-638268239

Abstract: The purpose of this paper is to show how the self-exciting threshold autoregressive (SETAR) model might be a suitable econometric framework for characterizing the dynamics of the US public debt/GDP ratio after the Bretton Woods collapse. Our preferred SETAR specifications are capable of capturing the main stylized facts of the US public debt/GDP ratio between 1974 and 2024. In addition, the estimated SETAR models are consistent with theoretical frameworks that look to explain the behavior of the US public debt/GDP ratio before and after the Global Financial Crisis (GFC). Finally, under the assumption of public debt/GDP ratio stationarity, for which we find only limited and inconclusive evidence, this paper provides some arguments for why previous studies, which use the exponential smooth threshold autoregressive (ESTAR) models, logistic smooth threshold autoregressive (LSTAR) models or SETAR-type models for the first differences of the US public debt/GDP ratio, are potentially misspecified on both econometric and economic grounds.

Keywords: SETAR model; United States; sovereign debt; persistence; nonlinear fiscal adjustment

MSC: 62M10; 91B84

Citation: Andric, V.; Bodroza, D.; Djukic, M. A Commentary on US Sovereign Debt Persistence and Nonlinear Fiscal Adjustment. *Mathematics* **2024**, *12*, 3250. https://doi.org/10.3390/math12203250

Academic Editor: María del Carmen Valls Martínez

Received: 31 August 2024
Revised: 11 October 2024
Accepted: 11 October 2024
Published: 17 October 2024

1. Introduction

In a comprehensive review on the use of threshold autoregressions (TARs) and self-exciting threshold autoregressions (SETARs) in economics, Hansen (2011) [1] provides an overview of seventy-five papers that employ the (SE)TAR econometric framework in modeling and forecasting output growth, interest rates, prices, stock returns and exchange rates. This paper shows how the SETAR model might be useful in capturing asymmetries in the dynamics of the public debt/GDP ratio. We focus on modeling nonlinearities in the US public debt/GDP ratio after the Bretton Woods collapse. Not only are our estimated two-regime and three-regime SETAR specifications capable of capturing stylized facts concerning the US public debt/GDP ratio between 1974 Q1 and 2024 Q1, they are also consistent with several theoretical predictions regarding the behavior of the US public debt, starting, most notably, from the seminal paper by Barro (1979) [2]. Finally, we also show how earlier contributions that use the exponential smooth transition autoregression (ESTAR) models, logistic smooth threshold autoregression (LSTAR) models and SETAR models for the first differenced US public debt/GDP ratio might be potentially misspecified on both econometric and economic grounds.

To familiarize the reader with the earlier literature on public debt behavior in the case of the US, we organize the rest of this introductory section by reviewing the literature on the potential non-stationarity of the US public debt/GDP ratio first and then discuss papers that model the nonlinearities in the dynamics of the US public debt. As Bec et al. (2004) [3] caution, establishing the stationarity of the time series in question prior to estimating the appropriate SETAR model is essential, since the underlying econometric methods developed by Hansen (1996, 1997, 2017) [4–6], González and Gonzalo (1997) [7] and Gonzalo and Pitarakis (2002) [8] crucially rest on the ergodicity and global stationarity assumptions for the stochastic process under investigation.

1.1. Persistence

Barro (1979) [2] was the first to claim that there are no underlying economic forces that would cause the public debt/GDP ratio to converge to a steady-state target value. In other words, in Barro's (1979) [2] tax smoothing model, the US public debt behaves like a random walk after World War I. The public debt/GDP ratio shows unpredictable movements governed by only transitory government spending (mostly during wars) and countercyclical output shocks (mostly during recessions). There is also no effect of both unanticipated and expected (anticipated) inflation on the public debt/GDP ratio. The stated results do not change, regardless of whether one measures public debt at nominal (par) or market values.

Hamilton and Flavin (1986) [9] refute Barro's (1979) [2] conclusion that the US public debt/GDP ratio shows random walk-type behavior, although for a much shorter period spanning from 1960 to 1984. By applying the standard Dickey–Fuller unit root test of Dickey and Fuller (1981) [10], Hamilton and Flavin (1986) [9] reject the unit root non-stationarity hypothesis for the US public debt/GDP ratio at the 10% significance level.

Kremers (1988) [11], however, shows that one cannot reject the non-stationarity of the US public debt/GDP ratio in the post-World War II data. Contrary to Hamilton and Flavin (1986) [9], Kremers (1988) [11] implements an augmented Dickey–Fuller unit root test to appropriately model the autocorrelation present in the residual values of the US public debt/GDP ratio and consequently overturns the results of Hamilton and Flavin (1986) [9] by not being able to reject the non-stationarity hypothesis at any critical level of up to 90%. In addition, Kremers (1989) [12] further shows that even for the combined inter- and post-war period, one cannot firmly reject the non-stationarity hypothesis in the case of the US public debt/GDP ratio.

Wilcox (1989) [13] argues that the measure of the US public indebtedness that Hamilton and Flavin (1986) [9] use is inappropriate since it refers to the undiscounted public debt. Contrary to Hamilton and Flavin (1986) [9], Wilcox (1989) [13] uses stochastic real interest rates to compute the discounted present value of the US public debt at a particular point in time. Wilcox (1989) [13] uses the discounted value of the US government debt to define a public debt sustainability criterion, which says that overall fiscal policy is sustainable if the projected discounted value of the public debt ratio approaches zero, i.e., if the expected present value of the sum of future primary surpluses equals the current market value of the US public debt. As in Hamilton and Flavin (1986) [9], Wilcox (1989) [13] operationalizes his sustainability criterion by comparing the current market value of the public debt with the sum of the expected discounted primary surpluses and denotes the difference between the two as $A_t = \lim_{n\to\infty} \rho^n E_t\big(B_{t+n}\big)$ where B_{t+n} is the projected infinite-horizon market value of public debt and $\rho = 1/(1+r_t)$ is the stochastic discount factor that is inversely related to the stochastic time-varying market interest rate r_t measured in real terms. Wilcox (1989) [13] further argues that the behavior of A_t is influenced by the behavior of the discounted public debt value B_t—if B_t is non-stationary, then A_t is stochastic (martingale), and if B_t is stationary, then A_t is constant (possibly zero)—see Wilcox (1989) [13] (p. 296) for further development of this argument. The conclusion of Wilcox (1989) [13] is that for the period after 1974, the discounted market value of the US public debt is non-stationary.

Given the inconclusive evidence of earlier unit root studies in assessing the sustainability of the US public debt, Bohn (1998, 2007) [14,15] criticizes unit root-type regressions on two grounds. First, Bohn (1998) [14] argues that unit root test regressions suffer from an omitted variable bias since they do not account for cyclical output changes and transitory government spending. By aiming to explain the variations in the primary fiscal balance as a function of movements in the public debt, output gap and transitory government spending, Bohn (1998) [14] proposes a fiscal reaction function (FRF) regression approach to evaluate the mean reversion in the stochastic process for the US public debt. Using data for the US between 1916 and 1995, Bohn (1998) [14] concludes that the US public debt/GDP ratio behaves as a highly persistent, but overall mean-reverting, stationary stochastic process. Regardless of how interest rates and growth rates compare, a positive response of the primary fiscal balance to public debt movements is a sufficient condition for public debt sustainability since a positive primary fiscal balance response would reverse any upward movement in the public debt/GDP ratio. Second, the sustainability notion of Wilcox (1989) [13], $A_t = \lim_{n\to\infty} \rho^n E_t\left(B_{t+n}\right) \to 0$, is always satisfied, since the exponential growth in the denominator, $1/(1+r_t)^n$, of the expression for the real discount factor ρ^n asymptotically dominates the m-th order polynomial in the numerator, B_{t+n}, irrespective of the order of integration for B_{t+n}—see Proposition 1 of Bohn (2007) [15] (p. 1840) for a detailed proof.

Contrary to Bohn (1998) [14], who estimates a single equation ordinary least squares (OLS) FRF, Cochrane (2020, 2022) [16,17] estimates a vector autoregressive (VAR) model with the public debt and primary fiscal surplus and finds a 0.98 value for the first lag debt coefficient. In other words, Cochrane (2020, 2022) [16,17] reaffirms the findings of Bohn (1998) [14] that the public debt/GDP ratio is a stationary, but highly persistent, near-unit root stochastic process. The claims of Cochrane (2020, 2022) [16,17] are based on a positive, statistically significant, response of the primary fiscal balance to changes in public debt/GDP, which ensures a mean reversion in the stochastic process for the public debt/GDP ratio.

On the other hand, Campbell et al. (2023) [18] argue that the US public debt/GDP ratio after World War II must be non-stationary since it has little ability to predict its own dynamics, as well as future fiscal developments in taxes and spending. Campbell et al. (2023) [18] instead propose a stationary government surplus/debt ratio as a useful predictor of future fiscal outcomes. Campbell et al. (2023) [18] use the relationship between surplus and debt in the US to show that the US government responded to the shrinking fiscal space between 1947 and 2022 by cutting spending, not by raising taxes.

Finally, although Jiang et al. (2024) [19] find that the US public debt/GDP ratio is persistent, close to a unit root, stochastic process, the authors exclude the possibility that there is an actual unit root in the autoregressive representation for the public debt/GDP ratio on several grounds. First, a non-stationary public debt/GDP would breach any upper bound given an arbitrarily long forecast horizon. Second, a unit root stochastic process would also imply an ever-increasing variance of the public debt/GDP ratio with the passage of time. Third, large increases in the public debt/GDP ratio in US fiscal history have usually led to (i) discretionary fiscal adjustments; (ii) high inflation; (iii) financial repression in the form of interest rate caps on government borrowing; or (iv) corrections to the market prices of government bonds. In sum, Jiang et al. (2024) [19] conclude that the US public debt/GDP ratio shows highly persistent, near-unit root, behavior, but more importantly, the authors contribute such an autocorrelation profile to the 2007 structural break due to the Global Financial Crisis (GFC). However, as Jiang et al. (2024) [19] acknowledge, they impose the break exogenously on the dynamics of the US public debt/GDP ratio in the sense that "…this analysis leaves the large, permanent increase in the D-O ratio (as well as its timing) unexplained" (Jiang et al. (2024) [19] p. 4).

Jiang et al. (2024) [19] use the Chow structural break test to date the break in the US public debt/GDP ratio in 2007. The reader should note that even if the timing of the 2007 structural break had been endogenous, i.e., explained, both in terms of the timing and size, by the underlying forces that govern the dynamics of public debt/GDP ratio, Carrasco (2002) [20] warns that endogenous structural change tests have no power if the data are generated by a nonlinear threshold-type model. Put differently, the nonlinear threshold-type tests for parameter stability have greater power in comparison to tests that deal with structural change in parameters. Consequently, Carrasco (2002) [20] advises that evaluating the null hypothesis of linearity against a threshold alternative is the most robust approach to detecting parameter instability in macroeconomic and financial time series.

The recommendations of Carrasco (2002) [20] regarding the use of nonlinear threshold-type models in economics are crucial from the standpoint of this paper, even more so given the results reported by Gonzáles and Gonzalo (1997) [7] and Lanne and Saikkonen (2002) [21], who caution about the observational equivalence between the actual unit root stochastic processes and respective nonlinear alternatives, especially in relatively small samples. The question is, however, which nonlinear threshold alternative is the most suitable one for describing the dynamics of highly persistent, potentially unit root, stochastic processes such as the one governing the dynamics of the US public debt/GDP ratio after the Bretton Woods collapse.

1.2. Nonlinearities

One of the first contributions that model the nonlinearities in the dynamics of the US public debt/GDP ratio is Sarno (2001) [22]. Sarno (2001) [22] estimates the ESTAR model of the following form:

$$\Delta d_t = \alpha + \rho d_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta d_{t-j} + \left(\alpha^* + \rho^* d_{t-1} + \sum_{j=1}^{p-1} \phi_j^* \Delta d_{t-j}\right)\Phi[\theta; d_{t-k} - c] + \epsilon_t \quad (1)$$

in which Δ stands for the first difference operator, d_t is the ergodic and globally stationary public debt/GDP ratio, α and α^* are regime-dependent level shifts, the residuals are $\epsilon_t \sim iid(0,\ \sigma^2)$, while k is the delay parameter. The transition function between the two regimes takes the form $\Phi[\theta; d_{t-k} - c] = 1 - exp\left[-\theta(d_{t-k} - c)^2\right]$, where θ measures the speed of transition between the two regimes and c denotes the threshold public debt/GDP ratio. The sum of the autoregressive coefficients, $\sum_{j=1}^{p-1} \phi_j$, describes the persistence and the order of autoregression (p), while ρ and ρ^* represent the respective regime-dependent autoregressive slope coefficients. Although it is admissible for $\rho \geq 0$, the global stationarity condition for the described ESTAR model of Sarno (2001) [22] demands that $\rho^* \leq 0$ and $\rho + \rho^* \leq 0$.

As in Bohn (1998) [14], Sarno (2001) [22] estimates Equation (1) on a sample spanning from 1916 to 1995 to discover that the US public debt/GDP ratio behaves as a nonlinear mean-reverting ESTAR stochastic process. There are, however, potential problems with the underlying ESTAR econometric estimates by Sarno (2001) [22].

First, since Equation (1) of Sarno (2001) [22] from above is parameterized and estimated in first differences, and not levels of public debt/GDP ratio, the estimates from (1) might be prone to an omitted variable bias. Equation (1), in essence, represents a nonlinear reaction function of Δd_t on d_{t-1} in which the response of Δd_t to d_{t-1} is regime-specific and determined by the estimated values of k, θ and μ, as well as by the shape of the transition function $\Phi[\theta; d_{t-k} - \mu]$, which, in the case of Sarno (2001) [22], is an exponential transition function. Since Δd_t is equal to the overall fiscal balance corrected for the potential stock-flow adjustments, the ESTAR Equation (1) is a nonlinear FRF of the overall fiscal balance to regime-specific lagged d_{t-k} values. To the extent that Δd_t approximates the dynamics of the US primary fiscal balance, Equation (1), similarly to the unit root test regressions, also does not incorporate transitory government spending and cyclical output shocks on its right-hand side. More importantly, Bohn (1998) [14] explicitly says that Δd_t is a function of both lagged public debt/GDP and non-debt components, most notably the output gap and transitory government spending. Equation (4) from Bohn (1998) [14] reads as follows:

$$\Delta d_t = d_t - d_{t-1} = -[1 - x_t(1-\rho)]d_{t-1} - x_t \mu_{t-1} \quad (2)$$

in which $x_t = 1 + r_t - y_t$ holds for the real interest rate r_t and the real growth rate y_t, and where μ_{t-1} represents the lagged output gap and lagged transitory government spending, under the realistic assumption that both variables are strictly bounded stochastic processes. In Table 2 (p. 956), Bohn (1998) [14] provides estimates of Equation (2) from above. In addition, when evaluating a nonlinear response of the primary fiscal balance to changes in the public debt/GDP ratio in Table 3 (p. 958), Bohn (1998) [14] explicitly controls for the variations in the output gap and transitory government spending. Like Bohn (1998) [14], Mendoza and Ostry (2008) [23] and Mauro et al. (2015) [24] quantify the extent of the omitted variable bias that results from neglecting the output gap and transitory government spending in a FRF of the primary budget balance on public debt in a broader international and historical context.

Second, a claim by Sarno (2001) [22] (p. 120) that "there is growing evidence that governments respond more to primary deficits (surpluses) when public debt is particularly high (low)" is a valid empirical fact in the case of the US for the sample period from 1916 to 1995, which both Bohn (1998) [14] and Sarno (2001) [22] use in their respective studies. However, there is a statistically significant structural shift in the primary balance FRF coefficient after the GFC, as D'Erasmo et al. (2015) [25] document in the case of the US for the period 1791–2014. Using the extended sample period that ends in 2014, D'Erasmo et al. (2015) [25] manage to overturn the results originally reported by Bohn (1998) [14]. Due to an unprecedented public debt build up after the 2008 GFC, D'Erasmo et al. (2015) [25] quantify a much lower primary balance FRF coefficient to public debt upward movements. This finding of D'Erasmo et al. (2015) [25] contradicts the statement of Sarno (2001) [22] (p. 121) "...that governments react more strongly to primary deficits when the deviation of the debt/GDP ratio from equilibrium is large in absolute size suggests that the larger the deviation from the long-run equilibrium of the debt/GDP ratio, the stronger will be the tendency to move back to equilibrium".

The reader should note that the highlighted claims of Sarno (2001) [22] and the original estimates of Bohn (1998) [14] might not only be sample-specific, as they are also inconsistent with the theoretical model of rational expectations equilibrium of the sovereign borrower of Ghosh et al. (2013) [26], in which the fiscal behavior of the sovereign borrower follows a reduced form FRF with the characteristics of fiscal fatigue. The FRF with fiscal fatigue characteristics of Ghosh et al. (2013) [26] implies a cubic relationship between the primary fiscal balance and public debt such that at low levels of debt there is no, or even negative, relationship between the primary balance and public debt. With the increase of public debt, the response of the primary balance also increases, but the size of the response eventually weakens and finally decreases at extremely elevated levels of debt. To summarize, it is unlikely that governments can react more aggressively to increased primary deficits when government debt/GDP ratios are particularly high, if only because the primary surplus/GDP ratios cannot exceed 100%, while interest payments and government debt as a % of GDP can.

Third, some novel econometric findings of Heinen et al. (2012) [27] and Buncic (2019) [28] are in contrast with the claims of Sarno (2001) [22] about the desirable properties of the exponential transition function, most notably the properties of its boundedness between 0 and 1 and its symmetrically inverse-bell-shaped transition function around zero. Sarno (2001) [22] claims (p. 120, below Equation (1)) that "these properties are attractive in the present context because they allow symmetric adjustment of d_t for deviations above and below the equilibrium level". Put differently, the symmetric adjustment property of the ESTAR transition function and the property that the exponential transition function increases with absolute deviations of the dependent variable from the estimated threshold imply the inverted bell shape of the exponential transition function. But, as Heinen et al. (2012) [27] show, such properties of the exponential transition function also imply that it might be impossible to uniquely identify the exponential transition function as it has comparable properties to the quadratic transition function. More precisely, Heinen et al. (2012) [27] argue that one cannot uniquely identify the exponential transition function in relation to extreme parameter combinations, which is especially true for small or exceptionally large values of the error term variance, or when certain model parameters tend to their limiting values. The consequence of this identification problem are strongly biased estimators in the case of the ESTAR model specification.

Like Heinen et al. (2012) [27], Buncic (2019) [28] emphasizes an additional identification problem in the case of the ESTAR model, which implies observational equivalence between the exponential transition function and the quadratic transition function in cases when the speed of transition parameter θ takes on small values. On the other hand, for large values of the speed of transition parameter θ, there is an observational equivalence between the exponential transition function and the indicator outlier fitting function. In other words, the exponential transition function acts as a dummy variable that removes the influence of outlier observations at and near the threshold. As the simplest practical alternative to the ESTAR model specification, Buncic (2019) [28] recommends the use of (SE)TAR-type threshold models.

Fourth, as Sarno (2001) [22] claims (p. 120, footnote number 3), an alternative smooth transition function to the exponential one of the ESTAR process is the logistic transition function of the LSTAR model specification. Sarno (2001) [22] opts for an exponential transition function on statistical grounds and further argues that the LSTAR model "seems relatively less appropriate for modeling the dynamics of the public debt/GDP ratio", since it implies the asymmetric behavior of public debt/GDP with respect to the endogenously estimated threshold. Cochrane (2022) [17], however, claims (p. 31) that "the s-shaped surplus/GDP process is a crucial lesson" for the post-World War II US fiscal dynamics. In other words, today's deficits precede future surpluses since the surplus/GDP follows an s-shaped process in a VAR setting with public debt/GDP and surplus/GDP ratios. But even if the statements of Cochrane (2022) [17] about the s-shaped surplus/GDP process are correct, which Campbell et al. (2023) [18] and Jiang et al. (2024) [19] question on the basis of the (near) unit root process for public debt/GDP, the additional problem with the LSTAR model, as Ekner and Nejstgaard (2013) [29] claim (p. 17), is that "a large and imprecise estimate of the speed of transition parameter θ implies that the LSTAR model is effectively a TAR model". Moreover, Gao et al. (2018) [30] further show that the LSTAR model specification also suffers from identification issues since the value of its transition function, $\Phi[\theta; d_{t-k} - c] = 1/(1 + exp[-\theta(d_{t-k} - c)])$, converges to one for large values of the speed of transition parameter, i.e., $\Phi[\theta; d_{t-k} - c] = 1/(1 + exp[-\theta(d_{t-k} - c)]) \to 1$ when $\theta \to \infty$. In other words, the logistic transition function behaves as an indicator function of a discrete (SE)TAR model. The statements by Gao et al. (2018) [30] are also important from the economic perspective since the economic theory rarely (or ever) recommends which value the speed of transition parameter θ should assume. In sum, the recommendations of Buncic (2019) [28], Ekner and Nejstgaard (2013) [29] and Gao et al. (2018) [30] show that the (SE)TAR process has more desirable statistical properties in comparison to the ESTAR and LSTAR processes, respectively.

Gnegne and Jawadi (2013) [31] estimate a two-regime SETAR process for the public debt/GDP ratio in the case of the US between 1970 and 2009. However, similarly to Sarno (2001) [22], Gnegne and Jawadi (2013) [31] model the nonlinear behavior in the changes, not levels, of the public debt/GDP ratio, which effectively implies investigating asymmetries in the stock-flow-adjusted overall fiscal balance. The choice of Gnegne and Jawadi (2013) [31] to focus on changes, instead on levels, of the public debt/GDP ratio is a consequence of a potentially inappropriate choice of respective unit root tests. Gnegne and Jawadi (2013) [31] (p. 158, Table 1) assert the following:

According to Table 1, the great majority of unit root tests indicate that public debt/GDP ratio in the case of US is an I (1) stochastic processes. To check the robustness of our findings for the presence of structural breaks, we further apply a ZA unit root test, but the main conclusion about I (1) behaviour remains unchanged.

Gnegne and Jawadi (2013) [31], hence, use the Zivot–Andrews (ZA) unit root test of Zivot and Andrews (1992) [32] with a single endogenous structural break to strengthen their findings about the I (1) nature of the stochastic process for the US public debt/GDP ratio between 1970 and 2009. Chortareas et al. (2008) [33], however, caution that the results of unit root tests with structural breaks often do not agree with the results of unit root tests that posit a nonlinear mean reversion (stationarity) under the alternative hypothesis. In other words, since unit root tests with structural breaks capture the different time series characteristics of the stochastic process in question, one should use them only as complementary tests to the nonlinear unit root tests, as Chortareas et al. (2008) [33] recommend.

Since the choice of a particular alternative hypothesis in unit root tests affects their ability to reject the null hypothesis, one testing strategy for attaining the desirable power of unit root testing procedures would be to use an *F*-test of Enders and Granger (1998) [34] for the null hypothesis of a unit root against an alternative of a stationary two-regime SETAR process. The reader should note, however, that the Monte Carlo simulations of Enders (2001) [35] report that the *F*-test of Enders and Granger (1998) [34] has lower power than the traditional Dickey–Fuller unit root test of Dickey and Fuller (1981) [10], which ignores the threshold break under the alternative hypothesis. The problem with the Dickey–Fuller unit root test, on the other hand, is that it has extremely low power in the case of highly persistent near-unit root AR (1) processes, which is precisely the case for the US public debt/GDP ratio. Since both the *F*-test of Enders and Granger (1998) [34] and the Dickey–Fuller test of Dickey and Fuller (1981) [10] have low power in the case of the US public debt/GDP ratio, one potential solution is to use the efficient unit root tests of Elliott et al. (1996) [36], since Bec et al. (2022) [37] find that these unit root tests have higher power than traditional unit root tests, the single threshold-type unit root tests of Enders and Granger (1998) [34] and the two threshold-type unit root tests of Kapetanios and Shin (2006) [38] in the case when the AR (1) coefficient is larger than 0.95.

Although Gnegne and Jawadi (2013) [31] do not report the results of efficient unit root tests from Elliott et al. (1996) [36], they present, in line with the recommendations of Bohn (2007) [15], the results of the stationarity KPSS test of Kwiatkowski et al. (1992) [39]. Bohn (2007) [15] asserts that assessing the null hypothesis of stationarity against the alternative of a unit root can be of economic interest, since one can, after concluding that the null hypothesis of stationarity cannot be rejected, proceed to evaluate potential nonlinearities in the stochastic process for public debt. However, Gnegne and Jawadi (2013) [31] present only the results of stationarity testing for an intercept term without trend case, even though Figure 2 (p. 156) in their article clearly depicts the upward trending behavior in the US public debt/GDP ratio between 1970 and 2009. The realized value of the KPSS test statistics of 1.44 from Table 1 (p. 158) of Gnegne and Jawadi (2013) [31] rejects the null hypothesis of stationarity at the 5% significance level, but the results have to be interpreted with caution since the choice of an intercept term as the only deterministic component can influence the power of the stationarity test of Kwiatkowski et al. (1992) [39].

Before presenting the methodological econometric framework in the next section of this paper, it would be useful to summarize the main points about the time series properties of the US public debt/GDP ratio after the Bretton Woods collapse. First, the US public debt/GDP ratio is a (near) unit root stochastic process with a first lag autocorrelation coefficient higher than 0.95. Second, in finding the order of integration of the US public debt/GDP ratio, one should place emphasis on efficient unit root tests from Elliott et al. (1996) [36] and the stationarity test from Kwiatkowski et al. (1992) [39], using both the intercept and linear time trend as deterministic components in testing regressions. Third, to model the threshold nonlinearities in the dynamics of the US public debt/GDP ratio one should opt for the SETAR model specification instead of the ESTAR or LSTAR model specifications. Fourth, the SETAR model should be estimated in levels, not first differences, of the US public debt/GDP ratio since (i) the first differenced public debt/GDP approximates the overall fiscal balance corrected for the stock-flow adjustments and consequently has an alternative economic interpretation in comparison to the public debt/GDP ratio measured in levels; and (ii) bond investors, credit rating agencies, policymakers and international financial institutions are primarily interested in monitoring and forecasting public debt/GDP ratio in levels, not first differences (Badia et al. (2022) [40]).

2. Methods

Tsay (1989) [41] presents an early contribution to detecting the number and location of thresholds using an intuitive graphical approach on the residuals from the arranged linear AR (p) autoregression. More recently, Hansen (1996, 1997, 2017) [4–6] develops an asymptotic p-value-based approach that supports testing, estimation, and inference for general two-regime SETAR type models of order p. Following closely, almost verbatim, the exposition in Hansen (1997) [5], this section acquaints the reader with the theoretical econometric background on which the empirical estimates from Section 3 are based. Since this part of the paper is mathematical and methodological in nature, readers who are only interested in econometric model estimates can skip this section and focus exclusively on the next section, which discusses the results and their interpretation.

Following Hansen (1996, 1997, 2017) [4–6], a two-regime SETAR model with an autoregressive order p has the following form:

$$y_t = \left(a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p}\right) 1(y_{t-1} \le c) + \left(b_0 + b_1 y_{t-1} + \cdots + b_p y_{t-p}\right) 1(y_{t-1} > c) + e_t \tag{3}$$

in which $1(\cdot)$ denotes the indicator function, y_{t-1} is the threshold variable with the delay parameter d set equal to one ($d = 1$), and c is the value of the threshold. The parameters $a_0,\ a_1, \ldots, a_p$ are autoregressive slopes for the lower regime ($y_{t-1} \le c$) while $b_0,\ b_1, \ldots, b_p$ are autoregressive slopes for the upper regime ($y_{t-1} > c$). The error e_t, potentially heteroscedastic, is a martingale difference sequence.

We present only the case for the delay parameter d set equal to one ($d = 1$) since (i) the partial autocorrelation function of the US public debt/GDP ratio points to an AR (1) process ($p = 1$) for the US public debt/GDP ratio for the period in question and since (ii) by definition, we have $d \le p$, where d takes on discrete values only. Note that the estimation problem in the case of $1 < d \le p$ still implies a super consistent least squares (LS) estimate of d, since the parameter space over which one must conduct the grid search for d would be discrete.

To estimate the threshold parameter and slope coefficients, Hansen (1997) [5] introduces the following notation:

$$x_t = \begin{pmatrix} 1 & y_{t-1} & \ldots & y_{t-p} \end{pmatrix}' \tag{4}$$

and

$$x_t(c) = \begin{pmatrix} x_t' 1(y_{t-1} \le c) & x_t' 1(y_{t-1} > c) \end{pmatrix}' \tag{5}$$

that yields the following representation for Equation (3):

$$y_t = x_t' a 1(y_{t-1} \le c) + x_t' b 1(y_{t-1} > c) + e_t \tag{6}$$

or

$$y_t = x_t(c)'\theta + e_t \tag{7}$$

where

$$\theta = \begin{pmatrix} a' & b' \end{pmatrix}'.$$

Hansen (1997) [5] estimates the parameters of Equation (7), c and θ, with the conditional LS estimator. For a given value of c, the LS estimate of θ is

$$\hat{\theta}(c) = \left(\sum_{t=1}^{n} x_t(c) x_t(c)' \right)^{-1} \left(\sum_{t=1}^{n} x_t(c)\, y_t \right) \tag{8}$$

with residuals

$$\hat{e}_t(c) = y_t - x_t(c)'\hat{\theta}(c) \tag{9}$$

and residual variance

$$\hat{\sigma}_n^2(c) = \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t(c)^2. \tag{10}$$

The LS estimate of c is the value that minimizes the right-hand side of Equation (10):

$$\hat{c} = \underset{c \in \Gamma}{\operatorname{argmin}} \hat{\sigma}_n^2(c) \tag{11}$$

where $\Gamma = [\underline{c}, \overline{c}]$ describes the lower ($\underline{c}$) and the upper ($\overline{c}$) percentiles of the probability distribution of the ordered threshold variable.

Hansen (1997) [5] solves the minimization problem from Equation (11) using a direct search. The reader should note that the residual variance from Equation (10), $\hat{\sigma}_n^2(c)$, takes on at most n distinct values when c is varied, and these values correspond to $\hat{\sigma}_n^2(y_{t-1})$, $t = 1, 2, \ldots n$. Consequently, to find the LS estimate of Equation (11), Hansen (1997) [5] estimates the OLS regressions of the form $y_t = x_t(c)\prime\theta + e_t$ setting $c = y_{t-1}$ for each $y_{t-1} \in \Gamma$, which amounts to slightly less than n regressions. For each regression, Hansen (1997) [5] calculates the residual variance $\hat{\sigma}_n^2(c)$ and concentrates on the value of c that corresponds to the smallest variance:

$$\hat{c} = \underset{y_{t-1} \in \Gamma}{\operatorname{argmin}} \hat{\sigma}_n^2. \tag{12}$$

Hansen (1997) [5] finds the LS estimate of θ as $\hat{\theta} = \hat{\theta}(\hat{c})$, while the LS residuals are equal to $\hat{e}_t = y_t - x_t(\hat{c})\prime\hat{\theta}$ with sample variance $\hat{\sigma}_n^2 = \hat{\sigma}_n^2(\hat{c})$. To construct asymptotically valid confidence intervals for c, Hansen (1997) [5] recommends the use of the following likelihood ratio statistics:

$$LR_n(c) = n\left(\frac{\hat{\sigma}_n^2(c) - \hat{\sigma}_n^2(\hat{c})}{\hat{\sigma}_n^2(\hat{c})}\right) \tag{13}$$

for which $LR_n(c_0)$ is the likelihood ratio statistic to test the null hypothesis $H_0:\ c = c_0$, and for which, trivially, $LR_n(\hat{c}) = 0$ for $c = \hat{c}$. Hansen (1997) [5] further denotes the β-level critical value for ξ as $c_\xi(\beta)$ in which ξ is the random variable with the following probability density function:

$$P(\xi \leq x) = \left(1 - e^{-x/2}\right)^2 \tag{14}$$

and for which the 95% critical value equals 7.35 (see the second row of Table 1 in Hansen (1997) [5]). To find the 95% confidence interval for c, Hansen (1997) [5] sets

$$\hat{\Gamma} = \{c:\ LR_n(c) \leq c_\xi(\beta)\} \tag{15}$$

and finds the 95% confidence interval graphically by plotting the likelihood ratio $LR_n(c)$ sequence against c and then drawing a flat line at 7.35, i.e., at the 95% critical value for $c_\xi(\beta)$. Since the set $\hat{\Gamma}$ can be disjointed in practice, it is possible to perform a threshold search over a convexified region $\hat{\Gamma}^c = [\hat{c}_1, \hat{c}_2]$, where $\hat{c}_1 = \min_c \hat{\Gamma}$ and $\hat{c}_2 = \max_c \hat{\Gamma}$. Hansen (1997) [5] provides Monte Carlo evidence in favour of using $\hat{\Gamma}^c$ over $\hat{\Gamma}$.

Hansen (1997) [5] also shows how to construct β-level confidence intervals for the slope parameters of the SETAR model in question. If z_β is the β-level critical value for the normal distribution and if $\hat{s}_\gamma$ is the standard error for $\hat{\theta}(c)$, then the β-level confidence interval for θ, conditional on c being fixed, is

$$\hat{\Theta}(c) = \hat{\theta}(c) \pm z_\beta \hat{s}(c). \tag{16}$$

When c is known and equals c_0, Equation (16) trivially becomes $\hat{\Theta}(c_0) = \hat{\theta}(c_0) \pm z_\beta \hat{s}(c_0)$. Since $\hat{c}$ is a consistent estimator of c_0 at a fast rate, Hansen (1997) [5] continues as if $\hat{c} = c_0$ and use $\hat{\Theta}(\hat{c})$ as an asymptotically valid confidence interval for θ. However, since in small samples the estimates of c might not be precise, this sampling error would also affect the precision of the $\hat{\theta}$ estimates. To reduce the sampling error, Hansen (1997) [5] proposes to first construct a ϕ-level, $\phi < 1$, confidence interval for c, and for each c in this ϕ-level confidence interval, calculates a confidence interval for θ, and then forms the union of all these sets. More formally, if $\hat{\Gamma}(\phi)$ is a confidence interval for c for a given asymptotic coverage ϕ, $\phi < 1$, and if, for each $c \in \hat{\Gamma}(\phi)$, one can construct the confidence interval $\hat{\Theta}(c)$ as in Equation (16) and denote the union of these sets to

$$\hat{\Theta}_{\phi} = \bigcup_{c \in \hat{\Gamma}(\phi)} \hat{\Theta}(c) \tag{17}$$

so it is possible to reduce the sampling error of the slope parameter estimates, as Hansen's (1997) [5] Monte Carlo experiment shows. As Hansen (1997) [5] notes, by construction, $\hat{\Theta}_{\phi}$ increases with ϕ in the sense that $\hat{\Theta}_{\phi 1} \subset \hat{\Theta}_{\phi 2}$ for $\phi_1 < \phi_2$. In addition, the smallest member of this class is $\hat{\Theta}_0 = \hat{\Theta}(\hat{c})$, the confidence interval formed by ignoring the sampling variation in $\hat{c}$, so that $\hat{\Theta}_{\phi}$ is by construction larger than $\hat{\Theta}(\hat{c})$ for any $0 < \phi < 1$.

3. Results

This section consists of four subsections. Section 3.1 presents the main stylized facts for the US public debt/GDP ratio for the period 1974 Q1–2024 Q1. Section 3.2 presents the results from the unit root tests. Section 3.3 presents baseline econometric estimates. Section 3.4 provides the results of the sensitivity analyses.

3.1. Stylized Facts

Figure 1 plots the dynamics of the seasonally adjusted federal US public debt as a % of the GDP for the period 1974 Q1–2024 Q1. Following Acalin and Ball (2024) [42], we use the data from the US Office of Management and Budget. We downloaded the data from the Federal Reserve Bank of St. Louis website under code number GFDEGDQ188S.

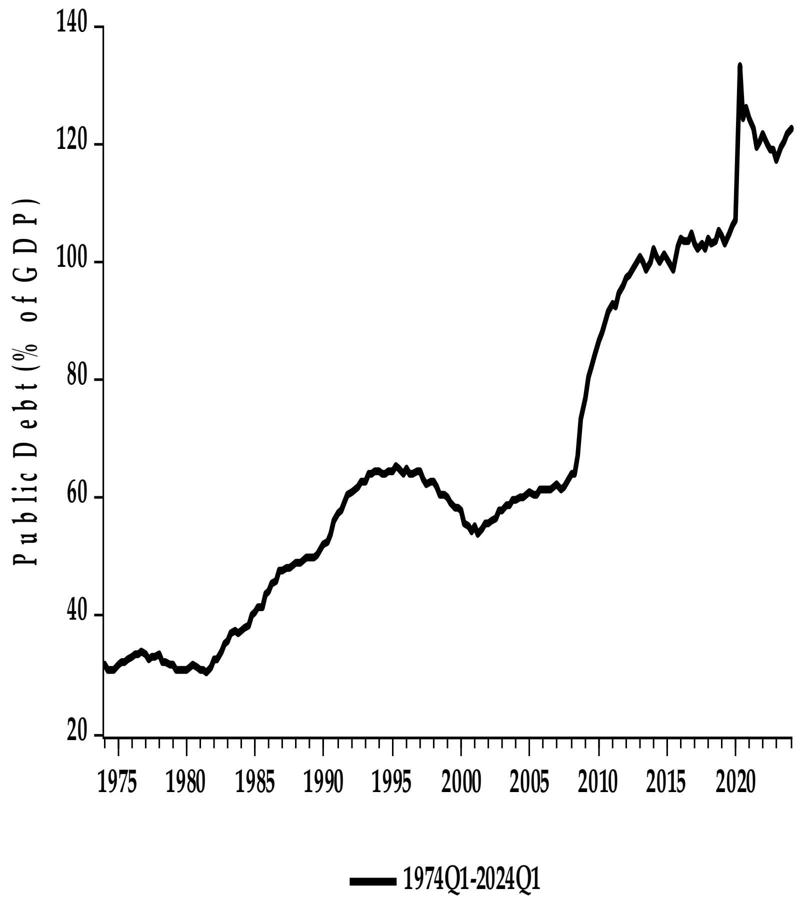

Figure 1. The US public debt/GDP ratio, 1974 Q1–2024 Q1.

Since the US public debt/GDP ratio between 1946 and 1974 fell from 106% of GDP to 23% of GDP, primarily due to the FED's interest rate pegging policy between 1942 and 1951 and unanticipated inflation during the 1960s and 1970s, as Acalin and Ball (2024) [42] document in detail, the sample starts in 1974 Q1. In addition, the 1974 Q1 sample start enables us to investigate the nonlinearities in the dynamics of the US public debt/GDP ratio for a homogeneous period of flexible exchange rates after the Bretton Woods collapse. The end of the sample, 2024 Q1, corresponds to the last publicly available data point on the 23rd of July when we downloaded the data from the FRED website.

Following the advice of Bohn (2005) [43], we measure public indebtedness as a % of GDP, since unscaled time series for the federal US public debt show non-stationary variance. In addition, we opt for the par, instead of the market, value of the federal US public debt, since Gale [44] (p. 210) reports that the market value and par value of the outstanding US federal debt move closely for the period under scrutiny.

As Figure 1 shows, the dynamics of the US public debt/GDP ratio is highly persistent. The series shows a potential structural break in 2007/2008 because of the GFC. Between 2008 Q2 and 2020 Q1, the US public debt showed a staggering increase of approximately forty percentage points of GDP. In 2008 Q2, the US public debt stood at around 65% of GDP, while in 2020 Q1, it was equal to 107% of GDP. In addition, in 2020 Q2, due to the COVID-19 pandemic, the US public debt ratio showed an upward spike of approximately twenty-five percentage points of GDP, so that in 2020 Q2, the US public debt/GDP ratio equaled 133% of GDP.

The examination of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the US public debt/GDP ratio shows that the first lag partial autocorrelation coefficient equals 0.98. The first lag partial autocorrelation coefficient is statistically significant at the 1% significance level. The second lag partial autocorrelation coefficient equals 0.01 and it is not statistically significant even at 10% significance. In sum, both the ACF and PACF point in such a direction that the US public debt/GDP ratio follows a (near) unit root AR (1) process. To formally check whether there is an actual unit in the autoregressive polynomial of the US public debt/GDP ratio, we turn to formal unit root and stationarity testing in Section 3.2 below.

3.2. Unit Root Tests

Table 1 presents the results of the conventional unit root and stationarity tests from Elliott et al. (1996) [36], Kwiatkowski et al. (1992) [39] and Ng and Perron (2001) [45]. In all the tests, we use both an intercept and linear time trend as deterministic components and choose the number of lags in the test regressions according to the modified Akaike criterion (MAIC) of Ng and Perron (2001) [45], given that we fix the maximum number of lags to four due to the quarterly business cycle data frequency. In the case of all the tests, we opt for the AR spectral GLS detrending as a long-run variance spectral estimation method.

The results of the conventional unit root and stationarity tests are mixed and identical to the results reported in Table CI of Appendix C of Jiang et al. (2024) [19] (p. 45). Jiang et al. (2024) [19] test for the unit root in the annual log US public debt/GDP ratio between 1947 and 2022 and reach the following conclusion:

Having considered the ultramodern univariate tests of the null of a unit root and the null of stationarity, we cannot reject either null hypothesis for the log D-O (debt output, added emphasis) ratio. This conclusion is perhaps not surprising in light of the difficulties that univariate tests face in distinguishing a unit root from a near-unit-root process.

Table 1. Conventional unit root and stationarity tests for the US public debt/GDP ratio.

Tests	KPSS	DF−GLS	ERS	MZ_{α}	MZ_{t}	MSB	MP_{T}
B_t	54.89 ***	−1.11	28.59 ***	−3.01	−1.11	0.37 ***	27.39 ***
Verdict	$I(1)$	$I(1)$	$I(0)$	$I(1)$	$I(1)$	$I(0)$	$I(0)$

Notes: Author's calculations. B_t-US public debt/GSD (%). *** 1% significance level, ** 5% significance level, * 10% significance level. Test regressions include a constant and linear time trend. Max number of lags in the test regressions is four due to the quarterly data frequency. Optimal number of lags in the test regressions found by the MAIC criterion of Ng and Perron (2001) [45]. Long-run variance estimation method is AR spectral GLS detrending.

One of the problems with the conventional unit root and stationarity tests from Table 1 is that they do not allow for the possibility of a structural break in the alternative hypothesis. The absence of a structural break in the intercept, trend or both can influence the power of the conventional tests from Table 1. Table 2, hence, presents the results of unit root tests that allow for the single, endogenously determined, structural break in one or both deterministic components. Analogously to the test regression specifications from Table 1, we use both an intercept and linear time trend as deterministic components and choose the number of lags in the test regressions according to the modified Akaike criterion (MAIC) of Ng and Perron (2001) [45], given that we fix the maximum number of lags to four due to the quarterly business cycle data frequency. We perform the endogenous choice of breakpoints in all the test specifications by minimizing the corresponding Dickey–Fuller test statistics. The first row outlines the results from the ZA innovational outlier unit root test model of Zivot and Andrews (1992) [32], which implies the gradual break in the intercept (first column), trend (second column) and both (third column). The second row outlines the results of the Vogelsang–Perron (VP) additive outlier unit root test model of Vogelsang and Perron (1998) [46], which implies immediate abrupt break in the intercept, trend or both the trend and intercept.

Table 2. Unit root tests with a single structural break.

Test/Break Type	Breaking Intercept	Breaking Trend	Breaking Trend and Intercept
ZA IO Model	-3.74 $(p = 0.54)$	-2.42 $(p = 0.93)$	-3.41 $(p = 0.85)$
VP AO Model	-3.40 $(p = 0.75)$	-2.25 $(p = 0.92)$	-3.62 $(p = 0.75)$
Verdict	$I(1)$	$I(1)$	$I(1)$

Notes: Author's calculations. *** 1% significance level, ** 5% significance level, * 10% significance level. Unit root test regressions include a constant and linear time trend. Max number of lags in the test regressions is four due to the quarterly data frequency. Optimal number of lags in the test regressions found by MAIC criterion of Ng and Perron (2001) [45]. Breakpoint choice: min Dickey–Fuller test statistic. First row: innovation outlier (IO) model of Zivot and Andrews (1992) [32]. Second row: additive outlier (AO) model of Vogelsang and Perron (1998) [46]. Vogelsang (1993) [47] asymptotic one-sided p-values.

Using one-sided p-values from Vogelsang (1993) [47], we cannot reject the unit root hypothesis for any of the test specifications presented in Table 2. The results in Table 2 are, hence, consistent with the findings of Gnegne and Jawadi (2013) [31]. In sum, contrary to the mixed results in Table 1, the findings in Table 2 unequivocally cannot reject the presence of a unit root in the AR polynomial of the US public debt/GDP ratio between 1974 Q1 and 2024 Q1.

The reader should, however, interpret the results in Table 2 with caution. First, both the ZA innovational outlier model and VP additive outlier model do not allow for more than one break in the respective deterministic components. Second, both the ZA innovational outlier model and the VP additive outlier model do not allow for the breaking trend under the null hypothesis—see Vogelsang and Perron (1998) [46] for a detailed discussion on this undesirable restriction which is needed to obtain distributional results for the respective Dickey–Fuller test statistics.

To allow for more than one break in the deterministic components of the unit root test regression from Table 2, but also to allow for the presence of breaks under the null hypothesis, we have also implemented the minimum Lagrange multiplier unit root test with two structural breaks of Lee and Strazicich (2003) [48] presented in Table 3. The first row of Table 3 corresponds to the "crash" model of Lee and Strazicich (2003) [48], which allows only for an abrupt change in levels. The second row of Table 3 corresponds to the "break" model, which allows for simultaneous changes in both level and trend. The optimal number of lags in both test regressions is determined with the parametric general-to-specific procedure with the inclusion of only those lagged differences of the US public debt/GDP ratio that are statistically significant at the 10% level, given that we set the maximum number of lagged differences to four, as in the case of the tests presented in Tables 1 and 2. Note that the break dates from the second column of Table 3 do not correspond to the "best fitting" breaks in terms of the residual sum of squares (RSS) like in Bai and Perron (1998) [49]. We choose the break dates in the unit root test of Lee and Strazicich (2003) [48] to provide the most evidence against the unit root null hypothesis, i.e., to be the least favorable to the unit root null. But even with a such design of break date selection, the results in Table 3 uniquely show that it is not possible to reject the presence of a unit root in the dynamics of the US public debt/GDP ratio. Overall, given the results of the unit root tests presented not only in Table 3 but also in Tables 1 and 2, the only implication is that one cannot reject a unit root null hypothesis for the US public debt/GDP ratio between 1974 Q1 and 2024 Q1.

Table 3. Lee–Strazicich (LS) unit root test with two endogenous structural breaks.

Test	LS Test Statistic	Break Dates	Verdict
LS Crash Model	−1.97	2008Q2, 2015Q3	$I(1)$
LS Break Model	−3.39	1996Q4, 2009Q3	$I(1)$

Notes: Author's calculations. *** 1% significance level, ** 5% significance level, * 10% significance level. Crash model: breaks in intercept. Break model: breaks in both intercept and trend. Max number of lags: 4 (quarterly data). Optimal number of lags: parametric general-to-specific procedure with the marginal 10% significance level.

The reader should remember, however, that all the tests from Tables 1–3 are not immune to the Bohn's (1998) [14] omitted variable bias critique. In other words, none of the tests presented so far does not control for the variations in the output gap and transitory government spending, which Bohn (1998) [14] considers crucial in recovering the mean reversion property of the US public debt/GDP ratio between 1916 and 1995. The other problem is, however, that the OLS FRFs estimates from Bohn (1998) [14] could also be potentially biased. Jiang et al. (2024) [19] show that the positive response of the primary fiscal balance to changes in the public debt/GDP ratio disappears when one implements the Stambaugh (1999) [50] small-sample bias OLS correction. Put differently, Jiang et al. (2024) [19] show that, due to the different degrees of persistence between the stationary primary fiscal balance (FRF dependent variable) and potentially non-stationary public debt/GDP ratio (FRF predictor), the positive FRF response of the primary balance to changes in the public debt is a consequence of a small-sample bias. Unfortunately, as previously said in the introductory remarks, the situation is further complicated by the findings of Cochrane (2020, 2022) [16,17], who recovers a positive primary fiscal balance response in a multivariate, dynamically richer, VAR setting for the post-World War II US fiscal data. In sum, although unit root tests might be misleading in establishing the degree of persistence in the US public debt/GDP ratio, one cannot obtain conclusive evidence in favor of rejecting the unit root null hypothesis by applying different, on the economic grounds founded, univariate and multivariate econometric techniques designed to potentially correct for the shortcomings of unit root tests.

If the econometric evidence gathered so far from various unit root, FRF and VAR statistical exercises cannot yield an unambiguous answer to the question of the unit root presence in the US public debt/GDP ratio, we can potentially find the answer in alternative economic theories. Alternative economic theories, however, also do not yield a definitive answer to the question at hand. A seminal tax-smoothing model of Barro (1979) [2] predicts a random walk behavior for the US public debt. On the other hand, Aiyagari et al. (2002) [51] show that the results can diverge in important ways from those presented in Barro (1979) [2] if one accepts the Lucas–Stokey (1983) [52] state-contingent complete markets framework.

Since economic theory also does not provide a clear-cut answer concerning how to characterize the persistence in the public debt/GDP ratio, we must resort, as Tsay (2005) [53] (Section 4.3, p. 191) advises in the case of univariate nonlinear modeling, to a subjective judgement using historical and institutional characteristics of the US fiscal policy making after World War II. We reiterate once more that the issue of non-stationarity is paramount in the context of nonlinear (SE)TAR modeling since the underlying econometric frameworks of Hansen (1996, 1997, 2017) [4–6], González and Gonzalo (1997) [7] and Gonzalo and Pitarakis (2002) [8], which we apply in the next section of this paper, are all based on the assumption that the underlying stochastic process under investigation is ergodic and globally stationary.

Following the lead of Jiang et al. (2024) [19] (p. 13), who claim that "It is also worth emphasizing that it seems less plausible (added emphasis) for an economic model to impute a unit root to the market value of the debt by output ratio", we treat the US public debt/GDP ratio as stationary, near-unit root, stochastic process, and we are reluctant to accept the hypothesis that there is an actual unit root in the AR polynomial of the US public debt/GDP ratio due to the following reasons. First, if the US public debt/GDP ratio is an actual $I(1)$ unit root stochastic process, then, as Engle and Granger (1987) [54] show, the public/GDP ratio could breach any upper bound with certainty, i.e., with probability one. Moreover, from an institutional perspective, one has to consider if it is plausible, given the fiscal ceilings on the US federal debt, to argue that public debt/GDP ratio could grow without bound. Put differently, if the US public debt/GDP ratio is an actual $I(1)$ unit root stochastic process, then one must accept that its variance, given enough time, would become infinite. Second, if the US public debt/GDP ratio is an actual $I(1)$ unit root stochastic process, then innovations have a permanent effect on its dynamics. In other words, the shocks to the US public debt/GDP ratio would never show a transitory decay. From a historical perspective, we find extraordinarily little evidence for such public debt/GDP behavior in the case of the US. Yoon (2012) [55] (p. 2, Figure 2), for example, asserts that the US public debt/GDP ratio was explosive during World War II. The explosiveness of the US public debt, i.e., a larger than one root in the AR polynomial of the US public debt/GDP ratio, would imply the not only permanent but increasing effect of shocks on the dynamics of US public debt during World War II. Acalin and Ball (2024) [42] show, however, that between 1946 and 1974, the FED and the US government implemented an unconventional mix of fiscal and monetary policy actions to reduce the share of public debt in the GDP from 106% in 1946 to 23% in 1974 and to accommodate the unprecedented public debt from the World War II buildup. Some of the policy actions included (i) the FED's pegging of interest rates at low levels between 1946 and 1951, when the FED–Treasury Accord was enacted; and (ii) the surprise inflation in 1960s and 1970s that reduced the share of US public debt in the GDP due to its relatively longer maturity with respect to its current maturity structure.

Overall, we are inclined to follow the calibrations of Bhandari et al. (2017) [56] (p. 619), which impute that the optimal policy for government debt implies a slow mean reversion with a half-life of almost 250 years. In other words, Bhandari et al. (2017) [56] argue that it takes 250 years for a shock to the US public debt/GDP ratio to lose 50% of its contemporaneous impact. Given such an extraordinary degree of persistence in the US public debt/GDP ratio, it becomes more understandable why earlier $I(1)$ results from the literature in the case of much smaller samples (such as, for example, ours) can treat a (potentially nonlinear) near-unit root stochastic process with slow mean reversion as an observationally equivalent actual unit root in the US public debt/GDP ratio.

The contributions of González and Gonzalo (1997) [7], Lanne and Saikkonen (2002) [21] and Hansen (2017) [6] provide an econometric rationale for the case of observational equivalence in the context of globally stationary (SE)TAR models for highly autocorrelated time series with (potential) partial unit roots. In particular, González and Gonzalo (1997) [7] present a class of (self-exciting) threshold unit root-(SE)TUR- models that, while keeping the structure and properties of the globally stationary (SE)TAR models, allow for unit roots in some of the regimes. González and Gonzalo (1997) [7] show that least squares (LS) estimates of the parameters of these models are consistent and asymptotically normal. Lanne and Saikkonen (2002) [21] consider a threshold autoregressive (TAR) process with the threshold effect only in the intercept term. Although these processes are globally stationary, they closely resemble those of (near) integrated processes for small sample sizes. The idea behind the Lanne and Saikkonen's (2002) [21] approach is that if level shifts cause (near) integratedness, the series purged of these shifts should be stationary. Jiang et al. (2019) [24] use the approach of Lanne and Saikkonen (2002) [21] to demean the log annual US public debt/GDP ratio data between 1947 and 2022 to provide more evidence in the case of US public debt/GDP ratio stationarity. Jiang et al. (2019) [24] write the following on page 4: "We define the transitory component of the D-O ratio as the raw D-O series minus the different subsample mean before and after 2007 (added emphasis). Using this transitory D-O series, we find stronger evidence for surplus predictability (added emphasis) but not return predictability. Fundamentals account for about 50% of the variation in the transitory component of the D-O ratio at the 10-year horizon (added emphasis). Removing a structural break removes a low-frequency component in the D-O ratio and creates more room for fundamentals in explaining the now-more-transitory nature of the variation in the D-O ratio (added emphasis). The resulting transitory D-O ratio is less persistent (added emphasis), and predictive coefficients have smaller small-sample biases (added emphasis)". Finally, Hansen (2017) [6] estimates the continuous TAR model of economic growth and public debt/GDP in the case of US between 1791 and 2009. Hansen (2017) [6] finds the threshold following Hansen (1996) [4], as is the case in this article. We ground our baseline econometric estimates in the subsection that follows on the results from Hansen (1996) [4], González and Gonzalo (1997) [7], Lanne and Saikkonen's (2002) [21] and Hansen (2017) [6].

3.3. Baseline Econometric Estimates

Table 4 presents the conditional LS estimates of Equation (3) for the $p = d = 1$ case. The online Supplementary Material contains all details regarding data and econometric estimates for all the results presented in this paper. The estimated value of the threshold for the one-quarter lagged US public debt/GDP ratio is 65% of GDP and corresponds to the second quarter of 1995, 1995 Q2. All the estimated coefficients, except the intercept term in the lower regime ($B_{t-1} \leq 65.31\%$), are statistically significant at the 1% significance level. We do not use the heteroscedasticity consistent standard errors since White's heteroscedasticity test does not detect the presence of heteroscedasticity in the residual values of the US public debt/GDP ratio from the estimated AR (1) linear autoregression. In terms of the statistical significance of the estimated coefficient values, the results do not change, however, even if we use White's heteroscedasticity corrected standard errors.

The 95% confidence interval for the estimated threshold break is $[65.04, 67.29]$. Note that the upper value for the threshold of 67.29% in the case of the 95% confidence interval corresponds to the third quarter of 2008, 2008 Q3, when the macro and fiscal effects of the GFC in the United States started to unravel. Figure 2 shows the estimated threshold break with an associated 95% critical value and the 95% confidence interval for the threshold break. More precisely, the fixed solid line in Figure 2 stands for the 95% critical value from Hansen (1997) [5]. The dotted line stands for the values of the likelihood ratio (LR) test statistics for different lagged values of the public debt/GDP ratio (B_{t-1}) ranging from zero to more than 120% of public debt/GDP ratio. The value of the threshold for the one-quarter lagged public debt/GDP ratio corresponds to the zero value of the LR test statistic, i.e., it is found in the intersection of the LR test statistic sequence and the x-axis on which we order the values of the one-quarter lagged public debt/GDP ratio in the ascending fashion. We obtain the 95% confidence interval boundaries in the intersection of the LR test statistic sequence and the 95% critical value, as represented by the solid black horizontal line in Figure 2.

Table 4. Baseline SETAR (2, 1, 1) model for the US public debt/GDP ratio, 1974 Q1–2024 Q1.

Regressors	Coefficients	Standard Errors	95% Interval
	$B_{t-1} \leq 65.31\%$		
Intercept	0.37	0.74	[−1.10, 1.82]
B_{t-1}	0.99 ***	0.01	[0.97, 1.03]
	$65.31\% < B_{t-1}$		
Intercept	9.47 ***	2.10	[1.30, 16.59]
B_{t-1}	0.92 ***	0.02	[0.85, 0.99]

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. B_t: dependent variable (public debt/GDP ratio). B_{t-1}: threshold variable (15% trimming percentage for threshold search with ordinary standard errors and one thousand bootstrap repetitions).

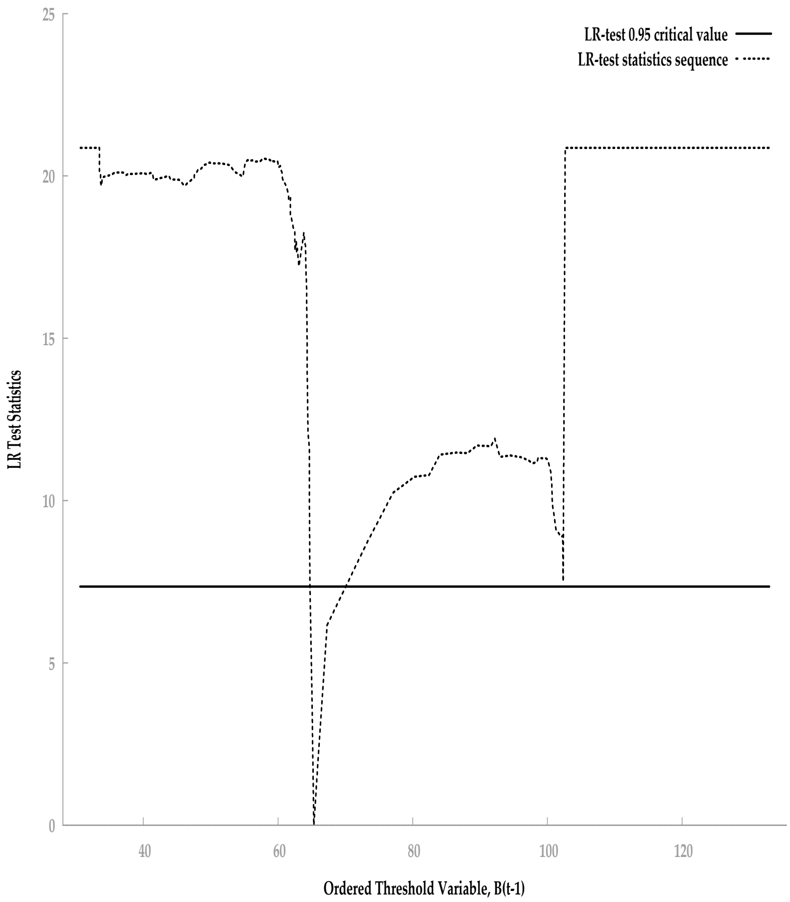

Figure 2. The 95% confidence interval for the threshold variable B_{t-1} from the baseline SETAR (2, 1, 1) model.

The residual values from the estimated SETAR (2, 1, 1) model, which stands for the two-regime AR (1) process with delay lag one in each of the regimes, are not normally distributed due to a single twenty-five percentage points COVID-19 outlier in the second quarter of 2020, 2020 Q2. Note, however, that assumptions from Hansen (1996, 1997, 2017) [4–6] only require that the residuals be a (potentially heteroscedastic) martingale difference sequence, i.e., the residuals should only be iid (independent and identically distributed) with zero mean and (potentially) constant variance. Consequently, the results of the Breusch–Godfrey serial correlation LM test that we use to test for the residual autocorrelation reject its presence up to four lags at a 1% significance level. Furthermore, the results of White's heteroscedasticity test cannot reject the null hypothesis of homoscedasticity for the residual values of the US public debt/GDP ratio from the estimated two-regime SETAR model presented in Table 4. Finally, in spite of the detected COVID-19 outlier, the coefficient stability tests based on the recursive residuals report that the estimated coefficients are stable. Figure 3 shows the recursive residuals associated with the CUSUM stability test statistics, while Figure 4 presents the recursive residuals associated with the CUSUM of the squares test statistics. The results of the CUSUM test, which is, according to Hansen (1992) [57], a test of constancy in intercepts, do not show any signs of instability in the estimated level shifts. On the other hand, the results of the CUSUM of the squares stability test, which is, again according to Hansen (1992) [57], a test of constancy in the residual variance, detect signs of instability in the variance of the model's residuals between 2012 Q2 and 2020 Q2, which is consistent with the documented sovereign debt build-in in the case of the US in the aftermath of the GFC. Note, however, that the variability of the residual variance eventually stabilizes in the post-COVID-19 period, hence supporting our earlier claim that the COVID-19 outlier does not significantly influence the model's estimates on statistical grounds. Jiang et al. (2023) [60] report that the FED bought all the new issuance of long-term bonds in the quantitative easing programs mounted in response to the COVID-19 pandemic, which provided more stability in the residual variability of the US public debt/GDP ratio.

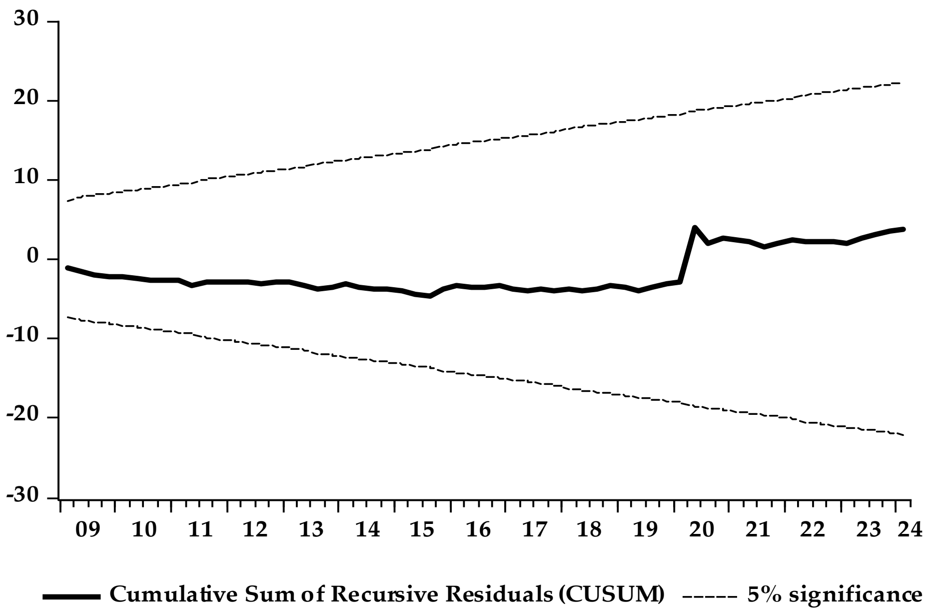

Figure 3. Cumulative sum of recursive residuals (y-axis) with respect to time (x-axis) from the baseline SETAR (2, 1, 1) model.

The estimated AR (1) slope coefficients are highly persistent: the AR (1) slope coefficient in the lower regime ($B_{t-1} \leq 65.31\%$) equals 0.99 with the 95% confidence interval of $[0.97,\ 1.03]$, while the AR (1) slope coefficient in the upper regime ($B_{t-1} > 65.31\%$) equals 0.92 with the 95% confidence interval of $[0.85,\ 0.99]$. Although both AR (1) slope coefficients are highly persistent, the Wald coefficient restriction test rejects the null hypothesis of their equality—the realized value of the chi-squared test statistics with one degree of freedom, $\chi^2(1)$, equals 7.90, with an associated p-value of 0.005. In other words, the AR (1) slope coefficient in the lower regime is not statistically different from one, according to the results of the Wald coefficient restriction test ($\chi^2(1) = 0.015,\ p = 0.90$), while the AR (1) coefficient in the upper regime is statistically different from one, according to the results of the Wald coefficient restriction test $(\chi^2(1) = 13.06,\ p = 0.00)$. Even though the lower regime AR (1) slope coefficient shows potential partial unit root behavior, note that the use of the LS estimator is justifiable, i.e., the LS estimates are consistent and asymptotically normal, as shown in González and Gonzalo (1997) [7].

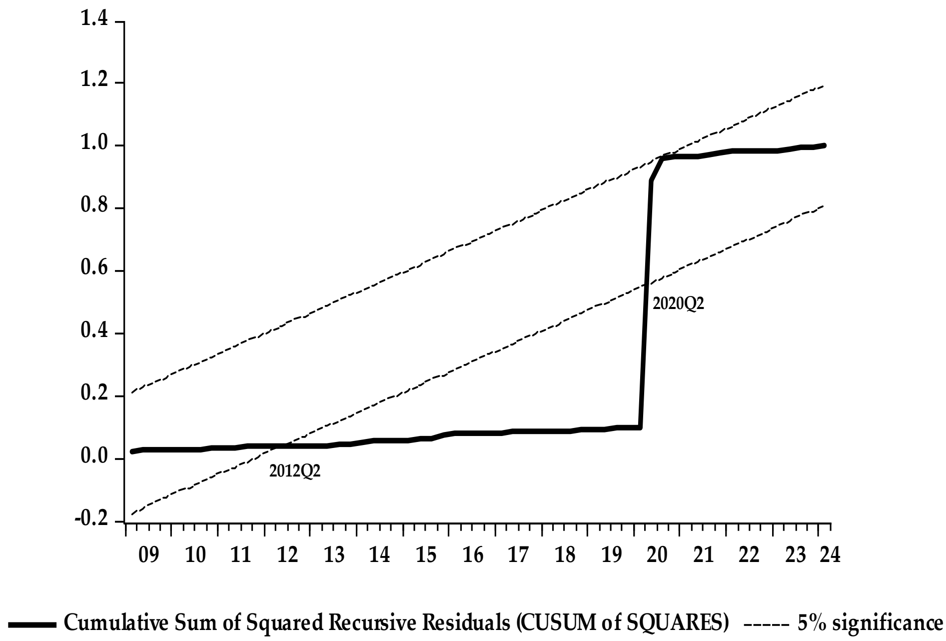

Figure 4. Cumulative sum of squared recursive residuals (y-axis) with respect to time (x-axis) from the baseline SETAR (2, 1, 1) model.

The AR (1) slope coefficient for one-quarter lagged public debt/GDP ratio, B_{t-1}, in the lower regime ($B_{t-1} \leq 65.31\%$), equals 0.99, i.e., it exhibits (near) unit root behavior. This finding is in line with the seminal tax-smoothing model of Barro (1979) [2] briefly reviewed in the introduction to the paper. In other words, it could be the case that the dynamics of the US public debt/GDP ratio below the 65.31% threshold behave randomly due to unanticipated shocks in the output gap and transitory government spending. In addition, Aiyagari et al. (2020) [43] further argue that the high persistence of public debt is due to market incompleteness, which implies that the government cannot issue state-contingent debt, i.e., it can only issue risk-free debt.

The AR (1) slope coefficient for the one-quarter lagged public debt/GDP ratio in the upper regime ($B_{t-1} > 65.31\%$) equals 0.92., i.e., it shows highly persistent, but not unit root, behavior. If one inspects the data for the US public debt/GDP ratio, it becomes clear that the public debt/GDP ratio in the case of the United States trends above the 65.31% threshold after 2008 Q3. In fact, if one disregards the 2008 Q3–2024 Q1 subsample, the only quarter when the public debt/GDP ratio is higher or equal than the 65.31% endogenous threshold is 1995 Q2. This finding, hence, confirms that our endogenously estimated threshold corresponds to a Chow identified exogenously imposed 2007 structural break in Jiang et al. (2024) [19]. In other words, our result shows that Jiang et al. (2024) [19] were correct to place the structural break in the US public debt dynamics in 2007.

More importantly, the estimated persistent AR (1) slope coefficient for the upper regime of 0.92 is also in line with theoretical conjectures and predictions from Jiang et al. (2024) [58–60]. In particular, Jiang et al. (2021, 2022, 2023) [58–60] argue that the persistent behavior of the US public debt/GDP after the GFC is due to (i) the perception of foreign investors that the US Treasury bonds represent a risk-free asset, i.e., that the US government is a supplier of safe assets; (ii) the FED's inelastic demand for the US Treasury bonds after 2008 through its quantitative easing programs; and (iii) biased subjective expectations and beliefs on the part of bond investors.

Jiang et al. (2021, 2022, 2023) [58–60] further argue that there is a belief in the international investment community that the US government is a supplier of risk-free government bonds. This exorbitant privilege enables the US government to countercyclically issue risk-free debt at low interest rates, further accentuating the persistency in the autocorrelation profile of the US public debt/GDP ratio. In other words, the inelastic foreign demand for the US Treasury bonds implies that, contrary to other economies around the world, the US government can issue new debt even in "bad" times, and hence "insure" the US taxpayers in the short-to-medium run by not raising taxes or cutting spending to cover the pro-cyclical increase in fiscal deficit due to an aggregate recession shock.

The foreign investment community is not the only inelastic buyer of the US Treasury bonds. The FED, through its quantitative easing programs, managed to buy large shares of the US government debt. In fact, from 2008 Q3 to 2024 Q1, the amount of the US federal debt held by the FED as a % of the GDP increased by approximately 14 percentage points, from 3.2% of GDP in 2008 Q3 to 17.6% of GDP in 2024 Q1, as readers can see from the chart under the ticket code HBFRGDQ188S on the official website of the Federal Reserve Bank of St. Louis. Note that the fourteen percentage points increase in the FED's holdings of the US public debt/GDP ratio since 2008 Q3 is remarkably close to the intercept level shift between the two regimes of our preferred SETAR model from Table 4. In particular, while in the lower ($B_{t-1} \leq 65.31\%$) regime the intercept is not statistically significant, the 95% confidence interval for the intercept in the upper ($B_{t-1} > 65.31\%$) regime is $[1.30,\ 16.59]$, capturing potentially the level-shift in the FED's asset purchases.

The likelihood that the FED would prevent a default of the US government might also influence the subjective beliefs of bond investors. In other words, the investors in US Treasury bonds might continue to buy the US federal debt if they have optimistic expectations about the future fiscal outlook in the United States. These overly optimistic expectations, supported by the FED's actions, can induce highly persistent shocks in the US public debt/GDP ratio after the GFC.

Since the estimated baseline SETAR (2, 1, 1) model from Table 4 is potentially useful for confirming the latest theoretical work on the US fiscal capacity, the next subsection provides sensitivity analyses for the baseline estimates presented in Section 3.3. We first show how our results are robust when we use a piece-wise linear AR (2) process for the US public debt/GDP ratio irrespective of whether the delay lag parameter is one-quarter lagged or two-quarters lagged public debt/GDP ratio. In addition, we also estimate a three-regime SETAR (3, 1, 1) model and show that its estimates are consistent with the theoretical framework of Elenev et al. (2024) [61]. Finally, we compare the estimates from our baseline SETAR (2, 1, 1) model with estimates from the corresponding ESTAR and LSTAR model specifications of the same order.

3.4. Sensitivity Analyses

Following Sarno (2001) [22], Cochrane (2001) [62] and Jiang et al. (2022) [59], we experiment with the AR (2) autoregression in each of the regimes of the estimated SETAR (2, 2, 1) model. Table 5 presents the LS estimates of the SETAR model with two piece-wise linear AR (2) regimes for which the one-quarter lagged public debt/GDP ratio serves as a threshold variable. Figure 5 shows the 95% confidence interval for the threshold variable B_{t-1} for the estimated SETAR (2, 2, 1) model outlined in Table 5.

Table 5. SETAR (2, 2, 1) model for the US public debt/GDP ratio, 1974 Q1–2024 Q1.

Regressors	**Coefficients**	**Standard Errors**	**95% Interval**
	$B_{t-1} \leq 65.31\%$		
Intercept	0.34	0.73	[−1.10, 1.79]
B_{t-1}	1.29 ***	0.26	[0.78, 1.79]
B_{t-2}	−0.29	0.26	[−0.79, 0.21]
	$65.31\% < B_{t-1}$		
Intercept	9.88 ***	2.07	[−9.50, 36.45]
B_{t-1}	0.72 ***	0.07	[0.30, 1.04]
B_{t-2}	0.20 ***	0.02	[−0.15, 0.59]

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. B_t: dependent variable (public debt/GDP ratio). B_{t-1}: threshold variable (15% trimming percentage for threshold search with ordinary standard errors and one thousand bootstrap repetitions).

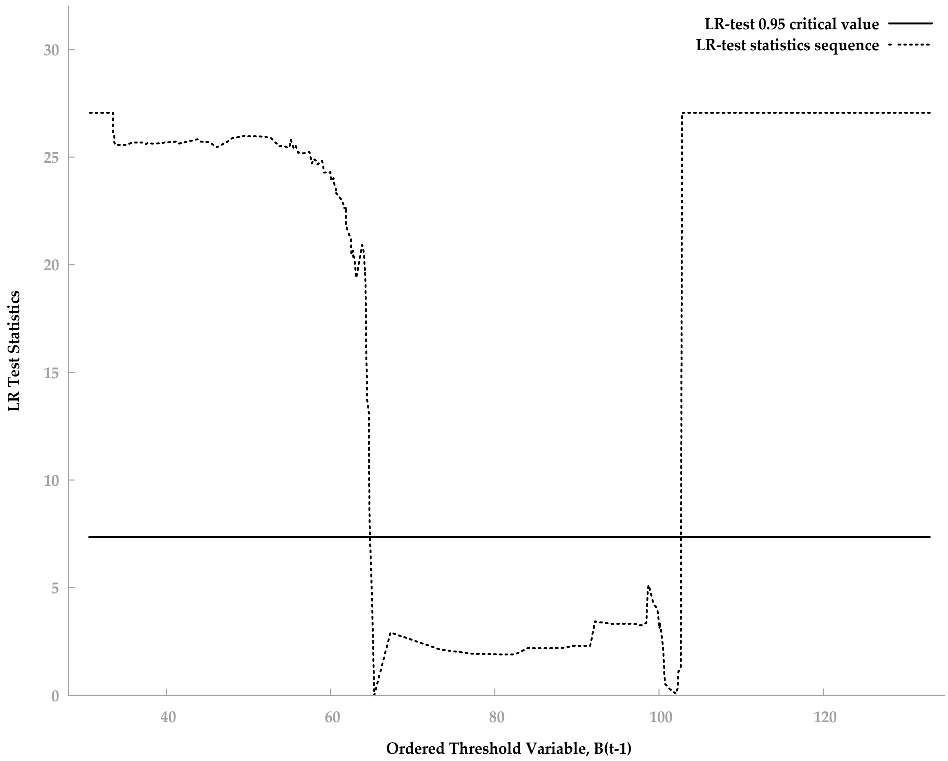

Figure 5. The 95% confidence interval for the threshold variable B_{t-1} from the SETAR (2, 2, 1) model.

From Table 5 and Figure 5, we can see that the estimated threshold is identical to the one from the baseline SETAR (2, 1, 1) model presented in Table 4 and Figure 2. Note, however, that the 95% confidence interval is much wider in comparison to the one from the baseline SETAR (2, 1, 1) model. More precisely, the 95% confidence interval for the threshold is [65.04, 102.64]. Moreover, a closer inspection of Figure 5 shows that there is potentially a second threshold at the upper 95% confidence interval boundary, i.e., the LR test statistics sequence equals zero not only for the 65.04% public debt/GDP threshold but also for the 102.64% public debt/GDP. In other words, the US public debt/GDP ratio between 1974 Q1 and 2024 Q1 might show a three (lower, middle, and upper) regime behavior, as conjectured in the theoretical model of Elenev et al. (2024) [61].

As for the coefficient estimates in the lower regime of the SETAR (2, 2, 1) model from Table 5, the AR (2) slope coefficient, along with the intercept term, is not statistically significant. On the other hand, the AR (1) slope coefficient in the lower regime equals 1.3, and it is statistically significant at 1% level, and more importantly, it points in the direction of potentially explosive AR (1) dynamics below the estimated threshold. The statistically significant estimates for the AR (1) and the AR (2) slope coefficients in the upper public debt regime support the findings in Table 4. In particular, the sum of the estimated AR (1) and AR (2) slope coefficient equals 0.72 + 0.20 = 0.92, which is identical to the degree of persistence captured by the upper-regime AR (1) slope coefficient in the baseline SETAR (2, 1, 1) model from Table 4. In summary, the results from the estimated SETAR (2, 2, 1) model presented in Table 5 and Figure 5 are consistent with the baseline SETAR (2, 1, 1) estimates from Table 4 and Figure 2.

Since the SETAR (2, 2, 1) model is a piece-wise linear AR (2) model, we further experiment with the variant of this model for which the threshold break occurs in a two-quarter lagged public debt/GDP ratio. In other words, we estimate a SETAR (2, 2, 2) model specification for which the delay lag equals two (d = 2). Table 6 outlines the coefficient estimates while Figure 6 shows the 95% confidence interval for the threshold variable B_{t-2} for the estimated SETAR (2, 2, 2) model from Table 6.

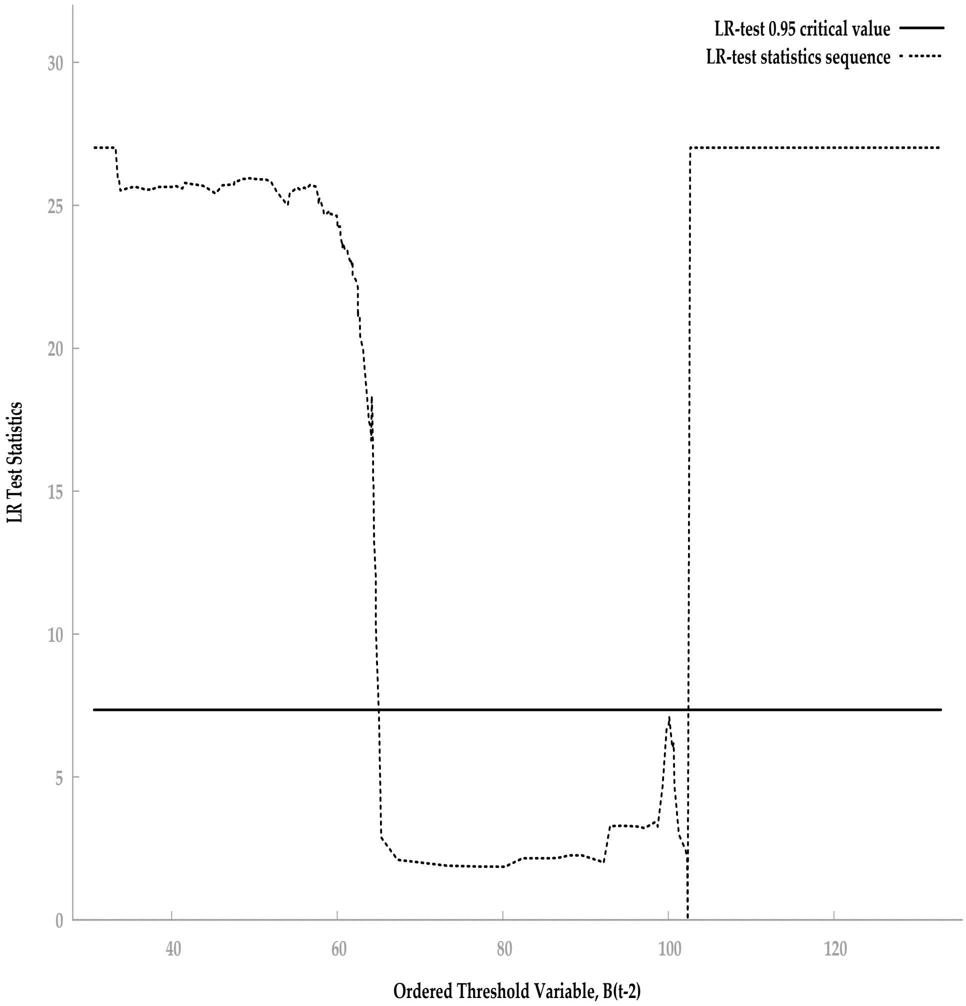

Figure 6. The 95% confidence interval for the threshold variable B_{t-2} from the SETAR (2, 2, 2) model.

Table 6. SETAR (2, 2, 2) model for the US public debt/GDP ratio, 1974 Q1–2024 Q1.

Regressors	Coefficients	Standard Errors	95% Interval
	$B_{t-2} \leq 102.3\%$		
Intercept	0.13	0.48	[−0.81, 1.06]
B_{t-1}	1.44 ***	0.16	[1.14, 1.75]
B_{t-2}	−0.44 ***	0.16	[−0.75, 0.13]
	$102.3\% < B_{t-2}$		
Intercept	14.66 ***	4.80	[−9.18, 37.45]
B_{t-1}	0.67 ***	0.08	[0.30, 1.04]
B_{t-2}	0.21 ***	0.08	[−0.16, 0.57]

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. B_t: dependent variable (public debt/GDP ratio). B_{t-2}: threshold variable (15% trimming percentage for threshold search with ordinary standard errors and one thousand bootstrap repetitions).

From Table 6 and Figure 6, we can see that the estimated threshold equals 102.3% of GDP and is identical to the upper 95% confidence interval boundary in the case of the SETAR (2, 2, 1) model from Table 5 and Figure 5. In other words, the finding of a potentially alternative public debt/GDP threshold of 102.3% of GDP reinforces the findings in Table 5 and Figure 5 that the US public debt/GDP between 1974 Q1 and 2024 Q1 might be characterized as a three-regime (two-threshold) SETAR (3, 1, 1) stochastic process. Note also that the 95% confidence interval for the threshold in the case of the SETAR (2, 2, 2) model is [65.04, 102.64], i.e., it is identical to the 95% confidence interval from the previously reported SETAR (2, 2, 1) model. In fact, the lower 95% confidence interval boundary of 65.04 exactly equals the threshold from Table 5 and Figure 5.

As the coefficient estimates in the lower regime of SETAR (2, 2, 2) model from Table 6 show, both the AR (1) and the AR (2) slope coefficients in the lower public debt regime are statistically significant at the 1% significance level. The fact is that the sum of the two autoregressive slope coefficients below the 102.3% public debt/GDP threshold, $1.44 - 0.44 = 1$, approximates the random walk behavior envisioned by Barro (1979) [2]. The sum of the statistically significant autoregressive coefficients in the upper regime equals 0.88 $(0.67 + 0.21)$ and is almost identical in terms of the captured persistence to the estimates from the SETAR (2, 1, 1) and SETAR (2, 2, 1) models presented in Tables 4 and 5, respectively. In summary, the findings from the SETAR (2, 2, 2) model from Table 6 and Figure 6 support the previously reported estimates from the SETAR (2, 1, 1) and SETAR (2, 2, 1) models.

Given the results in Tables 4–6, we finally turn our attention to estimating a two-threshold SETAR (3, 1, 1) model using the sequential LS estimator of Gonzalo and Pitarakis (2002) [8]. We present the estimates of a three-regime SETAR model in Table 7. All the estimated coefficients, except the intercept term in the lower ($B_{t-1} \leq 65.31\%$) regime, are statistically significant at the 1% significance level. Our three regime SETAR model from Table 7 is, to a certain extent, consistent with the theoretical model of Elenev et al. (2024) [61]. Elenev et al. (2024) [61] construct a dynamic stochastic general equilibrium model (DSGE) for the US economy with the novel feature that fiscal policy in the US does not respond continuously, but rather discretely, to the movements in the public debt/GDP ratio. Furthermore, Elenev et al. (2024) [61] calibrate that in "normal times", when only productivity shocks govern the economic dynamics and monetary policy is "conventional", the upper austerity threshold for the US public debt/GDP ratio equals 115%, while the lower profligacy threshold equals 47.5%. In the structural model of Elenev et al. (2021), the fiscal authority runs countercyclical spending and pro-cyclical tax policies if the debt/GDP ratio is below an austerity threshold. Once the debt/GDP ratio exceeds the threshold, fiscal policy switches from active (macroeconomic stabilization) to passive (controlling the debt). Elenev et al. (2024) [61] find the austerity and profligacy thresholds endogenously to keep the public debt/GDP ratio risk-free and stationary. Our respective endogenously estimated thresholds of 65.31% and 100.5% could potentially correspond to the calibrated threshold values of Elenev et al. (2024) [48]. Note, however, that there is little difference between the estimated coefficient values in the middle ($65.31 < B_{t-1} \leq 100.5\%$) and the upper ($100.5\% < B_{t-1}$) regimes, implying that the two-regime SETAR model from Table 4 is probably a better characterization of the nonlinearities and asymmetries in the dynamics of the US public debt/GDP ratio between 1974 Q1 and 2024 Q1.

Table 7. Three-regime SETAR (3, 1, 1) model for the US public debt/GDP ratio, 1974 Q1–2024 Q1.

Regressors	Coefficients	Standard Errors	95% Interval
	$B_{t-1} \leq 65.31\%$		
Intercept	0.37	0.74	[−1.07, 1.82]
B_{t-1}	0.99 ***	0.01	[0.97, 1.03]
	$65.31 < B_{t-1} \leq 100.5\%$		
Intercept	12.69 ***	4.25	[4.38, 21.02]
B_{t-1}	0.88 ***	0.05	[0.79, 0.97]
	$100.5\% < B_{t-1}$		
Intercept	11.64 ***	4.11	[3.58, 19.70]
B_{t-1}	0.90 ***	0.04	[0.83, 0.97]

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. B_t: dependent variable (public debt/GDP ratio). B_{t-1}: threshold variable (15% trimming percentage for threshold search with ordinary standard errors and one thousand bootstrap replications). Sequential threshold estimation of Gonzalo and Pitarakis (2002) [8].

Finally, Table 8 compares all the estimated SETAR models in terms of the AIC and BIC. For comparison purposes, Table 8 also includes the values of the AIC and BIC for the AR (1), AR (2), SETAR (3, 1, 1), SETAR (3, 2, 1) and SETAR (3, 2, 2) model specifications. We include the last two model specifications to allow for the possibility of AR (2) autoregressions in each of the regimes and to compare various delay lags (d = 1 vs. d = 2). From Table 8, it is evident that the best performance in terms of the minimized AIC and BIC values is achieved in the case of the SETAR (2, 1, 1) and SETAR (2, 2, 1) model specifications, with a slightly better performance in the case of the SETAR (2, 2, 1) model. The AR (2) specification is also a preferred choice by Sarno (2001) [22], Cochrane (2001) [62] and Jiang et al. (2022) [59], although in the context of linear autoregressions and in the case of annual data.

Table 8. Information criteria for alternative linear and nonlinear autoregressive model specifications.

Tests	***AR* (1)**	***AR* (2)**	***SETAR* (2,1,1)**	***SETAR* (2,2,1)**	***SETAR* (2,2,2)**	***SETAR* (3,1,1)**	***SETAR* (3,2,1)**	***SETAR* (3,2,2)**
AIC	323.29	321.71	89.57	81.78	92.37	307.7	302.3	305.35
BIC	329.9	331.59	99.68	96.93	105.96	327.48	331.9	335.0

Notes: Author's calculations. AIC: Akaike information criterion. BIC: Bayesian information criterion.

In further analyses, we opt, however, for the SETAR (2, 1, 1) model specification for the following reasons. First, the PAC function shows that the quarterly data for the US public debt/GDP ratio between 1974 Q1 and 2024 Q1 are consistent with the AR (1) model specification. Second, the SETAR (2, 1, 1) model specification is more comparable with the results of the unit root tests that search for the presence of a unit root in the AR (1) polynomial. Third, the values of the AIC and BIC in the case of both SETAR (2, 1, 1) and SETAR (2, 2, 1) model specifications are almost identical, although they should be taken only as a preliminary guide in the context of nonlinear model selection, as emphasized by Psaradakis et al. (2009) [63]. Finally, the SETAR (2, 1, 1) model specification aids comparison with the respective first-order ESTAR and LSTAR model specification, a task that we undertake in the rest of this subsection.

Table 9 reports the estimates from the first-order ESTAR model characterized by the smooth exponential transition function of the form $\Phi[\theta; d_{t-k} - c] = 1 - exp\left[-\theta(d_{t-k} - c)^2\right]$, where θ measures the speed of transition between the two regimes and c denotes the threshold public debt/GDP ratio. The following important findings are worth emphasizing about the ESTAR estimates from Table 9. First, none of the estimated coefficients are statistically significant, but the estimated 110.7% public debt/GDP threshold is, a logically inconsistent finding. If the statistically significant threshold estimate supports a nonlinear adjustment mechanism, how come the estimated coefficients across regimes do not support the same type of adjustment? More importantly, the estimated coefficient values are unrealistically large and show perfect offsetting across regimes. This finding is precisely the sort of paradox that Buncic (2019) [29] (Table 3, p. 677) talks about when he replicates the unconstrained estimates of Taylor et al. (2001) [64]. Intuitively, a relatively small value for the estimated (statistically significant) speed of transition parameter implies that when $\theta \to 0$, then the exponential transition function $\Phi[\theta; d_{t-k} - c] = 1 - exp\left[-\theta(d_{t-k} - c)^2\right] \to 0$ as well. In other words, even though the estimated threshold implies a nonlinear adjustment, the slow speed of transition effectively implies only one regime. To familiarize the reader more closely with this finding, we plot the exponential transition function of the estimated ESTAR (2, 1, 1) model (y-axis) with respect to the threshold variable (x-axis). From Figure 7, one can infer the following: (i) the exponential transition function assigns value 0 for the threshold estimate; (ii) the exponential transition function assigns value 1 for almost all the other values of the threshold variable (one-quarter lagged public debt/GDP ratio); and (iii) it assigns a non-zero value for the values of the threshold variable between 105% and 117% of GDP, which amounts to only four sample observations! In other words, the exponential transition function effectively acts as an indicator outlier dummy function, as Buncic (2019) [29] correctly asserts, and effectively assigns all the values of the threshold variable in just one regime, failing to adequately capture the nonlinear adjustment mechanism.

Table 9. ESTAR (2, 1, 1) model for the US public debt/GDP ratio, 1974 Q1–2024 Q1.

Regressors	Coefficients	Standard Errors	t−Stat
	$B_{t-1} \leq 110.7\%$ ($\theta = 0.15$)		
Intercept	−19,243.01	13,087.19	−1.47
B_{t-1}	182.96	123.95	1.48
	$110.7 < B_{t-1}$ ($\theta = 0.15$)		
Intercept	19,243.59	13,087.42	1.47
B_{t-1}	−181.9602	123.95	−1.47

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. B_t: dependent variable (public debt/GDP ratio). B_{t-1}: threshold variable (15% trimming percentage for threshold search with ordinary standard errors and one thousand bootstrap repetitions). θ : speed of transition parameter.

The LSTAR (2, 1, 1) model specification suffers from similar identification issues. Table 10 reports the estimates from the first-order LSTAR model characterized by the smooth logistic transition function of the form $\Phi[\theta; d_{t-k} - c] = 1/(1 + exp[-\theta(d_{t-k} - c)])$, where θ again measures the speed of transition between the two regimes and c denotes the threshold public debt/GDP ratio. The difference with respect to the ESTAR (2, 1, 1) model is that the transition mechanism is not exponentially smooth but rather s-shaped smooth.

The estimates from Table 10 are much more intuitive and in line with our earlier SETAR estimates. The AR (1) slope coefficient in the lower regime equals one and implies random walk behavior in the case of US public debt/GDP ratio below the statistically significant endogenous threshold of 106.1% of GDP. The AR (1) slope coefficient in the upper public debt regime equals −1.16 and is also statistically significant at the 1% significance level. The sum of the two coefficients across regimes +1 − 1.16 = −0.16 satisfies the global stationarity condition in the case of the LSTAR (2, 1, 1) model and implies a slow mean reversion in the US public debt/GDP ratio between 1947 Q1 and 2024 Q1.

Table 10. LSTAR (2, 1, 1) model for the US public debt/GDP ratio, 1974 Q1–2024 Q1.

Regressors	Coefficients	Standard Errors	t−Stat
	$B_{t-1} \leq 106.11\%$ ($\theta = 20.27$)		
Intercept	0.13	0.23	0.59
B_{t-1}	1.00 ***	0.004	245.04
	$106.11 < B_{t-1}$ ($\theta = 20.27$)		
Intercept	140.52 ***	48.13	2.92
B_{t-1}	−1.16 ***	0.40	−2.93

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. B_t: dependent variable (public debt/GDP ratio). B_{t-1}: threshold variable (15% trimming percentage for threshold search with ordinary standard errors and one thousand bootstrap repetitions). θ : speed of transition parameter.

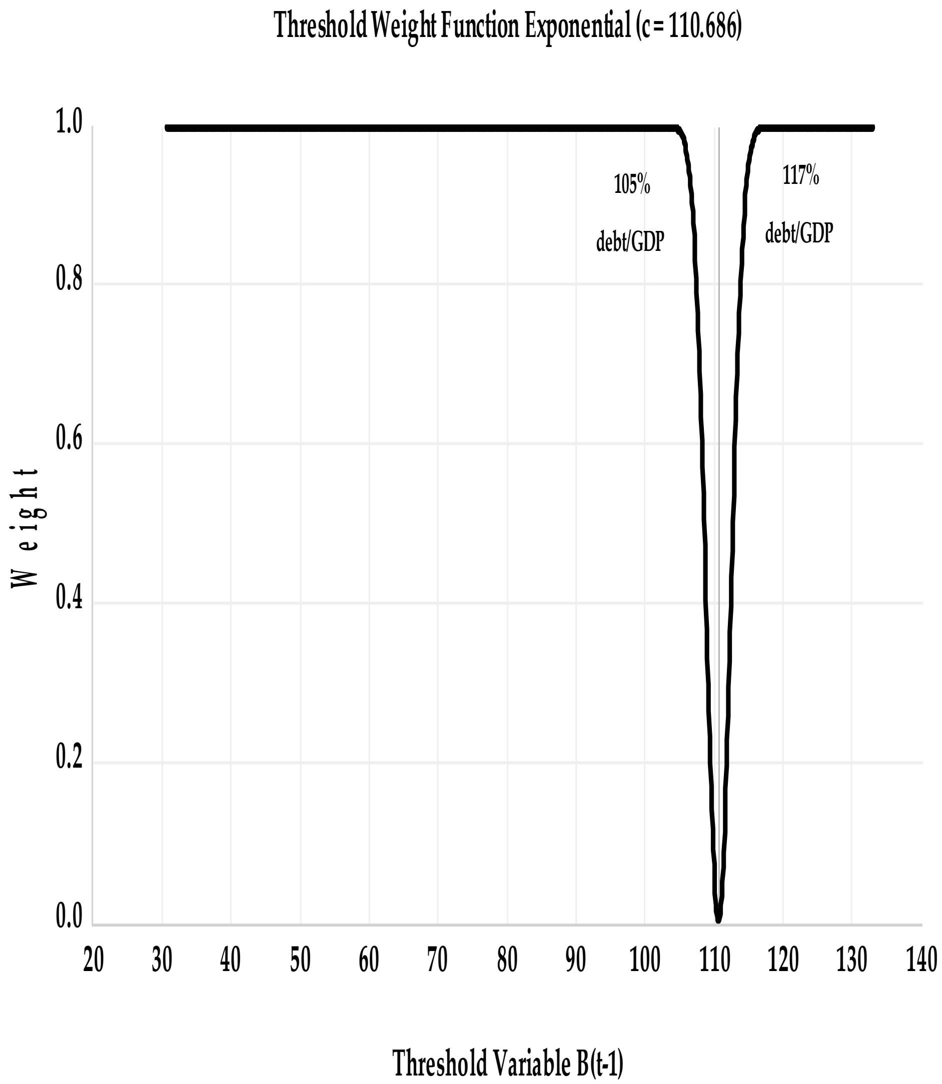

Figure 7. The ESTAR (2, 1, 1) transition function (y-axis) with respect to the threshold variable (x-axis).

The estimate for the speed of transition parameter is large, equaling 20.27, and it is significant at the 1% significance level. Note, however, that for large values of the speed of transition parameter, the logistic s-shaped transition function of the LSTAR (2, 1, 1) model behaves as an 0–1 discrete indicator dummy function of the corresponding SETAR model. More formally, when $\theta \to \infty$, then $\Phi[\theta; d_{t-k} - c] = 1/(1 + exp[-\theta(d_{t-k} - c)]) \to 1$. Figure 8 on the next page illustrates this property of the logistic transition function graphically. In other words, in the case of an extremely high speed of transition, the LSTAR (2, 1, 1) model has the properties of the SETAR (2, 1, 1) model, as Ekner and Nejstgaard (2013) [29] and Gao et al. (2018) [30] have already shown. The logistic transition function in the case of the LSTAR (2, 1, 1) model from Table 10 effectively behaves as a level-shift dummy variable that takes on a value of zero for the lagged public debt/GDP ratio values below the 106.1% threshold and a value of one for the lagged public debt/GDP ratio values above the 106.1% threshold. In other words, the estimated LSTAR (2, 1, 1) model is effectively a SETAR (2, 1, 1) model with the threshold value of 106.1% of GDP, which lies in the vicinity of the already estimated potential threshold values from the SETAR (2, 2, 1), SETAR (2, 2, 2) and SETAR (3, 1, 1) model specifications.

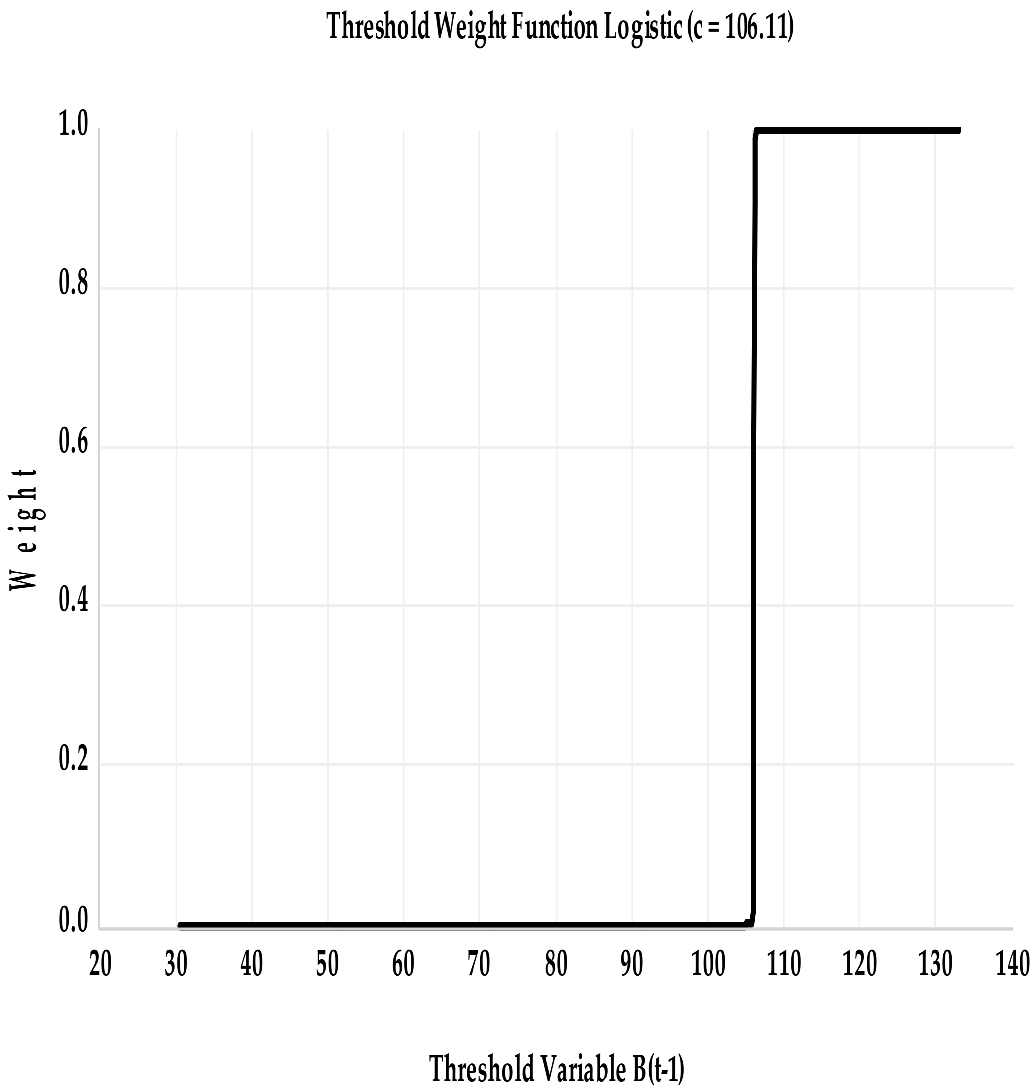

Figure 8. The LSTAR (2, 1, 1) transition function (y-axis) with respect to the threshold variable (x-axis).

4. Concluding Remarks

The goal of this paper is not to provide a definitive answer to the question of whether there is an actual unit root in the US public debt/GDP ratio. Our goal is much more modest: starting from the assumption that there is no actual unit root in the AR polynomial of the US public debt/GDP ratio between 1974 Q1 and 2024 Q1, our aim is only to provide an alternative view, i.e., a commentary, on the US public debt/GDP ratio persistence and nonlinear fiscal adjustment. Given the earlier contributions from the literature that we have reviewed in the introduction to this paper, we are inclined to conclude that the US public debt/GDP ratio is a border line near-unit root case.

As one of the reviewers correctly pointed out, the introductory section of this paper organized earlier contributions from the literature in a less traditional, unorthodox, way. Our primary motivation for such an approach was to familiarize the reader with the fact that the stationarity of the US public debt/GDP ratio is a necessary condition if one wants to try to estimate the SE(TAR)-like models presented in this paper. Starting from an initial assumption of no unit root, for which the evidence is inconclusive and limited, we have presented the estimates from four SETAR model specifications which, in our view, are more suitable in capturing the nonlinear behavior of the US public debt/GDP ratio than the respective ESTAR and LSTAR smooth transition model specifications.

To be consistent with the introductory remarks in this paper, we structure the concluding remarks in a similar fashion by (i) providing first an overview of our stance on why we are reluctant to except the notion of an actual unit root in the US public debt/GDP ratio from an econometric perspective but also (ii) by providing an economic rationale on why we are inclined to believe that the (SE)TAR models are more suitable econometric framework in capturing nonlinear fiscal policy behavior.

4.1. Econometric Perspective

We start by referring to the time series plot of the US public debt/GDP ratio. Campbell et al. (2023) [18] (p. 1) state the following:

We find that the debt-GDP ratio does not behave like a stationary time series in US data since World War II. As Figure 1, Panel a, shows, it has drifted persistently up and down for long periods of time. As one would expect, it shows no upward or downward trend; but it also shows no strong tendency to return to a constant mean.

We agree with the statement by Campbell et al. (2023) [18] above, having in mind, however, that the graphical analysis cannot stand as a substitute for a formal statistical test when deciding whether a time series in question is an actual $I\ (1)$ stochastic process. In other words, one must perform a unit root test to provide statistical evidence in favor of (or against) the unit root null hypothesis. As Chortareas et al. (2008) [33] (p. 649) argue, the choice of a particular alternative hypothesis in unit-root tests affects their ability to reject the null hypothesis. Since we start from the notion that the behavior of the US public debt/GDP ratio is a nonlinear globally stationary SETAR-like stochastic process, the most natural choice would be to opt for the Enders and Granger (1998) [34] unit root test with the alternative hypothesis of stationary discrete threshold adjustment. However, Enders (2001) [35] (p. 261) states that “the power of the tests for TAR adjustment with and without consistent estimates of the threshold are poor compared to that of Dickey-Fuller. The recommendation is to use the Dickey-Fuller test if TAR adjustment is suspected”. The problem with the linear Dickey–Fuller unit root test is, however, that it has notoriously low power in the case of highly persistent AR (1) stochastic processes with the AR (1) coefficient close to one. Following the recommendations of Bec et al. (2022) [37], who investigate the power of unit root tests against nonlinear and non-causal alternatives, we place an emphasis on unit root tests from Elliott et al. (1996), since Bec et al. (2022) [37] (pp. 2–4) argue that these tests have the highest power in the case of roots greater than 0.95 and lesser than one. The results for some of the tests from Elliott et al. (1996) [36], and their modified versions from Ng and Perron (2001) [45], all presented in Table 1, reject the unit root null hypothesis. In order not to rely on just one type of unit root tests, as Chortareas et al. (2008) [33] advise, we have employed complementary unit root tests with endogenous structural break(s). The results of these tests, presented in Tables 2 and 3, unanimously cannot reject the unit root null hypothesis for the US public debt/GDP ratio after the Bretton Woods collapse. In other words, we find only limited evidence against the $I\ (1)$ behavior when we use the unit root testing methodology. An additional problem is, however, that the unit root methodology also has its own share of problems, which we discussed in detail in other parts of this paper. In other words, even if one could claim with 100% certainty that there is no unit root in the US public debt/GDP ratio, theoretical and empirical evidence from other strands of the literature implies that the issue in question cannot be settled by relying only on the unit root testing methodology. Bohn (2007) [15] (p. 1845) summarizes it succinctly:

How serious are these challenges? Conceptually, I consider them a return to normalcy. Since the discovery of unit root testing, the economic (author's emphasis) analysis of debts and deficits has been overshadowed (added emphasis) by the notion that sustainability questions can be answered conclusively by running data through a battery of time series tests (added emphasis).

Bohn (2007) [15] (p. 1846) reiterates the quote from above on the next page of his article by saying that "research on fiscal deficits should focus more on questions of policy identification and stability and on questions of discounting than on testing for unit roots (added emphasis)". But more importantly, Bohn (2007) [15] (p. 1846) ends the article with the statement.

A second strategy is to consider stronger conditions on policy, e.g., upper bounds (added emphasis) on debt motivated by a limited capacity to service debt. Then stationarity in levels (added emphasis) is the most relevant econometric condition, and additional restrictions may apply.

Following Bohn's (2007) [15] advice, our assumption of no actual unit root in the US public debt/GDP ratio, on which we ground the estimates of our baseline globally stationary SETAR (2, 1, 1) model with a partial unit root in the lower regime, might, consistently with the stylized facts of the US public debt/GDP ratio, satisfy the emphasized assertions of Bohn (2007) [15]. With respect to the question of policy identification and stability, the estimated threshold (65% public debt/GDP ratio) produces two regimes: the first one (the lower regime) corresponds to the public debt/GDP dynamics before the GFC, while the second one (the upper regime) corresponds to the public debt/GDP dynamics after the GFC, the most severe recessionary period in the US history after World War II. Moreover, from the perspective of the upper bounds, one can easily compute a conditional long-run mean of the US public debt/GDP ratio for the upper regime by using the corresponding LS estimates $9.42/(1 - 0.92) = 118.38\%$ of GDP. Note that the estimated conditional long-run mean is remarkably close to the DSGE calibrated austerity threshold of 115% public debt/GDP that Elenev et al. (2024) [61] identify as an upper bound above which the public debt/GDP ratio explodes. This finding is also consistent with the remarks of one of the reviewers, for which we are grateful, that one cannot exclude the possibility of a unit root in the upper regime consistent with the notion of threshold stochastic unit roots (TSTUR) model specification, although we deem this scenario less probable for the economic reasons explained below.

4.2. Economic Perspective

If the calibrated values of Elenev et al. (2024) [61] and our post-GFC long-run conditional mean estimates for the US public debt/GDP ratio have some merit, then, given the current value of the US public debt/GDP ratio of 120%, the most import policy consequence of this paper is that US fiscal policymakers should implement an immediate program of fiscal consolidation. The results of D'Erasmo et al. (2015) [25] and Jiang et al. (2024) [19], who quantify a diminishing FRF primary balance response to debt accumulation after the GFC, further strengthen the arguments in favor of an immediate fiscal consolidation. As Balasundharam et al. (2023) [65] document in detail, fiscal consolidations on average last 3–4 years, and for such a brief time span, we believe it is more likely for one to find discrete breaks and thresholds in the public debt/GDP ratio dynamics, especially given the fact that macroeconomic theory does not provide an answer to how big or small the speed of transition parameter θ should be in the case of respective smooth transition model specifications. Consistent with the calibrations of Elenev et al. (2024) [61] that fiscal policymakers react in a discrete fashion to changes in public debt, we would also like to introduce the reader to the assertions of Legrenzi and Milas (2011) [66], who, in turn, reiterate some earlier points of Bertola and Drazen (1993) [67].

All above tests, nevertheless, are implicitly based on a linear model of continuous fiscal adjustment. However, Bertola and Drazen (1993) argue that, due to difficulties in reaching necessary consensus for fiscal retrenchments, fiscal authorities initiate a corrective action only when (added emphasis) the disequilibria reach a given trigger point, for instance when spending reaches levels high enough to be deemed critical. Only in this latter case, the necessary agreement can be reached and adjustment takes place. This suggests the opportunity of allowing for threshold behaviour of fiscal authorities, reacting only when (added emphasis) fiscal variables exceed an endogenously estimated threshold.

Although Legrenzi and Milas (2011) [66] implicitly refer to discrete corrective fiscal actions on two occasions in the passage quoted above, they still opt for the LSTAR-type model specification for the first differenced public debt/GDP ratio in the cases of Greece, Ireland, Italy, Portugal and Spain (GIIPS). However, in the case of Portugal, Legrenzi and Milas (2011) [66] do not obtain statistically significant coefficient estimates. In addition, in the notes to Table 1C, Legrenzi and Milas (2011) [66] (p. 10) report that they impose the value for the speed of transition parameter to obtain the best statistical fit because the precise estimation of this parameter is unlikely due to the extreme sensitivity of the LSTAR likelihood function to the values of the speed of transition parameter. Moreover, in footnote number two, Legrenzi and Milas (2011) [66] (p. 6) acknowledge that they did not obtain convergent estimates in the case of the ESTAR model specification for all the GIIPS economies.

Like Sarno (2001) [22] and Gnegne and Jawadi (2013) [31], Legrenzi and Milas (2011) [66] (p. 4) also use the first differenced public debt/GDP ratio since the results of the linear unit root tests point to a non-stationary public debt behavior in the case of the GIIPS economies. Contrary to the choice of Sarno (2001) [22], Gnegne and Jawadi (2013) [31], and Legrenzi and Milas (2011), who opt for the first differenced public debt/GDP ratio, we opt to work with the levels of the US public debt/GDP ratio. We emphasize two propositions, one from econometric and the other one from economic perspective. First, if there is no unit root in the US public debt/GDP ratio, and one proceeds to use the first difference operator, then it might be the case that the resulting differenced series could suffer from a unit root in the moving average (MA) polynomial due to over-differencing. Admittedly, this scenario is less probable in the case of the US, since the public debt/GDP ratio is a (near) unit root stochastic process. But, in general, the approach might be problematic in the case of economies with a less persistent public debt/GDP ratio. Second, and much more importantly, if there is a unit root in the US public debt/GDP ratio, and one uses the first difference operator, the resulting differenced series is an overall fiscal deficit corrected for the stock-flow adjustments. Now, if such a deficit series stands for a dependent variable in the SETAR/ESTAR/LSTAR framework, then the respective model estimates would quantify a deficit response to nonlinear changes in public debt. However, nonlinear public debt changes are not the only predictor of the deficit variations, so the resulting SETAR/ESTAR/LSTAR estimates will suffer from an omitted variable bias. Such regressions would have to account for, at least, the changes in transitory government spending and business cycle fluctuations—see Barro (1979) [2] and Bohn (1998) [14] on the theoretical foundations of FRFs.

The use of the first differenced public debt/GDP ratio would, hence, be econometrically proper if, and only if, one can be certain of the $I(1)$ nature of the stochastic process for the respective public debt/GDP ratio. We find no such evidence in the literature. In fact, Sarno (2001) [22] (p. 120) assumes, below Equation (1), that the public debt/GDP ratio is stationary and ergodic but still uses the first difference operator—see Chortareas et al. (2008) [33] (p. 650) for further critique. Gnegne and Jawadi (2013) [31] use the first differences given the results of the ZA unit root test, the results of the misspecified KPSS test (test regression with an intercept term only) and without resorting to nonlinear or efficient unit root tests. Legrenzi and Milas (2011) [66] also use the first differences based solely on the results of linear unit root tests.

Since the use of the first differenced debt would not be appropriate even economically, at least in the context of the fiscal policy behavior analyzed in this paper, we believe that the choice to model nonlinearities in the level of the US public debt/GDP ratio could yield potentially relevant theoretical and policy insights. The SETAR model in levels accentuates the importance of the GFC for the regime-specific nonlinear behavior of the US public debt/GDP ratio. Our results show that before the GFC the dynamics of the US public debt/GDP ratio follows a random walk behavior consistent with the tax-smoothing model of Barro (1979) [2]. The tax-smoothing model of Barro (1979) [2] implies a government that tries to smooth taxes, i.e., to minimize the costs of tax collection. The minimization of tax collections by the government can have enhancing welfare effects, at least in terms of the greater predictability and stability of average tax rates. In addition, countercyclical income shocks and pro-cyclical transitory government spending drive the variations in the US public debt/GDP ratio. After the GFC structural break, however, the dynamics of the public debt/GDP ratio show a highly persistent upward trend. As Jiang et al. (2024) [19] argue, the highly persistent innovations in the US public debt/GDP ratio after the GFC are due to (i) the inelastic demand of foreign investment community for US Treasury bonds because of their perception that the US government is a "manufacturer" of risk-free assets; (ii) the FED's inelastic quantitative easing asset purchases after 2008, which we potentially quantify both through level-shifts in the baseline SETAR (2, 1, 1) model, but also through the COVID-19 residual variance stabilization presented in Figure 4; and (iii) finally, biased subjective expectations and beliefs on the part of bond investors, which can be, to a certain extent, a consequence of the FED's actions. All three mechanisms imply that the US government does not necessarily have to resort to fiscal consolidation measures during recessions but can instead countercyclically "manufacture" risk-free debt. The manufacturing of risk-free debt actually "insures" taxpayers in the short run since, even in recessions, the US government does not need to resort to tax hikes and spending cuts to cover the cyclical deficit due to an aggregate recessionary output shock. The crucial question is how long can the US government imply such fiscal practices? Jiang et al. (2022) [59] show that the higher the persistence of the US public debt/GDP ratio, the longer the insurance horizon. In other words, the near-unit root behavior of the US public debt/GDP ratio after the GFC, due to the FED's actions, perceptions of international investment community and potentially biased beliefs of bond investors, enabled the US government to postpone the implementation of fiscal austerity measures after 2008 up to now. A high AR (1) slope coefficient for the US public debt/GDP ratio after the GFC could imply the absence of a positive primary balance response, a feature captured by the FRFs estimated in D'Erasmo et al. (2015) [25]. The lack of a positive primary balance response could show the lack of effort of US fiscal policymakers to implement credible fiscal consolidation measures in the years to come. But the longer the US government provides such an "insurance" to the taxpayers, the more fiscal problems accumulate in the back. Jiang et al. (2019) [19] report that the public debt/GDP ratio predicts higher inflation in the US after the GFC. The logic behind this result is that the increase in inflation, not primary surpluses, restores the intertemporal government budget constraint identity (In addition). In other words, Jiang et al. (2022) [59] show that the short-run "insurance" of taxpayers is not sustainable in the long run and the US Treasury would have to resort to fiscal austerity measures to curb the countercyclical public debt growth in order to prevent the default of the US government. We are inclined, hence, to believe that the fiscal consolidation package or default would enforce yet another break, i.e., a threshold, rather than a stochastic unit root in the upper public debt/GDP ratio regime, although we cannot exclude such a future scenario with absolute certainty.

The lack of attitude toward the implementation of credible fiscal adjustment program(s) is particularly worrisome in times of secular stagnation and after the GFC and COVID-19 pandemic shocks. In other words, even though the US government relied on a growth dividend before 2008, i.e., it took advantage of the fact that, on average, the real interest rates on US government bonds were below the real rates of growth (negative $r - g$ differential), there is no guarantee that such a trend will extrapolate into the future. In some states of nature, and in some periods of time, it will be the case that aggregate shocks would not work in favour of the negative interest rate-growth rate differential. The absence of discretionary fiscal measures in such scenarios could imply, according to the fiscal theory of the price level, an elevated price increase to ensure the workings of the intertemporal budget constraint of the government, especially in today's world, when the FED and the Treasury could not rely on the unconventional fiscal and monetary policy measures that accommodated the World War II spike in the US public debt/GDP ratio. In such circumstances, it becomes questionable for how long the international investor community will regard US Treasury bonds as risk-free, so the US government, as was the case with the UK government in the past, can in fact lose its exorbitant privilege and face difficulties in financing already high and growing spending needs. The absence of credible austerity measures might then impute an actual unit or explosive root in the US public debt/GDP ratio.

Supplementary Materials: The following supporting zip file with data, figures, tables, GRETL, STATA, EViews and RATS codes for reproducing the results from this paper is available at: https://www.mdpi.com/article/10.3390/math12203250/s1.

Author Contributions: V.A., D.B. and M.D.: investigation and conceptualization; V.A.: methodology and software; V.A.: writing—original draft preparation; D.B. and M.D.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The Ministry of Science, Technological Development, and Innovation of the Republic of Serbia funded this research under contract number 451-03-47/2023-01/200005.

Data Availability Statement: The data used in this study are from the US Office of Management and Budget and Federal Reserve Bank of St. Louis, Federal Debt: Total Public Debt as Percent of Gross Domestic Product [GFDEGDQ188S], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GFDEGDQ188S, accessed on 23 July 2024.

Acknowledgments: We thank the corresponding editor for the exceptionally professional handling of the manuscript. We have also received help from the detailed and constructive comments of the two anonymous reviewers. We bare sole responsibility for all the remaining errors. The corresponding author gratefully acknowledges the use of services and facilities of the European Center for Research in Economics and Business (ECREB) at the Faculty of Economics and Business Administration, West University of Timisoara, Timisoara, Romania, where he performed part of the work on this paper. The author especially thanks Mihai Mutascu, Bogdan Ianc, Stefana Maria Dima, Bogdan Dima, Claudiu Botoc, Gratiela Noja, Cosmin Enache and Gabriela Mircea for their help and hospitality during his research visit to the ECREB.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Hansen, B.E. Threshold autoregression in economics. *Stat. Interface* **2011**, *4*, 123–127. [CrossRef]
2. Barro, R.J. On the determination of the public debt. *J. Polit. Econ.* **1979**, *87*, 940–971. [CrossRef]
3. Bec, F.; Ben Salem, M.; Carrasco, M. Tests for unit-root versus threshold specification with an application to the purchasing power parity relationship. *J. Bus. Econ. Stat.* **2004**, *22*, 382–395. [CrossRef]
4. Hansen, B.E. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* **1996**, *64*, 413–430. [CrossRef]
5. Hansen, B.E. Inference in TAR models. *Stud. Nonlinear Dyn. Econom.* **1997**, *2*, 1–14. [CrossRef]
6. Hansen, B.E. Regression kink with an unknown threshold. *J. Bus. Econ. Stat.* **2017**, *35*, 228–240. [CrossRef]
7. González, M.; Gonzalo, J. *Threshold Unit Root Models*; DES Working Paper, No. 97–50; Departamento de Estadística y Econometría, Universidad Carlos III de Madrid: Madrid, Spain, 1997.
8. Gonzalo, J.; Pitarakis, J. Estimation and model selection-based inference in single and multiple threshold models. *J. Econom.* **2002**, *110*, 319–352. [CrossRef]
9. Hamilton, J.D.; Flavin, M.A. On the limitations of government borrowing: A framework for empirical testing. *Am. Econ. Rev.* **1986**, *76*, 808–819.

10. Dickey, D.A.; Fuller, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **1981**, *49*, 1057–1072. [CrossRef]
11. Kremers, J.J. Long-run limits on the U.S. federal debt. *Econ. Lett.* **1988**, *28*, 259–262. [CrossRef]
12. Kremers, J.J.U.S. federal indebtedness and the conduct of fiscal policy. *J. Monet. Econ.* **1989**, *23*, 219–238. [CrossRef]
13. Wilcox, D.W. The sustainability of government deficits: Implications of the present-value borrowing constraint. *J. Money Credit Bank.* **1989**, *21*, 291–306. [CrossRef]
14. Bohn, H. The behavior of U.S. public debt and deficits. *Q. J. Econ.* **1998**, *113*, 949–963. [CrossRef]
15. Bohn, H. Are stationarity and cointegration restrictions really necessary for the intertemporal budget constraint? *J. Monet. Econ.* **2007**, *54*, 1837–1847. [CrossRef]
16. Cochrane, J.H. *The Value of Government Debt*; NBER Working Paper, No. 26090; National Bureau of Economic Research: Cambridge, MA, USA, 2020.
17. Cochrane, J.H. The fiscal roots of inflation. *Rev. Econ. Dyn.* **2022**, *45*, 22–40. [CrossRef]
18. Campbell, J.Y.; Gao, C.; Martin, I.W.R. *Debt and Deficits: Fiscal Analysis with Stationary Ratios*; NBER Working Paper, No. 31224; National Bureau of Economic Research: Cambridge, MA, USA, 2023.
19. Jiang, Z.; Lustig, H.; Van Nieuwerburgh, S.; Xiaolan, M. What drives variation in the U.S. debt-to-output ratio: The dogs that did not bark. *J. Financ.* **2024**, *79*, 2603–2665. [CrossRef]
20. Carrasco, M. Misspecified structural change, threshold, and Markov-switching models. *J. Econom.* **2002**, *109*, 239–273. [CrossRef]
21. Lanne, M.; Saikkonen, P. Threshold autoregression for strongly autocorrelated time series. *J. Bus. Econ. Stat.* **2002**, *20*, 282–289. [CrossRef]
22. Sarno, L. The behavior of U.S. public debt: A nonlinear perspective. *Econ. Lett.* **2001**, *74*, 119–125.
23. Mendoza, E.G.; Ostry, J.D. International evidence on fiscal solvency: Is fiscal policy responsible? *J. Monet. Econ.* **2008**, *55*, 1080–1094. [CrossRef]
24. Mauro, P.; Romeu, R.; Binder, A.; Zaman, A. A modern history of fiscal prudence and profligacy. *J. Monet. Econ.* **2015**, *76*, 55–70. [CrossRef]
25. D'Erasmo, P.; Mendoza, E.; Zhang, J. *What Is a Sustainable Public Debt?* NBER Working Paper, No. 21574; National Bureau of Economic Research: Cambridge, MA, USA, 2015.
26. Ghosh, A.R.; Kim, J.I.; Mendoza, E.G.; Ostry, J.D.; Qureshi, M.S. Fiscal fatigue, fiscal space, and debt sustainability in advanced economies. *Econ. J.* **2013**, *123*, F4–F30. [CrossRef]
27. Heinen, A.; Michael, J.; Sibbertsen, P. Weak identification in the ESTAR model and a new model. *J. Time Ser. Anal.* **2013**, *34*, 238–261. [CrossRef]
28. Buncic, D. Identification and estimation issues in exponential smooth transition autoregressive models. *Oxf. Bull. Econ. Stat.* **2019**, *81*, 667–685. [CrossRef]
29. Ekner, L.E.; Nejstgaard, E. *Parameter Identification in the Logistic STAR Model*; Discussion Paper, No. 13–07; Department of Economics, University of Copenhagen: Copenhagen, Denmark, 2013.
30. Gao, Z.; Ling, S.; Tong, H. Testing for TAR Models vs. STAR Models—A Separate Family of Hypotheses Approach. *Stat. Sinica.* **2018**, *28*, 2857–2883.
31. Gnegne, Y.; Jawadi, F. Boundedness and nonlinearities in public debt dynamics: A TAR assessment. *Econ. Modell.* **2013**, *34*, 154–160. [CrossRef]
32. Zivot, E.; Andrews, D. Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *J. Bus. Econ. Stat.* **1992**, *10*, 251–270. [CrossRef]
33. Chortareas, G.; Kapetanios, G.; Uctum, M. Nonlinear alternatives to unit root tests and public finance sustainability: Some evidence from Latin American and Caribbean countries. *Oxf. Bull. Econ. Stat.* **2008**, *70*, 645–663. [CrossRef]
34. Enders, W.; Granger, C.W.J. Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. *J. Bus. Econ. Stat.* **1998**, *16*, 304–311. [CrossRef]
35. Enders, W. Improved critical values for the Enders-Granger unit-root test. *App. Econ. Lett.* **2001**, *8*, 257–261. [CrossRef]
36. Elliott, G.; Rothenberg, T.J.; Stock, J.H. Efficient tests for an autoregressive unit root. *Econometrica* **1996**, *64*, 813–836. [CrossRef]
37. Bec, F.; Guay, A.; Bohn Nielsen, H.; Saïdi, S. *Power of Unit Root Tests against Nonlinear and Noncausal Alternatives*; THEMA Working Paper, No. 2022-14; CY Cergy Paris Université: Cergy-Pontoise, France, 2022.
38. Kapetanios, G.; Shin, Y. Unit root tests in three-regime SETAR models. *Econom. J.* **2006**, *9*, 252–278. [CrossRef]
39. Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root. *J. Econom.* **1992**, *54*, 159–178. [CrossRef]
40. Badia, M.M.; Medas, P.; Gupta, P.; Xiang, Y. Debt is not free. *J. Int. Money Financ.* **2022**, *127*, 102654. [CrossRef]
41. Tsay, R.S. Testing, and modeling threshold autoregressive processes. *J. Am. Stat. Assoc.* **1989**, *84*, 231–240. [CrossRef]
42. Acalin, J.; Ball, L.M. *Did the U.S. Really Grow Out of Its World War II Debt?* NBER Working Paper, No. 31577; National Bureau of Economic Research: Cambridge, MA, USA, 2023.
43. Bohn, H. *The Sustainability of Fiscal Policy in the United States*; CESifo Working Paper, No. 1446; CESifo: Munich, Germany, 2005.
44. Gale, W. Comment on 'Measuring US Fiscal Capacity Using Discounted Cash Flow Analysis'. *BPEA* **2022**, *Fall*, 210–215.
45. Ng, S.; Perron, P. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* **2001**, *69*, 1519–1554. [CrossRef]

46. Vogelsang, T.; Perron, P. Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time. *Int. Ecom. Rev.* **1998**, *39*, 1073–1100. [CrossRef]
47. Vogelsang, T. Unpublished Computer Program. 1993. Available online: https://scholar.google.com.au/scholar?hl=en&as_sdt=0,5&q=Vogelsang,+T.+Unpublished+computer+program++,+1993.+&btnG= (accessed on 31 August 2024).
48. Lee, J.; Strazicich, M. Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks. *REStat* **2003**, *85*, 1082–1089. [CrossRef]
49. Bai, J.; Perron, P. Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* **1998**, *66*, 47–78. [CrossRef]
50. Stambaugh, R. Predictive Regressions. *J. Financ. Econ.* **1999**, *54*, 375–421. [CrossRef]
51. Aiyagari, S.R.; Marcet, A.; Sargent, T.J.; Seppala, J. Optimal taxation without state-contingent debt. *J. Polit. Econ.* **2002**, *110*, 1220–1254. [CrossRef]
52. Lucas, R.; Stokey, N. Optimal Fiscal and Monetary Policy in an Economy without Capital. *J. Monet. Econ.* **1983**, *12*, 55–93. [CrossRef]
53. Tsay, R. *Analysis of Financial Time Series*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2005; pp. 191–192.
54. Engle, R.; Granger, C. Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* **1987**, *55*, 251–276. [CrossRef]
55. Yoon, G. War and peace: Explosive U.S. public debt, 1791–2009. *Econ. Lett.* **2012**, *115*, 1–3. [CrossRef]
56. Bhandari, A.; Evans, J.; Golosov, M.; Sargent, T.J. Fiscal policy and debt management with incomplete markets. *Q. J. Econ.* **2017**, *132*, 617–663. [CrossRef]
57. Hansen, B. Testing for Parameter Instability in Linear Models. *J. Policy Model.* **1992**, *14*, 517–533. [CrossRef]
58. Jiang, Z.; Lustig, H.; Van Nieuwerburgh, S.; Xiaolan, M.Z. *Quantifying U.S. Treasury Investor Optimism*; Stanford Business School Working Paper, No. 3931; Stanford University: Stanford, CA, USA, 2021.
59. Jiang, Z.; Lustig, H.; Van Nieuwerburgh, S.; Xiaolan, M.Z. *Manufacturing Risk-Free Government Debt*; NBER Working Paper, No. 27786; National Bureau of Economic Research: Cambridge, MA, USA, 2022.
60. Jiang, Z.; Lustig, H.; Van Nieuwerburgh, S.; Xiaolan, M. Fiscal capacity: An asset pricing perspective. *Annu. Rev. Financ. Econ.* **2023**, *15*, 197–219. [CrossRef]
61. Elenev, V.; Landvoigt, T.; Shultz, P.J.; Van Nieuwerburgh, S. *Can Monetary Policy Create Fiscal Capacity?* NBER Working Paper, No. 29129; National Bureau of Economic Research: Cambridge, MA, USA, 2021.
62. Cochrane, J. Long-Term Debt and Optimal Policy in the Fiscal Theory of the Price Level. *Econometrica* **2001**, *69*, 69–116. [CrossRef]
63. Psaradakis, Z.; Sola, M.; Spagnolo, F.; Spagnolo, N. Selecting Nonlinear Time Series Models Using Information Criteria. *J. Time Ser. Anal.* **2009**, *30*, 369–394. [CrossRef]
64. Taylor, M.; Peel, D.; Sarno, L. Non-linear Mean-Reversion in Real Exchange Rates: Toward a Solution to the Purchasing Power Parity Puzzles. *Int. Ecom. Rev.* **2001**, *42*, 1015–1042. [CrossRef]
65. Balasundharam, V.; Basdevant, O.; Benicio, D.; Ceber, A.; Kim, Y.; Mazzone, L.; Selim, H.; Yang, Y. *Fiscal Consolidation: Taking Stock of Success Factors, Impact, and Design*; IMF Working Paper, No. 23/63; International Monetary Fund: Washington, DC, USA, 2023.
66. Legrenzi, G.; Milas, K. *Debt Sustainability and Financial Crises: Evidence from the GIIPS*; CESIfo Working Paper, No. 11/3594; CESifo: Munich, Germany, 2011.
67. Bertola, G.; Drazen, A. Trigger Points and Budget Cuts: Explaining the Effects of Fiscal Austerity. *AER* **1993**, *83*, 11–26.