Department of Economics
Athens University of Economics and Business

WORKING PAPER no. 18-2021

Exploring Okun's Law Asymmetry:
An Endogenous Threshold LSTR Approach

Dimitris Christopoulos, Peter McAdam and Elias Tzavalis

May 2021

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

Exploring Okun's Law Asymmetry: An Endogenous Threshold LSTR Approach*

Dimitris Christopoulos[a], Peter McAdam[b] and Elias Tzavalis[a]

[a] Athens University of Economics & Business; [b] European Central Bank

June 2, 2021

Abstract

Using a novel methodology, we offer new evidence that a threshold relationship exists for Okun's law (the well-known output-unemployment co-movement). We use a logistic smoothed transition regression (LSTR) model where threshold endogeneity is addressed using copulas. We also suggest a new test of the linearity hypothesis against the LSTR model. In line with Okun's insight (and that of the subsequent literature) that the trade off can be affected by different margins, we consider several potential threshold variables. Of these the labor share, the policy rate and the shadow rate appear to robustly reveal threshold effects in the Okun's parameter in the US in recent decades. This conclusion is bolstered by combing these threshold candidates into a single factor. Accordingly, we find that the unemployment gap is increasingly associated with a smaller output gap. Notably, whilst the Great Recession accelerated that rise, the bulk of the change occurred beforehand. Word Count (excluding online appendices: 9,478 words.)

Keywords: unemployment, output, asymmetries, logistic transition, structural threshold regression, endogeneity, copula, Monte Carlo, test for linearity.

JEL: C24, C46, E23, E24.

*We thank John Fernald (discussant) as well as Costas Azariadis, Larry Ball, Clint Cummins, Nicola Fuchs-Schündeln, Carlos Montes-Galdon, Isabel Hanisch, Michele Lenza, Gabriel Pérez-Quirós, Carmen Reinhart, Yves Schüler, Anders Warne, and numerous seminar audiences for helpful comments and discussions. The opinions expressed in this paper are those of the authors and do not necessarily reflect views of the European Central Bank or the Eurosystem. Dimitris Christopoulos and Elias Tzavalis acknowledge support for this research by the Hellenic Foundation for Research and Innovation under the 'First call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant' (project no: HFRI-FM17-3532).

1 Introduction

Okun's law (hereafter, OL) relates to an empirically observed comovement between cyclical unemployment and output (Okun, 1962). Despite its apparent simplicity, it has generated a vast, enduring literature, and is widely used among policy makers and applied economists.[1] The relationship provides a link between the labor and goods market over the business cycle, and is often considered a key empirical regularity.[2] It is a core part of many models, where the aggregate supply function is derived from combining OL with the Phillips curve, Mankiw (2015). In terms of policy, OL addresses the issue of how much output is 'lost' when unemployment exceeds its natural or trend rate, plus, it demonstrates that the effectiveness of disinflation policy depends on the responsiveness of unemployment to the rate of output growth.

One common and compelling criticism of OL in the literature, though, is the assumption of linearity. Many studies instead suggest that it is characterized by nonlinearities and asymmetries.[3] A nonlinear asymmetric OL would be an important finding for applied economists. It may affect other recognised economic relationships such as the price and wage Philips curves. Moreover, in the policy realm, it would suggest that the effectiveness (and required 'size') of stabilization policy on the real economy would depend in which 'regime' Okun's relationship then lies. It may also have a bearing on forecasting properties of the relationship.

We use a novel methodology to assess the case for asymmetry. Our econometric framework is an augmented version of the logistic smooth transition regression (LSTR) model.[4] The LSTR model nests the linear and standard threshold specification for 'low' and 'high' values of its identified regimes. The threshold variable allows the model to classify the data into two or more groups of observations. It captures smooth transition across regimes, which may be more reasonable in macroeconomics due to various adjustment mechanisms and frictions (in contrast to standard markov-switching models which imply a sharp switch).

Notably, though, we depart from the bulk of the LSTR literature since we allow the threshold variable to be *endogenous* (i.e., contemporaneously correlated with the disturbance term of the regression).[5] Indeed, in as fundamental and deep-seated a relationship as OL, this is intuitive; if there are asymmetries, then it is likely that, rather than be exogenous, they should arise from the workings of the economy itself – i.e., from the operation of booms and busts, or from policy shifts or more persistent structural changes. In so far as the threshold is endogenous, our empirical methodology explicitly seeks to capture that endogeneity. Failure to do so leads to inconsistent model estimates.

To deal with this problem, motivated by the recent work of Kourtellos, Stengos and Tan (2016) (hereafter KST), this approach includes bias correction terms conditional on each regime of the model to account for the endogeneity problem. However, instead of using the inverse Mills ratios (which assumes normality of the threshold variable), we draw on the recently suggested method of

[1] See Perman et al. (2015). Okun's law has been regularly referred to as a policy benchmark and used to appraise macroeconomic conditions, see Bernanke (2012).

[2] To cite Blinder (1997, p241): "... [A] truly sturdy empirical regularity [is] Okun's Law ... it closes the loop between real output growth and changes in unemployment with *stunning reliability*" (emphasis added).

[3] Perman et al. (2015) provides a meta study of the literature of the variety of approaches followed in the literature.

[4] See van Dijk et al., 2002 for a survey.

[5] Kourtellos, Stengos and Tan (2016) call this the "Structural Threshold Regression". As they discuss, the assumption of threshold exogeneity undermines the practical usefulness of threshold regression models, since many plausible threshold variables (their examples include modelling trade shares, political risk) are very likely to be endogenous to the process under consideration.

Christopoulos et al (2021) which calculates the bias correction terms based on the *copula* approach (which does not necessarily require normality of the threshold variable)[6]: the bias correction terms are then given by copula-based transformations of the threshold variable for each regime of the model truncated at the location parameter value.[7]

The copula approach has a number of interesting properties to deal with the problem of the endogeneity of the threshold variable in the framework of the LSTR model. It can capture the dependence between the threshold variable and the error term through the copula transformed terms (variables). The marginal distributions of the transformed variables can be estimated based on a non-parametric or density estimation procedure, in a first step. This makes our approach quite general and flexible. Indeed, we avoid the problem of "weak" instruments for the threshold variable in the case where one would like to estimate the model parameters based on an IV method. The suggested method can also allow for different variances of the LSTR regression disturbance term and its correlation structure with the threshold variable across the two regimes. Furthermore, to test for smooth transition effects in the data, we enhance our approach by suggesting a new likelihood ratio (LR) test to detect linearity against LSTR effects under threshold endogeneity. Both the estimation methodology of our approach and the power performance of the LR test are evaluated through a Monte Carlo (MC) study.

The paper is organized as follows. Section 2 discusses and the motivates the dynamic 'gap' version of the OL. Section 3 assesses in broad terms how well OL fits the data. These considerations establish its importance for applied economists and econometricians. Section 4 formally introduces the LSTR model, and presents our approach to adjust the model for the endogeneity of the threshold variables, based on copula theory. The adjusted model is estimated using a two-step concentrated nonlinear least squares method. The method is assessed in a Monte Carlo (MC) study in the appendix. We consider the cases that the threshold variable and disturbance term follow a normal as well as a Student-t distribution. The former allows us to make comparisons to KST's approach (as we have adapted it to the LSTR framework). Generating data from the Student *t* distribution will show the robustness of the method to fat tails of threshold variables often met in practice. Moreover, we also describe and motivate our new linearity testing procedure to test for smooth transition threshold effects against linearity under threshold endogeneity.

Section 5 explores threshold choices within the OL framework. Rather than impose a single preferred threshold variable, we pursue a multivariate approach. This is consistent with Okun's observation that *many factors* may affect the nature and strength of the trade off, in a possibly time-varying manner. Plus it is also consistent with much modern econometric discussion (see the seminal contribution of Stock and Watson, 2002) that instead combines and examines data information from many sources. This philosophy is also arguably more appealing than the researcher reporting a single threshold variable with no information provided as to the variable selection process.

Specifically, we test three broad categories drawn from the literature, classified as (i) *Demand & Cyclical* pressures, (ii) *Structural* features of the economy, and finally, (iii) *Policy & Financial* vari-

[6] See Patton (2006) for a discussion of, and an economic application of, copula methods.

[7] Moreover, to mitigate the effects of the normality assumption on the estimates of the model, Kourtellos et al (2021) use a semi-parametric approach.

ables.[8] As befits a quasi-structural relationship[9], there may be many feasible threshold candidates rather than merely a sparse outcome; in effect candidates which may reinforce or counteract one another's effects.[10] Accordingly, we also examine the performance of a composite indicator which can summarize common information of the most relevant threshold candidates.

That the Okun parameter is shaped by many different channels is something Okun himself recognized, Okun (1974). He discussed cyclical features (e.g., changes in hours worked and productivity and business cycles), changes in macroeconomics policy as well as more supply-side influences (changes in the labor force and labor supply). Our paper can be seen as an attempt to revive these additional channels envisaged by Okun and, where feasible, to discriminate between them.

Given this background, Section 6 estimates OL assessing the selection of the threshold candidates that indicated nonlinearity in the results of the previous section. These are mostly to be found in the last two categories (i.e., structural, and policy). We also plot the resulting transition probabilities alongside the threshold as a visual plausibility test. In Section 7, we then combine the most relevant threshold candidates into a composite index. We use this index as the threshold variable and re-estimate the endogenous threshold model. Section 8 concludes. Additional material is given in appendices.

2 Okun's Law: Overview and Margins

OL refers to an inverse business-cycle contemporaneous relationship between the unemployment rate (u) and real output (Y). We model this in a dynamic 'gap' form:[11]

$$y_t = \rho y_{t-1} + \beta u_t + \varepsilon_t \tag{1}$$

where $y_t = \tilde{y}_t - y_t^*$ is the cyclical component of real GDP – i.e., the deviation of log real GDP, $\tilde{y}_t$, from log potential real GDP, y^*. Likewise, the cyclical component of unemployment, $u_t = \tilde{u}_t - u_t^*$, is the deviation of the unemployment rate, $\tilde{u}_t$, from its 'trend', u_t^*. Relationship (1) captures the contemporaneous effect of u_t on y_t; y_{t-1} is included to capture dynamic adjustments due to lagged

[8] In past empirical studies of asymmetries in OL, a single threshold has been used, rather than (as here) allowance for a variety of potential channels. This focus may have led to a selection (or publication) bias in reported empirical studies given that researchers and journals have limited incentive to publish results where a prescribed threshold variable has no impact. See Perman et al. (2015) for an interesting discussion of publication biases in the context of OL.

[9] Prachowny (1993) argued that OL had a structural interpretation in the context of the production function approach. We choose to work with the simpler form, since it is the most common representation in the literature (and thus admits easy benchmarking). Moreover, although the appeal of the (production-function) approach is due to its inclusion of additional factor and factor utilization margins, many of the proposed controls (in cyclical form) such as hours worked are so heavily correlated with unemployment as to potentially dilute any intended structural interpretation of the parameters. Moreover, the Prachowny derivation is a special case assuming that aggregate production is Cobb Douglas and that the capital stock and a disembodied technology factor are always at their long-run levels (characteristics which are counter factual in the US aggregate data, see Klump et al., 2007).

[10] For an analysis of sparse and dense models, see Giannone et al. (2019).

[11] An additional 'first differences' form of OL is sometimes also used:

$$\Delta y_t = \beta_0 + \rho \Delta y_{t-1} + \beta \Delta u_t + \eta_t$$

where $\Delta y_t = \log(Y_t) - \log(Y_{t-1})$, $\Delta u_t = u_t - u_{t-1}$. Of the two OL forms, the gap model requires us to capture latent trends and is related more to understanding business-cycle trade offs. It has the advantage of taking into account the state of the economy relative to its trend or natural rate. The difference form posits a linear relationship between the first difference of the log of output and the first difference of the unemployment rate. The two versions are equivalent if potential growth and the natural rate of unemployment are constant. Since this is unlikely to hold, the gap version appears preferable (and, accordingly, is the one we emphasize).

unemployment or output own shocks.[12] Parameter β is the impact Okun coefficient (expected to be negative reflecting the trade off), and $\beta^* = \beta/(1-\rho)$ is its long-run equivalent. Term ε_t denotes a stochastic disturbance. For a given ρ, the higher is $|\beta|$, the *steeper* is Okun's relationship and thus the higher the output costs of a rise in cyclical unemployment.

As discussed, we can think of the OL coefficient as reflecting a number of margins of adjustment. Later in the paper, we explore the idea that such margins may be time varying and subject to threshold behavior. Here however as a preliminary step we can consider a simple formalization based around a production function. This is not meant to generate hypotheses per se, but more to anchor ideas for reflecting on the existence of different margins.

Consider the production function:

$$Y = \Gamma \times \mathbf{F}(k \times z_k \ , \ n \times h \times z_l) \tag{2}$$

where Y is output, k is the quantity of the capital stock, z_k is the utilization rate of capital and z_l represents usage (or hoarding) of the labor input, n represents the number of workers, h is hours worked, Γ represents Hicks neutral technical progress, and $\mathbf{F}$ represents some smooth and concave mapping function.[13]

Defining $\dot{y} = \dot{Y}/Y$ (akin to a continuous-time cyclical gap) and noting that growth in unemployment is the gap between the growth in the labor force and the working population, $\dot{u} = \dot{l}/l - \dot{n}/n$, we can re-express (2) as,

$$\dot{y} = \alpha\left(\frac{\dot{k}}{k} + \frac{\dot{z_k}}{z_k}\right) + (1-\alpha)\left(\frac{\dot{h}}{h} + \frac{\dot{l}}{l} - \dot{u} + \frac{\dot{z_l}}{z_l}\right) + \frac{\dot{\Gamma}}{\Gamma} \tag{3}$$

where $\alpha = \delta F/\delta K \times K/F \in (0,1)$ is the elasticity of output with respect to the capital input (reflecting the capital share of income), assuming constant returns. Noting variable-specific cyclicality admits a more compact OL framework: $|cor(\Xi, \dot{y})| = |\beta_\Xi| \geq 0$; $\Xi \in \{\dot{k}/k, \dot{z_i}/z_i, \dot{n}/n, \dot{h}/h, \dot{\Gamma}/\Gamma\}$:[14]

$$\dot{y} = \beta \times \dot{u} \tag{4}$$

where $\beta(\alpha, \beta_\Xi) = (1-\alpha) \times \left(\alpha(\beta_k + \beta_{z_k}) + (1-\alpha)(\beta_h + \beta_l + \beta_{z_l}) + \beta_\Gamma - 1\right)^{-1}$ is the Okun coefficient.

Accordingly, we can see that the OL can be written in a way that emphasizes the strengths of the various possible adjustment mechanisms (as recognized by Okun himself). Moreover, if these margins are time varying or subject to regime shifts, then the Okun parameter β should reflect these too. Indeed, even the production function "primitives", $\alpha, 1-\alpha$ may be time varying (as corroborated in various shifts in income shares in the major economies.)

3 How Well Does the basic Okun's Law fit the data?

To derive the cyclical component of the series to estimate the gap form of OL, we explored several different filters common in the literature. In many cases, they gave a relatively similar picture but we chose the asymmetric Christiano-Fitzgerald one since, for our series of interest, it identified the

[12] Following the bulk of the literature, we add a dynamic term. Thus the equation captures not only the contemporaneous correlation between changes in the unemployment rate and real output growth, but dynamic lagged effects.

[13] see León-Ledesma et al., 2010) for further discussion of production function types and identified technical progress.

[14] If $\beta_\Xi > 0$ then the appropriate element in Ξ is pro-cyclical with output growth; counter-cyclical < 0 or a-cyclical $= 0$ otherwise.

appropriate frequencies relatively well – i.e., when approximating the ideal band-pass filter over the standard Burns-Mitchell business-cycle frequency of $6 - 32$ quarters (for further discussion of this see Appendix D).

Estimates of OL are shown in Table 1 in static and dynamic form. In terms of the fit metrics R^2_{adj} and AIC (Akaike information criterion), the dynamic model dominates both the static OL model and a simple AR1 benchmark for y_t; the inclusion of u_t into the AR model absorbs a substantial part of the autoregressive persistence of y_t.

Results suggest a unit increase of cyclical unemployment is associated with a decline of just under 2% in output (in cyclical terms). Consistent with Okun's original work, most US studies locate the slope coefficient $|\beta|$ in a $(1, 3]$ interval. That the coefficient typically exceeds one in absolute value (i.e., cyclical output drops *by more* than the increase in cyclical unemployment) reflects different adjustment margins that can amplify movements in unemployment on output: e.g., some unemployed may cease job search thus contracting the labor force, labor productivity may fall (reflecting labor hoarding), hours worked may fall, and the economy may weaken through the normal Keynesian spending multiplier (associated to lower demand).[15]

Figure 1 shows both filtered series, and rolling (75 quarters) sample window regression estimates of β over 1950q1-2018q4. The basic fit is apparent: the cyclical turning points (as highlighted by the NBER recession dates) tends to be well aligned between cyclical unemployment and output. Some basic takeaways of asymmetry are revealed by the plots. For example negative output gaps are generally deeper and more abrupt than positive ones (see also Rothman, 1998). The gaps can also be quite changeable: in the first half (or at least second third) of the sample, output volatility far exceeded that of unemployment, but became closer in the subsequent decades. Finally, the rolling window regression analysis suggests that the Okun slope coefficient has been rising above its full-sample estimate from the mid 1990s.[16] Thus whilst OL fits the data well, it does so somewhat imperfectly: even on a cursory inspection, the Okun parameter appears to be shifting over time. However, to address this rigorously, we clearly need to go beyond simple plots. Whilst they indicate that there may be changes in the Okun parameter, they only roughly indicate when (given a particular choice of window size) and give no economic intuition for those changes (or shed light on the likely driver(s)).

Regarding such likely drivers (to which we alluded in the previous section), it goes without saying that over post-war period the US economy has experienced several major, complex developments which underpin the evolution of these series: the productivity slowdown of the early 1970s, the major oil shocks in that decade, the fall in the labor income share, swings in the stance of macroeconomic policies, the "Great Moderation" of reduced macroeconomic volatility (starting from the mid-1980s), as well as the "Great Recession" (2007-2009), followed in turn by extraordinary monetary policy accommodation.

There were also important shifts from the mid 1980s onwards (Fernald and Wang, 2016): labor productivity turned from pro to counter cyclical – largely reflecting the weakening pro cyclicality of factor utilization, which itself points to reduced factor hoarding. This may be due

[15] Some studies express OL not as $y = f(u; \beta)$ but instead as $u = f(y; 1/\beta)$. Using the latter normalization the interpretation would equivalently be that, for $\beta_b < \beta_a < 0$, a given fall in cyclical output would produce a smaller proportionate rise in unemployment under 'b' than 'a'.

[16] In the appendix as an additional exercise, Figure D.4 plots the fit of the dynamic OL regression from 2007q1-2018q4 (i.e., from the Great Recession onwards) against the realized outcomes. If the full sample forecast fitted the data exactly, all points would be on the $45°$ angle. However the bulk of the points turn out to be *below* the line indicating that the Okun coefficient is more strongly negative than required to fit the data (meaning that ideally the coefficient should be less negative).

Table 1: Okun Law Coefficients

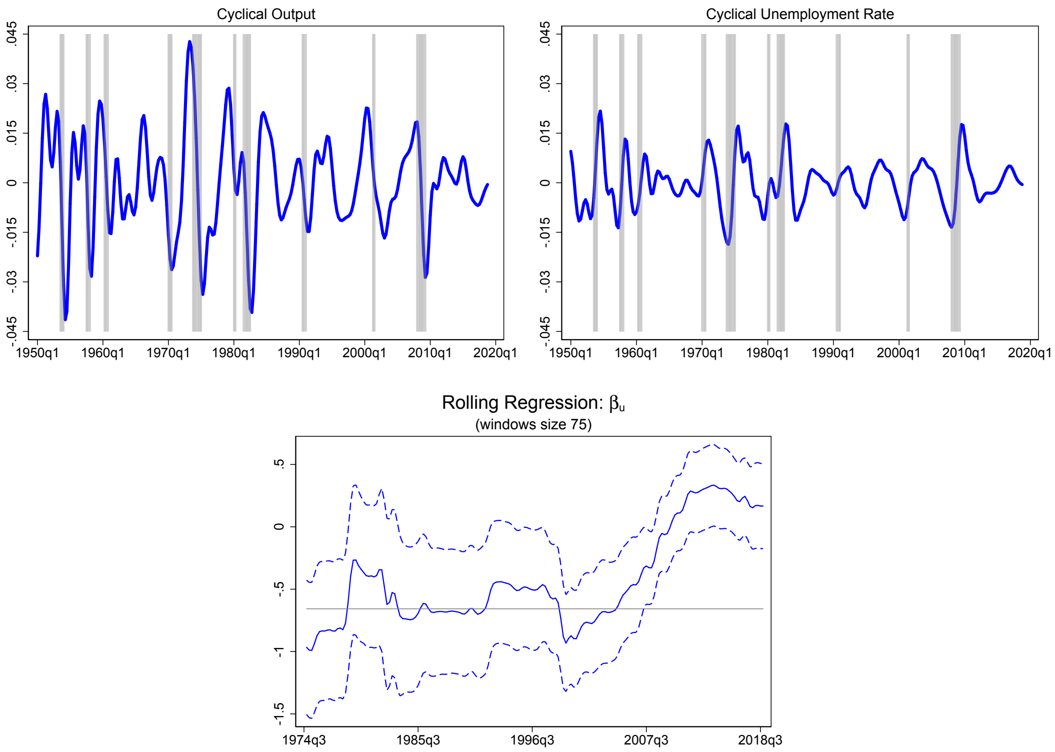
	$AR1$	Static OL	Dynamic OL
ρ	0.907^{***} (0.02)		0.607^{***} (0.07)
β		-1.815^{***} (0.06)	-0.658^{***} (0.13)
β^*		-1.815^{***}	-1.674^{***}
R^2_{adj}	0.830	0.796	0.843
AIC	-1015.521	-993.318	-1026.245

Note. This table shows the estimates of Okun's law over the sample 1950q1 to 2018q4 for the gap form:

$$y_t = \rho y_{t-1} + \beta u_t + \eta_t$$

where $y_t = \tilde{y}_t - y_t^*$ is the deviation of log real GDP from log potential real GDP (i.e., the cyclical component of real GDP: y_t). Likewise, $u_t = \tilde{u}_t - u_t^*$ is the deviation of the unemployment rate from its 'trend', where $\beta^* = \beta/(1-\rho)$. Numbers in parentheses below the coefficients represent bootstrapped standard errors. R^2_{adj} is the adjusted value of R^2. AIC denotes the Akaike Information criterion (for a given model size, the lowest score is preferred).

Figure 1: Okun Law Variables, Correlations and Stability



Note. Shaded gray areas represents NBER recession dates at quarterly frequency. In the rolling window coefficients graphs (for the gap version of the Okun's Law) the 95% standard errors are given by the blue dashed-dot lines (the window size is 75 quarters). The full sample estimate is indicated by the horizontal line.

to increased economic flexibility (i.e., expansion of female labor participation, perhaps declining labor power); it may also reflect the decline of manufacturing (where utilization was traditionally a more important margin of adjustment).[17] Whilst it is unrealistic to be able to map all these events to well-defined changes in Okun's coefficient, they do provide a persuasive case for examining the possible time variation in and non-linear nature of such a key relationship.

4 A LSTR Model with Threshold Endogeneity

Consider OL model (1) without the dynamic term y_{t-1} for analytical convenience. A structural threshold model that can be employed to interpret the structural changes in the slope coefficient β is the following two-regime logistic smooth transition regression (LSTR) model:[18,19]

$$y_t = x_t'\beta_1(1 - \boldsymbol{g}(z_t;\gamma,\delta)) + x_t'\beta_2\boldsymbol{g}(z_t;\gamma,\delta) + \varepsilon_t, \tag{5}$$

where $x_t = (1, u_t)'$ is a (2×1) vector of independent variables, when an intercept is also included, β_1 and β_2 respectively denote the vectors of the slope coefficients of x_t in two distinct regimes, denoted $h = \{1, 2\}$, where "**1**" stands for the first regime and "**2**" for the second, and $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$ is the disturbance term. Function $\boldsymbol{g}(z_t;\gamma,\delta)$ is a continuous logistic function of the observable variable z_t, known as the threshold variable, which governs the transition between the two regimes, at time t:

$$\boldsymbol{g}(z_t;\gamma,\delta) = \frac{1}{1 + \exp(-\gamma(z_t - \delta))} \in [0, 1]. \tag{6}$$

The value of $\delta \in \mathbb{R}$, known as the location or threshold parameter, defines the two regimes. Parameter $\gamma > 0$, the speed-of-transition parameter, determines the smoothness of the transition from one regime to the other. When $\gamma \to \infty$, $\boldsymbol{g}(z_t;\gamma,\delta)$ tends to indicator function $\mathcal{I}(z_t > \delta)$, for all i. In this case, the LSTR model can be approximated by the standard threshold model and thus, the transmission between regimes is abrupt: the shift from regime "**1**" to "**2**" becomes instantaneous at $z_t = \delta$. On the other hand, when $\gamma \to 0$, then $\boldsymbol{g}(z_t;\gamma,\delta) \to \frac{1}{2}$. In this case, the LSTR model reduces to a linear model, with vector of parameters $\frac{1}{2}(\beta_1 + \beta_2)$.

Through transition function $\boldsymbol{g}(z_t;\gamma,\delta)$, the slope coefficients of model (5) are time-varying, depending on the value of threshold variable z_t. One econometric problem often encountered in practice, though, when estimating model, is the bias of the estimates of the parameters of the model due to the contemporaneous correlation of z_t with the error term e_t, raising issues of the endogeneity of z_t. This is likely to happen in economic relationships, like the OL, where z_t and y_t (or e_t) are jointly determined or reflect exogenous common sources of shifts (shocks) due to policy changes.[20] See our discussion in Section 5.

Let us therefore split the sample of z_t across the two regimes as follows: $\mathcal{Z}_{1t} = (-\infty, \delta]$ and $\mathcal{Z}_{2t} = (\delta, \infty)$, for all t, based on a value of location parameter δ, and assume that the disturbance term ε_t is distributed differently across the two regimes, i.e., $\varepsilon_{ht} \sim IID(0, \sigma_h^2)$, $h = \{1, 2\}$. Endogeneity between z_t and ε_{ht} means that $\mathbb{E}(\varepsilon_{1t}|\mathcal{Z}_{1t}) \neq 0$ and $\mathbb{E}(\varepsilon_{2t}|\mathcal{Z}_{2t}) \neq 0$, which

[17] The average rate of capacity utilization has been declining for many decades, see Figure 2 below.

[18] See Teräsvirta (1994). Note that the inclusion of a (potentially regime-specific) intercept in the LSTR model algebra is for generality.

[19] We restricted ourselves to two regime models. Attempts to go beyond two regimes are extremely hard to estimate and take to the data.

[20] Note that some of these shifts may be deterministic in nature, i.e., shifts in the levels or the linear trends of the time series (see Terasvirta (1994).

implies that estimates of the location parameter δ will be inconsistent. This in turn implies that the estimates of all slope coefficients of model (5) will also be biased.

There are different estimation methods to tackle this endogeneity problem in econometrics (Antonakis et al. (2014) provide a survey). The method chosen here draws on that of Christopoulos et al. (2021) which employs copulas to capture the dependence between the disturbance term ε_t and the threshold variable z_t.[21] This method allows for the level of dependence between ε_{ht} and z_t, as well as the variance of ε_{ht} to be different across the two regimes. Copulas are functions which can express joint probability distributions (or densities) of random variables ε_{ht} and z_t in terms of their marginal probability density functions and a copula function capturing the dependence between ε_{ht} and z_t. Based on the copula function, we can derive a single-correlation structure of ε_{ht}, $h = 1, 2$, which can be used to control the endogeneity of z_t in model (5). To this end, consider the following definition:

Definition: Let $p_\delta \in (0, 1) = P(z_t \leq \delta) = F_z(\delta)$, where F_z is the distribution function of z_t. The joint distribution of the pair of random variables $(\varepsilon_{ht}, z_t \in \mathcal{Z}_{ht})$, $h = \{1, 2\}$, can be written using copulas as follows:

$$F_{\varepsilon_h z}(\varepsilon_{ht}, z_t|\mathcal{Z}_{ht}) = C_h\left(F_{\varepsilon_h}(\varepsilon_{ht}), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right), \tag{7}$$

where C_h is a bivariate appropriately scaled copula, with $C_h : [0, 1]^2 \to [0, 1]$, F_{ε_h} is the marginal distribution function of ε_{ht}, and

$$\begin{aligned} F_{z|\mathcal{Z}_1}(z_t|\mathcal{Z}_{1t}) &= \tfrac{F_z(z_t)}{p_\delta} \qquad \text{if} \quad 0 \leq F_z(z) \leq p_\delta \\ F_{z|\mathcal{Z}_2}(z_t|\mathcal{Z}_{2t}) &= \tfrac{F_z(z_t)-p_\delta}{1-p_\delta} \quad \text{if} \quad p_\delta \leq F_z(z) \leq 1, \end{aligned} \tag{8}$$

are the truncated from above and below location parameter value δ distribution functions of z_t, respectively.

From the above definition, it is clear that the truncated joint distribution of the pair of variables (ε_{ht}, z_t), $F_{\varepsilon_h z}(\varepsilon_{ht}, z_t|\mathcal{Z}_{ht})$, $h = \{1, 2\}$, constitutes a copula C_h on $[0, 1]^2$ with uniformly distributed on $[0, 1]$ margins, since $F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})$ is appropriately scaled to integrate to unity.[22] Based on copula theory, the conditional distribution function of ε_{ht} on $z_t \in \mathcal{Z}_{ht}$ can be derived from C_h as follows:

$$F_{\varepsilon_h|\mathcal{Z}_h}(\varepsilon_{ht}|\mathcal{Z}_{ht}) = \frac{\partial}{\partial F_{z|\mathcal{Z}_h}} C_h\left(F_{\varepsilon_h}(\varepsilon_{ht}), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right), \tag{9}$$

while the conditional probability density function related to this distribution is given as,

$$\begin{aligned} f_{\varepsilon_h|\mathcal{Z}_h}(\varepsilon_{ht}|\mathcal{Z}_{ht}) &= \frac{\partial^2}{\partial \varepsilon_h \partial F_{z|\mathcal{Z}_h}} C_h\left(F_{\varepsilon_h}(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right) \\ &= c_h\left(F_{\varepsilon_h}(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right) f_{\varepsilon_h}(\varepsilon_{ht}), \end{aligned} \tag{10}$$

[21] Although in their study they considered the simpler TAR model. The extension of the approach to the LSTR model makes increases the generality and usefulness of the framework since as already discussed it nests the linear and threshold model and allows for the smooth transitions between the regimes.

[22] The two copulas $C_h, h = 1, 2$, can be glued along z_t at point p_δ, and produce a single copula (see Siburg and Stoimenov, 2008; and Erdely, 2017).

where $c_h\left(F_{\varepsilon_h}(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right) = \frac{\partial^2}{\partial F_{\varepsilon_h}\partial F_{z|\mathcal{Z}_h}} C_h\left(F_{\varepsilon_h}(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right)$ is the copula density function corresponding to C_h and $f_{\varepsilon_h}(\varepsilon_t) = \frac{\partial}{\partial \varepsilon_h} F_{\varepsilon_h}(\varepsilon_t)$ is the probability density of ε_{ht}.[23] The last relationship of the conditional density $f_{\varepsilon_h|\mathcal{Z}_h}(\varepsilon_{ht}|Z_{ht})$implies that we can capture the contemporaneous dependence between ε_{th} and $z_t \in \mathcal{Z}_{ht}$, through the copula density function $c_h\left(F_{\varepsilon_h}(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right)$. For a Gaussian copula (which is often used in practice, e.g., Park and Gupta, 2012; Joe, 2014) this function becomes

$$c_h\left(F_{\varepsilon_h}(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right) = \varphi\left(\Phi^{-1}\left(F_{\varepsilon_h}(\varepsilon_t)\right), \Phi^{-1}\left(F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right)\right), \tag{11}$$

where $\Phi^{-1}$is the quantile function of the standard normal distribution and φ the probability density function of the normal distribution.[24]

If the error term ε_{ht} is normally distributed $\sim \mathcal{N}(0, \sigma_h^2)$, the Gaussian copula given by (11) implies a linear relationship between ε_{ht} and $\Phi^{-1}\left(F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right)$ which can deal with the endogeneity problem of the threshold variable z_t . Substitution of (11) into (10) implies that ε_{ht} has a single factor correlation structure[25]

$$\varepsilon_{ht} = \omega_h z_{ht}^* + Var(\varepsilon_{ht}|z_{ht}^*)^{1/2}\xi_{ht}, \quad h = \{1, 2\}, \tag{12}$$

and

$$\mathbb{E}(\varepsilon_{ht}|\mathcal{Z}_{ht}) = \omega_h z_{ht}^*, \text{ with } \omega_h = \sigma_h \rho_{\varepsilon_h z_h^*}, \tag{13}$$

where $z_{ht}^* = \Phi^{-1}\left(F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right)$ is distributed as $z_{ht}^* \sim \mathcal{N}(0,1)$, $\rho_{\varepsilon_h z_h^*}$ is the Pearson correlation coefficient between ε_{ht} and z_{ht}^*, $Var(\varepsilon_{ht}|z_{ht}^*) = \sigma_h^2(1-\rho_{\varepsilon z_h^*}^2)$ and ξ_{ht} is an $IID(0,1)$ a disturbance term which is independent of z_{ht}^*, implying $\mathbb{E}(\varepsilon_{ht}|z_{ht}^*) = 0$.

Using the above representation of relationship of ε_{ht}, model (5) can be written as

$$y_t = \begin{cases} x_t'\beta_1(1 - \boldsymbol{g}(z_t; \gamma, \delta)) & + \quad \omega_1 z_{1t}^* + e_{1t} \quad \text{if} \quad z_t \in \mathcal{Z}_{1t} : \text{Regime ``}\mathbf{1}\text{''} \\ \\ x_t'\beta_2 \boldsymbol{g}(z_t; \gamma, \delta) & + \quad \omega_2 z_{2t}^* + e_{2t} \quad \text{if} \quad z_t \in \mathcal{Z}_{2t} : \text{Regime ``}\mathbf{2}\text{''} \end{cases} \tag{14}$$

where $e_{ht} = -\omega_h z_{ht}^* + \varepsilon_{ht}$, $h = \{1, 2\}$, is a disturbance term with conditional mean $\mathbb{E}(e_t|z_{ht}^*) = 0$, since $\mathbb{E}(\varepsilon_{ht}|z_{ht}^*) = \omega_h z_{ht}^*$ and $\mathbb{E}(\xi_{ht}|z_{ht}^*) = 0$. This augmented with random variables z_{ht}^* version of model (5) can be employed to control for the endogeneity problem of z_t. The random variables z_{ht}^* correct the conditional mean of y_t on Z_h, $\mathbb{E}(y_t|Z_{ht})$, for the contemporaneous correlation between ε_{ht} and z_t for both regimes of the model. This is done, note, without making any assumption about the distribution of z_t. Furthermore, as already highlighted, it allows the distribution of ε_{ht} and its correlation structure with z_t to change across the two regimes.

---

[23] Note that the joint probability density corresponding to $F_{\varepsilon z}(\varepsilon_t, z_t|\mathcal{Z}_{ht})$ is given as follows:

$$f_{\varepsilon\mathcal{Z}}(\varepsilon_t, z_t|\mathcal{Z}_{ht}) = c_h\left(F_\varepsilon(\varepsilon_t), F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right) f_\varepsilon(\varepsilon_t) f_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})$$

where $f_{z|\mathcal{Z}_1}(z_t|\mathcal{Z}_{1t}) = \frac{f_z(z_t)}{p_\delta}$, $-\infty < z_t \leq \delta$, and $f_{z|\mathcal{Z}_2}(z_t|\mathcal{Z}_{2t}) = \frac{f_z(z_t)}{1-p_\delta}$, with $\delta < z_t \leq \infty$, are the truncated from above and below p_δ probability densities $f_z(z_t)$.

[24] As shown by Park and Gupta (2012), or Christopoulos et al (2021) the Gaussian copula can satisfactorily approximate different types (linear or nonlinear) dependencies between two random variables, under different distributional assumptions. In our Monte Carlo study we show that the Gaussian copula can satisfactorily approximate a dependent structure between the error term variable ε_{ht} and the threshold variable z_t when these two variables follow the Student-t distribution.

[25] See, for example, Joe (2014).

The augmented model (5) can be straightforwardly extended to include the lagged dependent variable y_{t-1}. This variable is not contemporaneously correlated with the error term e_t and, thus, does not raise any endogeneity regressor issues. In addition, it can be easily extended to include copula based transformations of the regressors including in vector x_t, if they are also correlated with e_t (see Christopoulos et al (2021)). This extension is an empirical matter. In our empirical OL application, we have found that it does not change the estimates of the model. Our interpretation for this is that the threshold variable z_t, being correlated with u_t, absorbs, substantially, any endogeneity effects of u_t on the estimates of the augmented OL relationship, by z_t.

4.1 Estimation Aspects and Monte Carlo Results

Model (14) can be employed to estimate the location parameter δ and its remaining parameters collected in vector $\theta(\delta) = (\gamma, \beta_1, \beta_2, \omega_1, \omega_2)'$ based on a two-step nonlinear least squares (NLLS) method, since its disturbance term e_{ht} is independent of the transformed variable z^*_{ht}, for $h = \{1, 2\}$.

In particular, δ can be estimated, in a first step, by solving the following NLLS optimization problem:

$$\hat{\delta} = \underset{\delta \in Q_z}{\arg\min} RSS(\delta), \tag{15}$$

where $RSS(\delta) = \sum_{t=1}^{T} \hat{e}^2_{h,t}$, is the residual sum of squares of (14), δ is an interior point of Q_z, since we assume $p_\delta = P(z_t \leq \delta) \in (0, 1)$. To estimate δ, note that we require values of the transformed variables z^*_{ht}, given by $\Phi^{-1}\left(F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})\right)$. This can be done based on non-parametric estimates of the marginal distribution $F_{z|\mathcal{Z}_h}(z_t|\mathcal{Z}_{ht})$ (see Silverman, 1986), or based on the empirical cumulative distribution function. Given the optimal estimate of δ, the slope parameters of model (5) and the speed-of-transition coefficient γ collected in vector $\theta(\hat{\delta})$ can be estimated, in a second step. Following the literature on threshold models (see, e.g., Chan, 1993; Samia and Chan, 2011), the estimator $\widehat{z}_\delta$ is T-consistent and the estimates of vector $\theta(\hat{\delta})$, which correspond to $\hat{\delta}$, are $\sqrt{T}$ asymptotically normal.

As a remark on the above estimation procedure note that, instead of the one dimensional grid search over δ, we can carry out a two dimensional grid search over δ and γ (see, e.g., Leybourn et al., 1986; and Franses and van Dijk, 2000). This procedure may mitigate optimization problems in estimating γ, due to the nonlinear nature of function $\boldsymbol{g}(z_t; \gamma, \delta)$. Furthermore, it can have better small sample properties, due to grid-search process in estimating δ and γ. Given the estimates of δ and γ, then we can estimate the remaining slope parameters, collected in vectors β_h, in a second step.

To evaluate the performance of our estimation approach to successfully control for the endogeneity of threshold variable, we carry out a small MC study, see Appendix A. We consider cases in which the disturbance term ε_{ht} and threshold variable z_t are jointly normally and then Student-t distributed, with different values of the speed-of-transition parameter γ and sample sizes T. Generating data from the Student's t distribution, which allows for tail dependence between ε_{ht} and z_t, can show the robustness of our method to such features in the data.

The results of our MC clearly supports the view that our method can successfully control for the endogeneity of z_t on the estimates of location parameter δ. They also show that ignoring this endogeneity leads to series biases in the estimates of δ.

These results hold for both the distributions of ε_{ht} and z_t considered. This means that our method is robust to misspecification of these distributions. For the comparator case that ε_{ht} and z_t are normally distributed, our method compares favorably to that of Kourtellos et al. (2016) based on the inverse Mills ratios, modified appropriately for the LSTR model (5). Also, our method is robust to the case that both ε_{ht} and z_t follow the Student-t distributed. Finally, another interesting result of our MC exercise is that our method can be implemented without any concern for bias or inefficiency of the estimates of δ in the case where the threshold variable z_t is exogenous.

4.2 New Test for Linearity With Threshold Endogeneity

Before estimating model (5), or its extended version (14), a critical prior testing procedure is to diagnose if the data supports our threshold model compared to its linear specification. To this end, we suggest a suitable testing procedure.

We follow recent work in the literature on threshold or LSTR models (see, e.g., Hansen (1996) and KılıÇ, 2016) which is focused on testing H_0: $\beta_1 = \beta_2$ (implying $\gamma \to 0$) against H_a: $\beta_1 \neq \beta_2$ (implying $\gamma > 0$). As noted by KılıÇ (ibid), compared to inference procedures testing for the exclusion restrictions on the threshold variable z_t or its product terms with regressors collected in vector x_t based on an approximation of model (5) under H_0, our suggested procedure may be proved more powerful for γ values far away from the $\gamma = 0$ neighborhood.[26] That is to say where the approximation of (5) is not accurate and depends on the value of the location parameter δ. Furthermore, in our simulations we found that endogeneity of threshold variable z_t makes the power performance of the inference procedures based on the above approximation of (5) even worse, due the estimation bias of the slope coefficients of the auxiliary regression.

More specifically, to test H_0: $\beta_1 = \beta_2$ against H_a: $\beta_1 \neq \beta_2$, we rely on the LR test statistic:

$$LR(\gamma, \delta) = 2 \times (\log L(\theta(\delta)) - \log L(\beta)) \tag{16}$$

where $\log L(\theta(\delta))$ and $\log L(\beta)$ constitute the maximum log-likelihood function of model (5) under the alternative and null hypotheses, respectively. Note that, in order to estimate the model under the alternative hypothesis, we will use the auxiliary regression (14), controlling for the endogeneity of the threshold variable z_t. Since the nuisance parameters δ and γ are not identified under the null, we next suggest a sup-version of statistic $LR(\gamma, \delta)$ (see Andrews and Ploberger, 1994), defined as follows:

$$LR^{\text{sup}} \equiv \sup_{(\gamma,\delta)\in Q_\gamma \times Q_z} LR(\gamma, \delta), \tag{17}$$

where Q_γ denotes a compact subspace on the real line, searching for an optimal value of γ and Q_z is defined as before.[27] Since the distribution of LR^{sup} is non standard under the null, its critical values can be obtained based on a parametric bootstrap procedure, generating data under the null of linearity (i.e., H_0: $\beta_1 = \beta_2$).

To evaluate the power performance of statistic LR^{sup}, we carry out a MC exercise in Appendix B. In this exercise, we also consider the case that we ignore the endogeneity of threshold variable z_t. The results of this exercise indicate that LR^{sup} has satisfactory power. This is true under alternative copula functions and marginal distributions of ε_t and z_t considered in the MC exercise. Another

[26] The initial literature of tests for linearity against threshold specifications (such as Luukonen et al., 1998) replaced the transition function by a Taylor series expansion which is estimable under null with the test amounting to testing for the significance of the interaction of the linear regressors with the polynomial terms.

[27] Note that, in practice, Q_γ can be set to $Q_\gamma = \left[\frac{1}{10\sigma_z}, \frac{1000}{\sigma_z}\right]$, where σ_z is the standard deviation of the threshold variable z_t, see KılıÇ (2011).

interesting finding is that our test does not lose significant power in the case where z_t is exogenous.

5 Data and Threshold Candidates

So far we have discussed the possibility of (endogenous) threshold variables leading to threshold effects in OL, without discussing which those variable(s) would be. As already indicated, the literature on asymmetries has concentrated on cyclical features. But, to recall, Okun himself identified many channels of influence on the Okun trade off. And it is with this broader more encompassing approach in mind that we proceed.

In that vein, Fernald et al. (2017) decompose Okun's coefficient emphasizing the different margins and adjustment channels that firms and households use to respond to different shocks.[28] Although these authors do not explore threshold nonlinearity, their framework is nonetheless suggestive. For instance, labor hoarding (an empirically well-established phenomenon) cushions the unemployment consequences of downturns perceived to be temporary or "small". However, a sufficiently large downturn (i.e., in our context, beyond some estimated threshold) may counter labor hoarding incentives, thus changing the output-unemployment nexus.

We classify these potential threshold candidates in three broad (and not necessarily mutually exclusive) categories (see Table 2)[29]:

1. *Demand & Cyclical* pressures;

2. *Structural* features of the economy, and;

3. *Policy & Financial* variables.

Our main data source is the Federal Reserve Bank of St. Louis (FRED) database, the Bureau of Labor Statistics (BLS) and the Congressional Budget Office (CBO). The shadow federal funds rate is taken from Wu and Xia (2016). As before, the sample mostly spans 1950q1 to 2018q4. In the final column we show the corresponding LR^{sup} test statistic (recalling Section 4.2) and probability values of the null hypothesis that the linear model (1) constitutes the correct specification of the data against the LSTR specification (5), controlling for the endogeneity of the threshold variable, which includes in its rhs the lagged dependent variable y_{t-1}.

Consider the first category of threshold candidates, *Demand & Cyclical*. Many studies suggest Okun's coefficient moves over the phases of the business cycle, and in a nonlinear, asymmetric fashion (e.g., good examples of this include Lee, 2000; Harris and Silverstone, 2001; Cuaresma, 2003; Silvapulle et al. 2004). Common rationales for this cyclical asymmetry include: labor hoarding; downward wage rigidity; employment regulations; patterns of cyclical firm expansion and contraction; non-constant factor substitution (Courtney, 1991) etc. In line with this business-cycle asymmetry literature, we consider as possible threshold variables: the unemployment and 'natural' unemployment rate, the output gap, output, and inflation rates (which similarly reflect cyclical demand pressures).[30]

[28] To illustrate, respectively, these are the extensive (e.g., labor force participation and migration etc.) and intensive margins (e.g., hours per worker, factor utilization etc.)

[29] Many of these are quite representative of macro US series commonly used for econometric studies, and most are included, for instance, in the large dat set of Marcellino et al. (2006) (their economy-wide variables anyway).

[30] Note that dichotomous (Heaviside) threshold variables (of the expansion/contraction variety) which are often used in

Our second category is dubbed *Structural*. Again this reflects various strands of the Okun literature as well as more general themes of structural change in the post-war US economy – the idea being that various secular (though not necessarily irreversible) trends may have imparted an effect on the slope coefficient of the OL relationship. For example employment and participation rates, the labor income share, macroeconomic volatility etc. The final category assesses *Policy & Financial* variables as threshold candidates, including a variety of public and market interest rates.[31]

Table 2 shows around half the threshold candidates reject linearity. Among the first category, "Demand" (or cyclical) factors, the evidence is mixed. Only inflation and capacity utilization appear important. The latter, though strongly significant, has a smaller sample (from 1967 onwards), and is not economy wide.[32] This finding is interesting since cyclical factors have been the favored candidates of past Okun threshold studies (e.g., Virén, 2001). However even a finding of weak demand threshold candidates need not imply that cyclical pressures do not impact OL. Downturns may have long and persistent effects, inducing protracted adjustments in Okun margins that extend well beyond the defined periods of recession, with the short run effects but one aspect of those influences.

In the "Structural" category, the labor share,[33] and long term unemployed (percentage unemployed for $\geq$ 27 weeks), illustrating important shifts in labor-market features, are also detected. Finally, "Policy & Financial" measures, yield a number of valid threshold candidates. Both long and short policy rates and the shadow rate reject the linear specification and thus are viable threshold candidates.

6 Estimation of the LSTR Model for Okun's Law

Having presented our empirical methodology to estimate a LSTR model under threshold endogeneity, and pretested for linearity, we now estimate the model. To measure directly the magnitude of the shift in the slope coefficient of unemployment, we estimate an augmented version of (5) with the regressors z_{1t}^* and z_{2t}^*, defined in Section 4, and where, for congruence with the literature, we assume that threshold behavior affects the Okun coefficient alone:[34]

$$y_t = \rho y_{t-1} + \beta u_t + b u_t \left[1 + e^{-\gamma(z_{1t}^* - \delta)}\right]^{-1} + \omega_1 z_{1t}^* + \omega_2 z_{2t}^* + e_t. \tag{18}$$

such analyses, such as

$$z_t = \begin{cases} 1: \text{Expansion} & \text{if } u_t < 0, \text{or } y_t > 0 \\ 0: \text{Contraction} & \text{otherwise} \end{cases}$$

are not feasible threshold variable types in our framework since they are not continuous.

31 We also considered the Gilchrist-Zakrajšek Corporate Bond Credit Spread, and the National Financial Conditions Index (from the FRB Chicago). But their short sample (available from 1973 and 1971, respectively) precluded their use. Although the credit spreads share similar information with the interest rates used since, for instance, they reflect similar monetary or fiscal shocks. Likewise, many variables which might have been of interest, for instance inequality measures, tax burden, 'economic and policy uncertainty' metrics etc were excluded since they tend to be of short sample and/or not available at a quarterly frequency. Further details of these data choices and restrictions are available on request.

32 Capacity utilization indexes reported by FRED, moreover, are not economy wide: they are constructed for 71 industries in manufacturing, 16 in mining, and 2 in utilities.

33 The variant of the labor share that we use is the quarterly indexed one provided by FRED where 2012 = 100. For comparison purposes we annualized this series and compared it to the Share of Labor Compensation series provided by University of Groningen (also available from FRED), see supplementary material Figure D.5.

34 We also tried a variety of structural break and recursive tests but we failed to detect any systematic pattern of time dependency in the autoregressive coefficient. This bolsters our case for concentrating on the Okun coefficient.

Table 2: Data: Definitions, Sources and Linearity Testing

Data Series	Symbol	Source (mnemonic)	LR^{sup}
Demand			
Cyclical Unemployment Rate	u	FRED (derived from *UNRATE*)	2.968 [0.315]
Cyclical Real GDP	y	FRED (derived from *GDPC1*)	4.250 [0.352]
Output Gap (% of potential GDP)	og	CBO	4.749 [0.310]
Capacity Utilization: Total Industry	cu	FRED (*TCU*)	10.830 [0.002]
Consumer Price Inflation (%Δ yoy)	π	FRED (*CPIAUCSL*)	6.259 [0.050]
Commercial Real Estate Price Inflation	π_H	FRED/Haver Analytics	3.017 [0.176]
Structural			
Rolling 10-year St. Dev. of Cyclical Output (Risk)	$\sigma_{y_{t(10)}}$	(derived from *GDPC1*)	5.307 [0.085]
Utilization-Adjusted TFP Growth	g_{TFP}	Fernald (2018)	3.542 [0.325]
Non-farm Business Sector: Labor Share	$labsh$	FRED (*PRS85006173*)	16.871 [0.013]
Employment Share: Manufacturing	$lmansh$	FRED (derived from *MANEMP/PAYEMS*)	5.224 [0.205]
Labor Force Participation Rate (Total)	pr	FRED (*CIVPART*)	4.239 [0.183]
Labor Force Participation Rate (Female)	pr_f	FRED (*LNS11300002*)	4.613 [0.123]
Natural Rate of Unemployment	u_n	CBO	6.223 [0.067]
Long-Term Unemployed[1]	u_l	BLS (*LNS13025703*)	22.951 [0.001]
Policy & Financial			
Federal Funds Rate	i_s	FRED (*FEDFUNDS*)	12.728 [0.040]
Shadow Policy Rate[2]	$i_{s,wx}$	Wu and Xia (2016)	12.481 [0.041]
Long Term Interest Rates (10y Gov. Bonds)[2]	i_l	FRED (*IRLTLT01USM156N*)	11.068 [0.003]
Corporate Bond Spread[3]	i_{Baa}	FRED (*BAAFFM*)	5.209 [0.255]

Note. This table tests threshold candidates for the LR^{sup} linearity test. All series are quarterly and, where relevant, seasonally adjusted and span 1950q1-2018q4 except i_s, $i_{s,wx}$ which start in 1954q3, i_l (1960q1), cu (1967q1). [1] Unemployed for $\geq$ 27 weeks (as a % of civilians unemployed). [2] Calculated as end of period. [3] Moody's Baa Corporate Bond minus Federal Funds Rate. Numbers in brackets in the final column are bootstrapped probability-values of the threshold linearity tests.

Parameter b captures the magnitude of the shift of the Okun coefficient of u_t from the first regime to the second. The no-threshold (linear) case of the dynamic OL is retrieved if $b = 0$. The long-run OL coefficients in the first and second regimes are, respectively, given by,

$$\beta_1^* = \frac{\beta}{1-\rho} \quad ; \quad \beta_2^* = \frac{\beta + b}{1-\rho}. \tag{19}$$

Table 3 presents estimates of dynamic (AR1) OL with the threshold variable treated as endogenous.[35] This is based on our preferred estimation approach just presented, including the copula-transformed variables z_{1t}^* and z_{2t}^* as regressors with the attendant parameters ω_1, ω_2. For better small sample estimation properties, we present bootstrapped standard errors (in brackets) and confidence intervals (in braces). These are calculated based on a wild parametric bootstrap method (see Davidson and MacKinnon, 2007). Note, the bootstrapped values are subject to the distribution of the threshold variable.

6.1 Estimation Results

table 3 shows the estimates of coefficients ρ, β and b of model (18), the implied long-run coefficients β_1^* and β_2^*, the threshold and speed of adjustment parameters (δ and γ, respectively) the AIC and the sample percentage residing in the second regime, denoted $\mathcal{D}_{\mathbf{2}} \in (0, 100)$. We also report the ω_h, $h = \{1, 2\}$, parameters (multiplied by 100 for legibility).

To aid interpretation, consider a case where

$$\beta_1^* \;\; < \;\; \beta_2^* < 0 \tag{20}$$

If condition (20) holds, it implies that a given unemployment gap is associated with a larger output gap in regime "**1**" relative to "**2**". Accordingly, we can think of regime "**1**" as the *steeper* regime, and regime "**2**" as the *flatter*. For instance, when the labor share of income exceeds the threshold value (which is around 70% of the time), the output gap is almost twice as sensitive to changes in the unemployment gap: $|\beta_2^*| = 1.863$. Otherwise, they move around one-to-one: $|\beta_1^*| = 1.156$. Note that a shift from the regime "**1**" to "**2**" does not always mean a change from the steeper to the flatter OL relationship. For cu and i_l, moreover, there is no clear-cut significant OL relationship in regime "**1**", at conventional significance levels. Although, to recall, capacity utilization has a far shorter sample and is not economy-wide in coverage.

Figures 2 – 4 plot the threshold time series, z_t (in solid blue lines using the lhs axis) overlaid with the transition probabilities $\boldsymbol{g}(z_t; \hat{\gamma}, \hat{\delta})$ (in red dashed lines using the rhs axis) and the scalar estimated threshold $\hat{\delta}$.[36] For a given series, the higher is γ the more rapid is the transition between regimes.[37] High values, i.e., $\gamma > 10$, as for threshold variables π and cu often have volatile and rapid adjustment transition probabilities. This conforms with their status as demand/cyclical variables. The labor share, and the short term policy rates by contrast exhibit intermediate transitions, i.e., $\gamma \in (1, 6)$. A lower adjustment speed, i.e., γ between 0 and 4, is associated to the long-term unemployment rate. One might expect this variable to reflect long run (slow moving) supply variables such as demographics, labor supply, work incentives, economic dynamism etc.

Accordingly, the transition probabilities demonstrate a mix of persistent (long lived) *and* tempo-

[35] The results for the exogenous case are shown in Appendix D.

[36] The same graphs with NBER dates overlaid are shown in Appendix D.

[37] This recalls the distinction made by Bernanke et al. (2005) who discuss 'slow' and 'fast-moving' variables.

TABLE 3: Threshold Okun Law Results: Endogenous Threshold

para./z_t	$labsh^{\dagger}$	u_l	π	$i_{s,xw}$	i_s	cu	i_l
ρ	0.546^{***}	0.571^{***}	0.602^{***}	0.692^{***}	0.693^{***}	0.751^{***}	0.748^{***}
	(0.059)	(0.063)	(0.071)	(0.071)	(0.071)	(0.079)	(0.074)
β	-0.525^{***}	-1.767^{***}	-0.492^{***}	-0.208	-0.210	0.278	0.564
	(0.122)	(0.311)	(0.222)	(0.143)	(0.138)	(0.214)	(0.382)
b	-0.321^{***}	1.346^{***}	-0.290	-0.396^{***}	-0.389^{***}	-0.750^{***}	-1.869^{**}
	(0.106)	(0.335)	(0.258)	(0.135)	(0.130)	(0.192)	(0.848)
δ	108.241	0.075	0.017	3.220	3.220	0.759	6.900
	{108.140-108.450}	{0.068-0.097}	{0.0106-0.066}	{2.83-3.460}	{2.98-3.480}	{0.754-0.780}	{2.530-9.02}
γ	5.500	0.100	47.400	1.400	1.600	44.20	0.100
	{4.10-6.70}	{0.010-0.421}	{45.30-48.500}	{0.80-3.10}	{1.01-3.20}	{42.00-46.10}	{0.010-3.900}
$100 \times \omega_1$	-0.001	0.290^{***}	-0.020	0.030^{**}	0.030^{**}	0.420^{***}	0.060^{*}
	(0.028)	(0.055)	(0.082)	(0.039)	(0.037)	(0.080)	(0.033)
$100 \times \omega_2$	-1.460^{***}	0.520	-0.780^{***}	-1.090^{***}	-10.900^{***}	0.140	0.700^{*}
	(0.243)	(0.733)	(0.250)	(0.362)	(3.647)	(0.144)	(0.392)
β_1^*	-1.156	-4.119	-1.236	-0.675	-0.684	3.359	2.238
β_2^*	-1.863	-0.981	-1.965	-1.961	-1.951	-1.080	-5.179
AIC	-1045.497	-1039.460	-1029.924	-977.443	-977.566	-712.802	-913.110
$\mathcal{D}_2$	69.12	86.64	83.48	60.93	60.43	81.31	36.62

Note. This table shows the results from estimation of (18) for the threshold candidates that rejected a linear specification. The ω_h terms are scaled up for readability. Numbers in parentheses below the coefficients represent bootstrapped standard errors, and those in braces represent bootstrapped 95% confidence intervals. The terms β_1^* and β_2^* are as defined in (19). $\mathcal{D}_2$ is the sample percentage residing in the second regime. † The labor share measure used is the quarterly indexed one provided by FRED where $2012 = 100$.

rary regime shifts of the Okun coefficient over the sample. In particular, the labor share ($labsh$) has triggered an apparently permanent shift to the flatter regime after year 2001. For $labsh$, this shift in its own profile has often been attributed to the effects of globalization and off-shoring (widening the global labor pool, weakening labor bargaining power), patterns of technology adoption and skill complementarities over recent decades etc. (inter alia, Manyika et al., 2019). These apparently unfavorable outcomes for labor can also arguably be seen in the increasing share of unemployed who remain unemployed after 27 weeks (Figure 3), which has been trending upward since the mid 1970s, and especially so during the Great Recession.[38]

The Federal Funds rate and the shadow rate ($i_s, i_{s,wx}$), overlaid in Figure 4, as well as inflation rate π (Figure 2) are associated with less permanent regime shifts of the OL relationship. These seem to cause shifts from the steeper to the flatter regime when they reach historically low levels. Large values of i_s or π (such as those in 1970s and 80s) triggered enduring shifts in the Okun parameter to the steeper regime, highlighting the effectiveness of monetary policy. By contrast, the long term interest rate i_l does not seem to trigger frequent regime shifts in the OL. Although, like the short rate, its decline over time reduces the probability of a steep OL regime, there is no statistically significant Okun parameter in the 2nd flatter regime. Moreover, the adjustment parameter appears weakly identified, with wide confidence intervals. Thus whilst it would appear that long term interest rates matter, in so far as they are a product of short rates, fundamentally the main channel is from actual and shadow short-run policy rates. In the aftermath of the recent Great Recession, thus, these threshold variables are associated with the flatter regime of the OL relationship or they triggered a shift to it. In other words, the unemployment and output gap

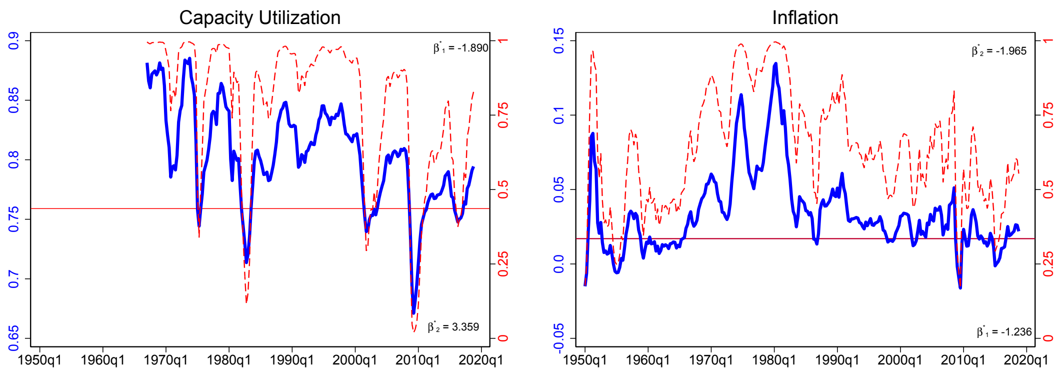
[38] This trajectory though has had limited impact on the Okun coefficient since, apart from some particular spells near the beginning of the sample, the coefficient is mostly around -1.

during that period co-moved more strongly.[39]

How can we interpret this flattening? Consider the $labsh$ threshold: the Okun coefficient was -1.863 roughly until the early 2000s, and -1.156 thereafter. Accordingly the unemployment gap is increasingly associated with a smaller output gap (they now co-move nearly 1-to-1). The corollary (or, literally, the inverse) of this is that the unemployment costs of output contractions are now substantially greater. This accords with outcomes: the period since 2000s were a period of historically low interest rates and, from 2007q4, one of extraordinary cyclical weakness where unemployment and long-term unemployment rose markedly. Of course if there had be no regime change and the value of -1.863 (or -0.537, its inverted variant) held throughout the sample, unemployment would still have risen markedly after 2007 reflecting the severity of the downturn, but not as much as it did.

Firms may have understood early on the severity of the Great Recession and hoarded labor less than otherwise (implying rising labor productivity), and perhaps substituted towards more capital-intensive production. Moreover various underlying structural changes in the US economy (such as declining labor share, globalization, changing sectoral composition patterns) may have lent weight to that change. Thus, if we had modelled the economy as having a fixed employment-output trade off, these developments would have been missed. Our examination of several potential threshold variables on the non-structural Okun relationship has provided robust evidence of that important change.

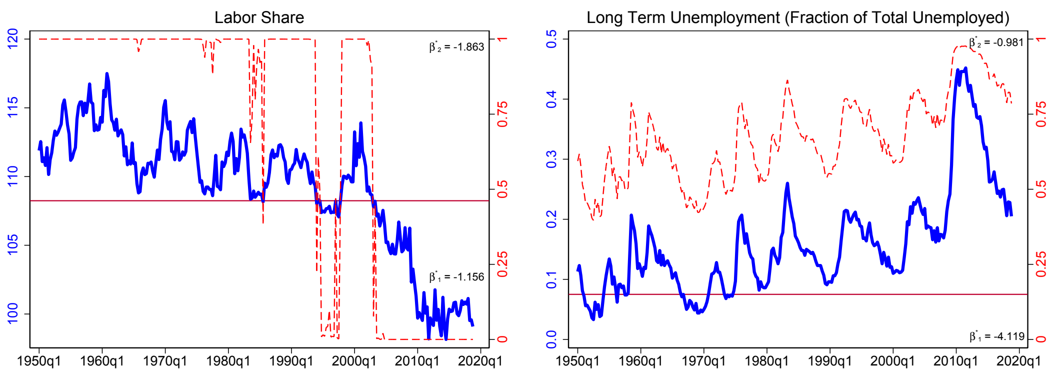
FIGURE 2: Endogenous Threshold Variables & Transition Probabilities (Demand)



Note. Blue solid line represent the threshold variable, z, red dashed line represent the $\{0, 1\}$ transition probabilities. The horizontal solid line represents the estimated threshold parameter, $\hat{\delta}$. The β^* values on the rhs edge are the Okun parameters in the respective regimes.

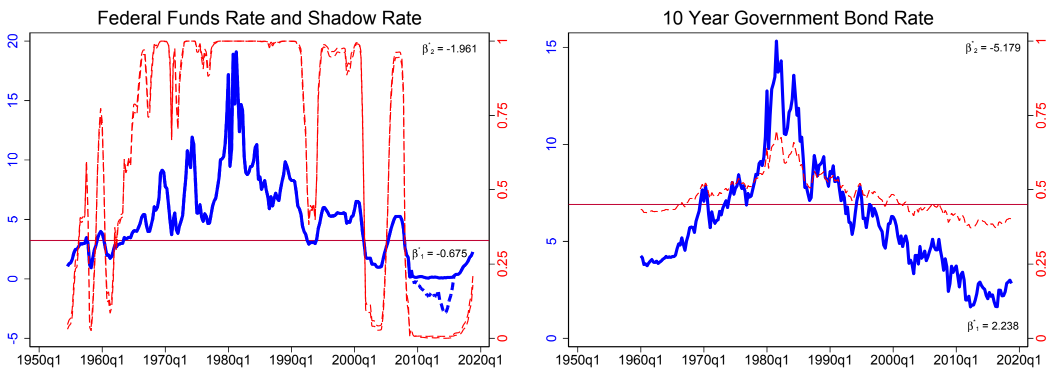
[39] More generally, Auerbach and Gorodnichenko (2012) found that for the US the size of the fiscal multiplier appeared to be state dependent; thus the state of the economy and policy effectiveness are innately liked. Moreover there is a growing literature on state dependency in monetary policy, particularly around the effective lower bound (see Woodford, 2012).

FIGURE 3: Endogenous Threshold Variables & Transition Probabilities (Structural)



Note. See notes to Figure 2.

FIGURE 4: Endogenous Threshold Variables & Transition Probabilities (Policy/Financial)



Note. In the first figure the federal funds rate (shadow rate) is in solid blue (dashed blue). Their transition probabilities are both in red dashed lines, for the β parameters on the rhs of the graphs, we use the i_s case (which is very close to the $i_{s,wx}$ case).

7 A Composite Threshold Variable

Our analysis suggests that many of the regime shifts triggered by the threshold variables examined co-move. This is true for both long-lived and more temporary shifts. Table 4 corroborates this showing the extent of the correlations among the $(0, 1)$ transition probabilities across regimes for all z candidates (all cross-correlations turn out to be significant at 1%). This co-movement can be attributed to common, or related, sources of information and economic events underlying the threshold variables triggering the threshold effects in OL.

Accordingly, we now combine these different information sources into a single common factor. The threshold factor effects considered can summarize and smooth out all the alternative sources of the OL regime shifts, and can potentially better describe and account for the total effects of such shifts over time. We can then use this to estimate (18) and exploit its usefulness as an index variable to assess predictive performance (see our exercises in the appendices).

Consider the K-dimension column vector of the relevant threshold variables considered:

TABLE 4: Correlation of the Transition Probabilities

	$labsh$	u_l	π	i_s	cu	i_l	$i_{s,wx}$
$labsh$	1						
u_l	−0.73	1					
π	0.58	−0.62	1				
i_s	0.69	−0.75	0.64	1			
cu	0.31	−0.57	0.40	0.58	1		
i_l	0.67	−0.48	0.69	0.76	0.27	1	
$i_{s,wx}$	0.69	−0.76	0.65	0.99	0.58	0.77	1

Note. The table presents estimates of the correlation coefficients of the transition probabilities from regime "**1**" to "**2**" for variables z_j which trigger significant threshold effects. All cross-correlations turn out to be significant at 1%.

$Z = [z_j]$, $j = 1, 2, ..., K$, (we drop time subscripts from z_j for notational convenience). The principal component (PC) analysis implies that there are common factors spanning Z, collected in column vector p, which can be obtained as

$$PC_i = p_i' Z, \quad i = 1, 2, ..., K \tag{21}$$

where p_i' is the i^{th}-row of an orthogonal matrix $P = [p_{ij}]$ (normalized, i.e., $P'P = I$) such that $P'\Sigma_z P = \Lambda = diag[\lambda_i]$, with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_K \geq 0$, where Σ_z is the covariance matrix of vector Z and Λ is the matrix of the eigenvalues of Σ_z, denoted λ_i, by the spectral decomposition theorem. Note that the first factor PC_1 (which accounts for the largest single share of the data variation) corresponds to the largest eigenvalue λ_1. The proportion of the variation of the data accounted for by the first v-PCs is given as $\rho_v^2 = \frac{\sum_{i=1}^{v} \lambda_i}{\sum_{i=1}^{K} \lambda_i}$. The covariance and correlation loading coefficients between PC_i and threshold variables z_j, $j = 1, 2, .., K$, are respectively given by

$$cov(PC_i, z_j) = p_{ij}\sqrt{\lambda_i} \text{ and } \rho_{ij} = \frac{p_{ij}\sqrt{\lambda_i}}{\sigma_{z_j}} \tag{22}$$

Table 5 reports estimates of model (18) using as threshold variable a weighted average of the first three $PC_i, i = 1, 2, 3$, denoted z_{ft}, using as weights their relative variance. The three PC_i explain 90% of the variation of the data.

The PCs are based on threshold variables z_j: $labsh, u_l, \pi, i_s, i_{s,wx}$ and i_l which, as we saw, imply significant threshold effects. Variable cu is excluded given its relatively shorter sample. To avoid the influence of differences in the measurement units and quantities of z_j's on the $PC_i s$, we rely on the standard-score transformation of the original variables z_j: $\tilde{z}_j = \frac{z_j - \mu_{z_j}}{\sigma_{z_j}}$ where μ_j and σ_{z_j} are the first and second moments of z_j (e.g., Timm, 2002). This implies that all variables z_j contribute equally to the sample variation of the $PC_i s$. In the table, we present the estimates of ρ_{1j} for the 1st factor PC_1 explaining roughly 80% of the variability of the data.

Results suggest the following. The AIC shows that the version of the threshold factor model constitutes a superior specification of the data compared to those of table 3. Coefficients β_1^* and β_2^* demonstrate that there are substantial differences in the OL slope between regimes (with the first being the flatter) and that the speed of transition between regimes is at very high levels if we control for the endogeneity of threshold variable. Such high transition rates were found for the $labsh, i_s$ and $i_{s,wx}$ threshold cases (recall table 3).

Figure 5 demonstrates that the threshold factor variable constructed, z_{ft}, provides smooth and broadly consistent patterns of movements of its underlying variables z_j and its accompanying transition probabilities implied by the model estimates. It demonstrates vividly that OL is less strong than we might have imagined. The steeper OL relationship can be mainly attributed to the economic changes at the end of sixties, and the early-middle of seventies and eighties, discussed previously. After the end-80s/early-90s, there essentially exists only one regime transition from a steeper to a flatter OL relationship implying a near 1-to-1 Okun coefficient.

The consistency of our regime switching resulting to those implied by the estimates with the individual threshold variables can be justified by the estimates of the loading coefficients ρ_{1j}. The positive values of ρ_{1j} for variables $labsh, \pi, i_s, i_{s,wx}$ and i_l mean that higher values of these variables (above their threshold levels) are associated with higher values of the threshold factor variablez_{ft}, which favor a regime shift to the steeper OL relationship. Negative values of ρ_{1j} for u_l imply the opposite effects. The values of z_{ft} decrease when u_l increase thus favoring switching to the flatter OL regime.

8 Conclusions

Okun's law is a building block of many macro-econometric models and often considered an empirical regularity. It is widely used in policy and forecasting environments, and macroeconomic model building where cyclical output and unemployment are a core relationship in the model structure. We examined the possibility of asymmetries in OL. Specifically, we considered a nonlinear, smooth transition regression model, exploring a variety of (endogenous) candidate threshold variables. These channels were informed by Okun's insight that the (nature and strength) of the trade off would be impact from the demand, supply and policy side (and from the subsequent literature).

Our paper makes two main contributions. The first is methodological, the second is applied and empirical.

In terms of **methodology**, our work has two distinct components:

(a) We allow for endogeneity in the threshold variable using copulas to capture the dependence between the disturbance term and the threshold variable in the context of a LSTR model. The last model constitute a more general framework to study transition effects in macroeconomic relationships, which nests the linear and the standard threshold model as special cases. The copula approach constitutes a quite flexible and general approach to deal with the problem of endogeneity; it does not depend on valid instruments for parameter estimation, nor does it necessarily require normality of the marginal distribution of the threshold variable.

(b) We also suggest a *new* testing procedure to test for smooth transition threshold effects against linearity under endogeneity of the threshold variable.

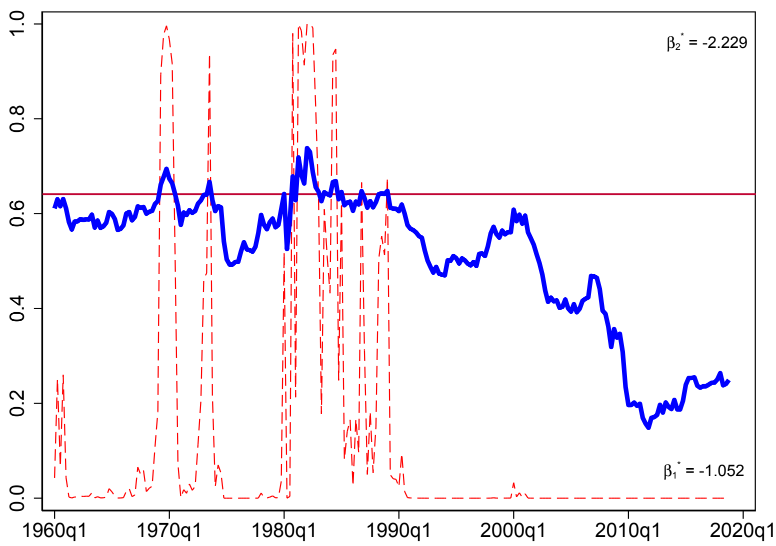
TABLE 5: Principal Component Analysis and Model Estimates

PC ESTIMATES										
		i, v	1	2	3	4	5	6		
		λ_i	5.1560	1.0400	0.2520	0.1130	0.0680	0.0048		
		ρ_v^2	0.7360	0.8850	0.9730	0.9890	0.9990	1.0000		
		j	$labsh$	π	i_s	$i_{s,wx}$	i_l	u_l		
		ρ_{1j}	0.7890	0.7620	0.9340	0.9500	0.8740	-0.7720		
THRESHOLD ESTIMATES										
ρ	β	b	δ	γ[1]	$100 \times \omega_1$	$100 \times \omega_2$	β_1^*	β_2^*	AIC	$\mathcal{D}_2$
0.751***	-0.262**	-0.555***	0.641	2.300	0.100	1.658***	-1.052	-2.229	-907.774	11.630
(0.066)	(0.134)	(0.160)	{0.620-0.642}	{2.26-2.33}	(0.376)	(0.718)				

Note. The table presents results of the PCA of the variables z_j triggering significant threshold effects and estimates of model (18) using the first three PC factors (weighted by their standard deviation) as the threshold variable. These estimates cover the case that the endogeneity of the threshold factor is controlled for based on our suggested copula method. Bootstrap standard errors and confidence intervals are, respectively, given in parentheses and braces. The terms β_1^* and β_2^* are as defined in (19).

[1]: Parameter γ and its confidence intervals should be divided by the variance 0.022 in order for the speed-of-transition parameter to be scale free.

Figure 5: Transition Probability (Principal Components)



Note. This figure graphically presents estimates of the threshold principle component factor model against its transition probability. The blue solid line represent the threshold PC factor, and the red dashed line represents the transition probability. The horizontal line represents the estimated threshold parameter, $\hat{\delta}$. The β^* values on the rhs edges are the long-run Okun parameters in the respective regimes. Note, in this case there is no need for both a lhs and rhs axis in the figure since both variables lie in the unit interval.

Both the copula based approach and the linearity testing procedure suggested are evaluated throughout a MC study.

Regarding **applied empirical** lessons, which is our second main contribution, we establish that threshold effects can be detected in Okun's relationship. We find mainly a combination of structural and policy-related variables accounts for changes in the Okun's law trade off in recent decades. Thus the notion that OL (though a cyclical comovement) is only affected by cyclical variables is incorrect. This conclusion is bolstered by combing these threshold candidates into a single factor. We found regime-like behavior with Okun's coefficient rising (or flattening) over time: from around -2.23 over the 1960s-1980s, then a slow transition to a value around -1 (i.e., more than halving in value). Thus the unemployment gap is increasingly associated with a smaller output gap. Put another way, the unemployment costs of output fluctuations are greater in more recent decades than before. This is an important finding.

Moreover, whilst the Great Recession accelerated that rise, interestingly, the bulk of the change occurred beforehand. This, in turn, corroborates our finding that both structural and non structural factors were at play. A key variable in that respect is the labor share of income. Perhaps this is not surprising since the importance of this variable in influencing growth and policy transmission has been widely discussed. Knowledge of the possibilities of these shifts and their determinants, improves the forecasting performance of Okun's law over methods missing the these shifts and their endogenous effects (see Appendix C).

Finally, the work done here could be fruitfully extended. An interesting controversy in the literature is whether OL holds outside the US, in particular for countries characterized by less flexible labor and product markets. Could the relationship be identified as asymmetric for countries of the European Union, for instance? If so, are there commonalities in the threshold variables found?

More generally, though, our endogenous threshold LSTR model could prove useful for additional applications such as in studies of growth, trade and finance where threshold models have often been employed to analyze asymmetries. The use of our new approach may yield fresh or more robust insights.

References

Andrews D. (1993). Tests for Parameter Instability and Structural Change With Unknown Change Point. *Econometrica*, 61, 821–856

Andrews, D. and W. Ploberger (1994). Optimal Tests When a Nuisance Parameter is Present only under the Alternative, *Econometrica*, 62, 1383–1414.

Antonakis, J., Bendahan, S., Jacquart, P. and R. Lalive (2014). Causality and Endogeneity: Problems and Solutions, in *The Oxford Handbook of Leadership and Organizations*, edited by D. Day, Oxford University Press.

Auer, P. and Gervini, D. (2008). Choosing principal components: A new graphical method based on model selection, *Communication in Statistics - Simulation and Computation*, 37, 962-977.

Auerbach, A. and Y. Gorodnichenko, (2012). Measuring the Output Responses to Fiscal Policy, *American Economic Journal: Macroeconomics*, 4(2), 1–27.

Attfield, C. L. F. and B. Silverstone (1997) Okun's Coefficient: A Comment, *Review of Economics and Statistics*, 79, 2, 326–329

Bernanke B. (2012) Recent Developments in the Labor Market, Speech given at the National Association for Business Economics Annual Conference, Arlington, Virginia, March 26, 2012.

Bernanke B, Boivin J. and P. Eliasz (2005). Measuring monetary policy: a factor augmented autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120: 387–422.

Blinder, A. (1997). Is there a core of practical macroeconomics that we should all believe? *American Economic Review*, 87, 240–243.

Chan, K. S. (1993). Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model, *Annals of Statistics*, 21, 520–533.

Christopoulos, D., McAdam, P. and E. Tzavalis (2021). Dealing with Endogeneity In Threshold Models using Copulas: Revisiting the Foreign Trade Multiplier, *Journal of Business & Economic Statistics*, 39, 1, 166-178.

Courtney, H. G. (1991). The Beveridge curve and Okun's law: a re-examination of fundamental relationships in the United States. MIT: PhD thesis.

Cuaresma, J. C (2003). Okun's Law Revisited, *Oxford Bulletin of Economics and Statistics*, 65(4), 439–451.

Davidson, R. and J. G. MacKinnon (2007). Improving the reliability of bootstrap tests with the fast double bootstrap, *Computational Statistics & Data Analysis*, 51(7), 3259–3281.

Erdely A. (2017). Copula-based piecewise regression, In: Ubeda F. M., de Amo Artero E., Durante F., Fernandez Sanchez J. (eds) *Copulas and Dependence Models with Applications*, pp. 69-81. Springer

Fernald, J. G. (2018). A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. FRBSF Working Paper 2012-19 (current update).

Fernald, J. G., Hall, R. E., Stock, J. H. and M. W. Watson (2017). The Disappointing Recovery of Output after 2009, *Brookings Papers on Economic Activity*, 48(1), 1–81.

Fernald, J. G. and C. Wang (2016) Why Has the Cyclicality of Productivity Changed? What Does It Mean?, *Annual Review of Economics*, 8(1), 465-496.

Franses, P. H. and D. van Dijk (2000). *Non-linear Time Series Models in Empirical Finance*, Cambridge University Press.

Freeman, D. G. (2001). Panel Tests of Okun's law for Ten Industrial Countries, *Economic Inquiry*, 39 (4), 511–523.

Giannone, D., M. Lenza and G. E. Primiceri (2019) Economic Predictions with Big Data: The Illusion of Sparsity, *Econometrica*, forthcoming.

Granger, C. W. and T. Teräsvirta (1993). *Modelling non-linear economic relationships*, Oxford University Press.

Hansen, B. E. (1996). Inference when a Nuisance Parameter is not Identified under the Null Hypothesis, *Econometrica*, 64, 413–430.

Hansen B. E. (1997). Approximate Asymptotic p Values for Structural-Change Tests. *Journal of Business and Economic Statistics*, 15, 60–67.

Hansen, B. E. (1999). Threshold Effects in Non-Dynamic Panels: Estimation, Testing and Inference, *Journal of Econometrics*, 93, 345–368.

Harris, R. and B. Silverstone (2001). Testing for asymmetry in Okun's law: cross-country comparison, *Economics Bulletin*, 5, 1–13

IMF (2010). Unemployment dynamics during recessions and recoveries: Okun's Law and beyond. WEO chapter, April.

Joe, H. (2014). *Dependence Modeling with Copulas*, Chapman and Hall.

KılıÇ, R. (2011). Testing for Cointegration and Nonlinear Adjustment in a Smooth Transition Error Correction Model, *Journal of Time Series Analysis*, 32, 647–660.

KılıÇ, R. (2016). Tests for Linearity in STAR Models: SupWald and LM-type Tests, *Journal of Time Series Analysis*, 37, 660–674.

Klump, R., McAdam, P. and A. Willman (2007). Factor Substitution and Factor-Augmenting Technical Progress in the United States, *Review of Economics and Statistics*, 89(1), 183-192.

Kourtellos, A., T. Stengos and C. M. Tan (2016). Structural Threshold Regression, *Econometric Theory*, 32, 4, 827–860.

Kourtellos, A., T. Stengos and Y. Sun (2021) Endogeneity in semiparametric threshold regression, *Econometric Theory*, forthcoming.

Lee, J. (2000). The robustness of Okun's law: Evidence from OECD countries, *Journal of Macroeconomics*, 22(2), 331–356.

León-Ledesma, M. A, McAdam, P. and A. Willman (2010). Identifying the Elasticity of Substitution with Biased Technical Change, *American Economic Review*, 100(4), 1330-1357.

Leybourne, S., P. Newbold and D. Vougas (1998). Unit Roots and Smooth Transitions, *Journal of Time Series Analysis*, 19, 83–97.

Luukonen, R., P. Saikkonen and T. Teräsvirta (1988). Testing Linearity against Smooth Transition Autoregressive Models, *Biometrika*, 75, 491–499.

Mankiw, N. G. (2015) *Macroeconomics*, Worth Publishers.

Manyika. J. et al. (2019). A new look at the declining labor share of income in the United States, Discussion Paper, McKinsey Global Institute

Marcellino, M., Stock, J. H. and M. W. Watson (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series, *Journal of Econometrics*, 135, 499–526

Neely, C. J. (2010). Okun's law: output and unemployment, *Economic Synopses*, Federal Reserve Bank of St. Louis.

Okun, A. M. (1962). Potential GNP: Its Measurement and Significance, *Proceedings of the Business and Economics Statistics Section*, 98–104, Alexandria, VA: American Statistical Association.

Okun, A. M. (1974). Unemployment and Output in 1974. *Brookings Papers on Economic Activity*, 5(2): 495–506.

Patton, A. J. (2006). Modelling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47(2), 527-556.

Park, S. and S. Gupta (2012). Handling Endogenous Regressors by Joint Estimation Using Copulas, *Marketing Science*, 31, 4, 567–586.

Peres, P. R., D. J. Jackson and K. M. Somers (2004). How many principal components? Stopping rules for determining the number of non-trivial axes revisited, *Computation Statistics & Data Analysis*, 49, 974-997.

Perman, R. and C. Tavéra (2007). Testing for convergence of the Okun's law coefficient in Europe. *Empirica*, 34(1), 45–61.

Perman, R., S. Gaetan and C. Tavéra (2015). Okun's law – A Meta-analysis, *Manchester School*, 83, 1,101-126.

Prachowny, M. F. J. (1993). Okun's law: Theoretical Foundations and Revised Estimates, *Review of Economics and Statistics*, 75 (2), 331–336.

Rothman, P. (1998). Forecasting Asymmetric Unemployment Rates, *Review of Economics and Statistics*, 80(1), 164–168.

Siburg, K.F and P.A. Stoimenov (2008) Gluing Copulas, *Communications in Statistics – Theory and Methods*, 37, 3124-3134.

Silvapulle, P., I. Moosa and M. Silvapulle (2004). Asymmetry in Okun's law, *Canadian Journal of Economics*, 37 (2), 353–374.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

Stock J. H. and M. W. Watson (2002) Forecasting Using Principal Components From a Large Number of Predictors, *Journal of the American Statistical Association*, 97, 1167-1179.

Teräsvirta T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association*, 89: 208–218.

Timm, N. H. (2002). *Applied Multivariate Analysis*, Springer.

van Dijk, D., T. Teräsvirta and P. H. Franses (2002). Smooth Transition Autoregressive Models – A Survey of Recent Developments, *Econometric Reviews*, 21(1), 1–47.

Virén, M. (2001). The Okun curve is non-linear, *Economics Letters*, 70(2), 253–257.

Woodford, M. (2012) Methods of policy accommodation at the interest-rate lower bound, in *The Changing Policy Landscape*, Federal Reserve Bank of Kansas City, Jackson Hole Wyoming: 185–288.

Wu, J. C. and F. D. Xia (2016) Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound, *Journal of Money, Credit, and Banking*, 48(2–3), 253–291.

A Monte Carlo

In this section, we present the data generating processes (DGP) and results of the Monte Carlo (MC) exercise touched upon in the main text. Our first exercise examines the performance of the estimation method suggested to estimate the location parameter δ accurately and the ability of our method to successfully control for the endogeneity of threshold variable z_t in the LSTR model. The second exercise (in Appendix B) evaluates the power performance of test statistic $LR^{\sup}$.

The DGP that we consider to estimate δ is as follows:[1]

$$y_t = (\beta_{11} + \beta_{12}x_{2t})(1 - \boldsymbol{g}(z_t;\gamma,\delta)) + (\beta_{21} + \beta_{22}x_{2t})\boldsymbol{g}(z_t;\gamma,\delta) + \varepsilon_t, \tag{A.1}$$

where $x_{2t} \sim IID\mathcal{N}(0,1)$ and $x_{3t} \sim IID\mathcal{N}(0,1)$ are exogenous, i.e., $\mathbb{E}(\varepsilon_t|x_{2t}) = 0$, and the values of the slope coefficients across the two regimes of the model are as follows:

$$\text{Regime 1:} \quad \beta_{11} = \quad 0.5, \quad \beta_{12} \quad = 2.0$$

$$\text{Regime 2:} \quad \beta_{21} = -0.7, \quad \beta_{22} \quad = \{1.5, 3.0\}.$$

These values imply quite small differences in the slope coefficients of model (A.1) across its two regimes, which are difficult to detect. They are chosen to highlight the ability of our method to lead to accurate estimates of δ even for very small-size shifts of the slope coefficients of the model. For simplicity and without any loss of generality, we drop the lagged dependent variable in these Monte Carlo exercises.

For the regression error term ε_t and threshold variable z_t of the above mentioned DGP, we consider the normal and the Student-t distribution, with four degrees of freedom. Using data from the Student's t distribution will show if our method, based on the transformations of the quantile function of the standard normal distribution (i.e., Φ^{-1}), can be proved robust to a misspecification of the true distributions of ε_{ht} and z_t, like the Student's t often assumed in econometrics. For reasons of space, we present results for the cases that the distribution of ε_t, and correlation structure between ε_{ht} and z_t do not change across the two regimes of the model. For z_t, we assume $z_t \sim IID(z_\mu, 1)$, with $z_\mu = 3.0$, while, for the correlation coefficient between ε_{ht} and z_t, we consider the following set of values:$\rho_{\varepsilon z} = \{0.0, 0.55, 0.75\}$, for $h = \{1,2\}$. The structure of the threshold variable is given as follows:

$$z_t = \delta + c_{\varepsilon z}v_t + \zeta_t \tag{A.2}$$

where $\varepsilon_t \sim t_{DF=4}$ and $\zeta_t \sim t_{DF=4}$, for all t. The value of coefficient $c_{\varepsilon z}$ is chosen to control for the degree of correlation between ε_t and z_t.

The threshold value δ is set to the 25% percentile (1st quantile) of the distribution of z_t. We also examined threshold values at the 75% percentile (3rd quantile) of this distribution, but the MC results do not change qualitatively. The values of the speed of adjustment considered are set to $\gamma = \{1.5, 3.5\}$. These values reflect the cases that the transition between the two regimes is low and high, respectively. Following other studies, we treat the above values of γ as known in our analysis, reflecting that our interest is focused on the estimation bias of δ. Since we are focused on

[1] Both the DGP and the values of its parameters considered in our simulation analysis are close to those considered in the simulation studies of, for instance, Lundbergh et al. (2003).

the performance of our method to control for the endogeneity of z_t on the estimates of δ, we do not report results for the remaining parameters of the model, i.e., ω_1, ω_2 for reasons of space.

We consider sample sizes of $T = \{50, 250\}$ observations and carry out $1,000$ iterations. For all iterations, we calculate the bias and the root mean square error of the estimator of δ. In Table A.1, we present average values of the above metrics, over all iterations, denoted $BIAS$ and $RMSE$, respectively. The table presents different sets of results. Panel A presents results ignoring the problem of the threshold variable endogeneity in the estimation. Panel B controls for this problem, based on our method by including in the rhs of (A.1) the bias correction terms z^*_{ht}, $h = \{1, 2\}$.

Panel C presents results for the case that Kourtellos's et al. (2016) approach for threshold models, appropriately modified for the LSTR model, is employed to control for the endogeneity problem of z_t. This approach adjusts equation (A.1) by estimating the expectation terms $\mathbb{E}(\varepsilon_t|z_t \leq \delta)$ and $\mathbb{E}(\varepsilon_t|z_t > \delta)$, capturing the bias of z_t across the two regimes of the model, based on the inverse Mills ratio terms assuming that both ε_t and z_t are normally distributed. To calculate these ratios, we assume that $z_t \sim IID\mathcal{N}(z_\mu, 1)$ has the following single factor presentation:

$$z_t = z_\mu + 0.95\zeta_t + e_{zt}, \tag{A.3}$$

where $\zeta_t \sim IID\mathcal{N}(0, 0.5)$ and $e_{zt} \sim IID\mathcal{N}(0, 0.5)$ and $z_\mu = 3.0$. Finally, Panel D presents results for the case that our method is applied to the case that ε_{ht} and z_t are jointly Student-t distributed.

Results lead to several interesting conclusions. First, they indicate that ignoring the endogeneity of threshold variable z_t causes serious biases in the estimates of threshold parameter δ. Specifically, it tends to overestimate the true value of δ. Moreover, as expected, the bias of δ is substantial in magnitude when the correlation between z_t and ε_t is high, i.e., 0.75. The magnitude of the bias also depends on the different values of γ considered. The bias is bigger, the smaller value of γ considered, i.e. $\gamma = 1.5$, than $\gamma = 3.5$, and it remains even if T increases.

A second conclusion is that our method can successfully control for the endogeneity problem of the threshold variable. The method can substantially reduce the estimation bias of δ. This is true even if the sample size is small (i.e., $T = 50$). As expected, the bias of δ reduces as the sample increases (i.e., $T = 250$). The bias also reduces, when γ increases (i.e., the LSTR model approaches the TAR, where the shifts of the model are faster across the two regimes). Similar conclusions to the above also hold for the $RMSE$. These results are also robust to the different distribution of ε_{ht} and z_t considered, namely the Student's t. This is clearer for the case of the larger size of $T = 250$.

Note that, for the case that both ε_{ht} and z_t are normally distributed, the performance of our method successfully compares to that of KST. Actually, our method seems to have better small sample (i.e., $T = 50$) performance relative to KST.

Finally, the results indicate that the estimation of the augmented regression model, which controls for the endogeneity of the threshold variable, leads to unbiased estimates of δ even if there is no threshold variable endogeneity (i.e., $\rho_{\varepsilon z} = 0$). This result is more clear cut for the cases that T is large.

References

Lundbergh, S., Teräsvirta, T. and D. van Dijk (2003). Time–Varying Smooth transition Autoregressive Models, *Journal of Business & Economic Statistics*, 21, 104–121.

Table A.1: Monte Carlo results of the BIAS and RMSE of the estimator of δ

γ	$T/\rho_{\varepsilon z}$	0.00		0.55		0.75		0.00		0.55		0.75	
		(i): $\beta_{22} = 1.5$						(ii) $\beta_{22} = 3.0$					
		$BIAS$	$RMSE$	$BIAS$	$RMSE$	$BIAS$	$RMSE$	$BIAS$	$RMSE$	$BIAS$	$RMSE$	$BIAS$	$RMSE$
A: Ignoring the endogeneity of z_t (z_t and ε_t are normally distributed)													
1.5	50	−0.613	1.732	−0.975	2.986	−1.002	3.238	−0.811	2.420	−0.990	3.257	−0.991	3.464
1.5	250	−0.213	0.417	−0.827	1.697	−1.043	2.113	−0.153	0.324	−0.529	0.982	−0.716	1.250
3.5	50	−0.398	0.961	−0.816	2.120	−0.929	2.533	−0.331	1.197	−0.647	1.651	−0.694	1.879
3.5	250	−0.020	0.057	−0.439	0.899	−0.782	1.528	−0.012	0.049	−0.174	0.343	−0.358	−0.592
B: Controlling for the endogeneity of z_t (z_t and ε_t are normally distributed)													
1.5	50	−0.775	1.611	−0.729	1.829	−0.567	1.799	−0.732	1.522	−0.650	1.588	−0.485	1.561
1.5	250	−0.456	0.662	−0.343	0.614	−0.293	0.647	−0.381	0.546	−0.245	0.425	−0.165	0.421
3.5	50	−0.619	1.331	−0.503	1.287	−0.409	1.289	−0.522	1.098	−0.424	1.007	−0.266	0.884
3.5	250	−0.149	0.204	−0.135	0.204	−0.114	0.167	−0.073	0.111	−0.065	0.091	−0.072	0.092
C: Controlling for the endogeneity of z_t based on Mill's ratios (z_t and ε_t are normally distributed)													
1.5	50	−0.891	2.363	−0.935	2.433	−0.928	2.423	−0.818	2.130	−0.863	3.202	−0.867	2.271
1.5	250	−0.398	0.811	−0.425	0.863	−0.428	0.873	−0.314	0.635	−0.315	0.661	−0.312	0.653
3.5	50	−0.691	1.559	−0.678	1.514	−0.667	1.500	−0.530	1.218	−0.572	1.295	−0.585	1.317
3.5	250	−0.091	0.189	−0.126	0.231	−0.129	0.234	−0.034	0.088	−0.057	0.125	−0.062	0.031
D: Controlling for the endogeneity of z_t (z_t and ε_t are Student-t distributed)													
1.5	50	−0.920	2.516	−0.993	2.615	−0.901	2.378	−0.878	2.425	−0.902	2.434	−0.846	2.214
1.5	250	−0.240	1.104	−0.516	1.381	−0.683	1.638	−0.133	0.847	−0.334	1.010	−0.518	1.295
3.5	50	−0.678	2.012	−0.696	1.954	−0.630	1.790	−0.599	1.838	−0.619	1.737	−0.499	1.393
3.5	250	−0.004	0.401	−0.231	0.508	−0.337	0.656	−0.035	0.218	−0.108	0.245	−0.150	0.271

B Power of the Test Statistic $LR^{\sup}$

We present values of the power performance of statistic $LR^{\sup}$ for the LSTR model (A.1), employed in our previous MC exercise, where the threshold parameter value δ is set at its 3rd quantile. For reasons of space, we do not report results for the case that δ is at the 1st quantile and we consider only the case of $\beta_{22} = 1.5$. The remaining slope parameters are set as in the previous MC exercise. We also consider the same simulation scenarios with that exercise for $\rho_{\varepsilon z} = \{0.0, 0.55, 0.75\}$ and $T = \{50, 250\}$.

To calculate the power of the test statistic, we need first to obtain its distribution under the null hypothesis H_0: $\beta_1 = \beta_2 = \beta$ and obtain its critical value, corresponding to the 95th-quantile of its simulated distribution. To this end, for each iteration we generate the error term $\varepsilon_t \sim IID(0,1)$ and threshold variable $z_t \sim IID(3,1)$ from a joint distribution of them, while x_{2t} and x_{3t} are generated as $x_{2t} \sim IID\ \mathcal{N}(0,1)$ and $x_{3t} \sim IID\mathcal{N}(0,1)$, respectively. Given these generated series, we then generate series y_t under null hypothesis based on model $y_t = x_t'\beta + \varepsilon_t$, where $\beta = \frac{1}{2}(\beta_1 + \beta_2)$; β_1 and β_2 are defined as in the previous exercise.

Based on the generated series y_t, next we estimate model (A.1) under both the null and alternative hypotheses and, then, we calculate the test statistic $LR^{\sup}$, over all possible values of γ and δ, based on $1,000$ iterations. In so doing, note that we trim out the top and bottom 10 percentiles of the distribution of δ, while for γ we rely on the set of values $Q_\gamma = \left[\frac{1}{10\sigma_z}, \frac{100}{\sigma_z}\right]$.[2] Also, under the alternative hypothesis, the model is adjusted by the bias correction terms z^*_{ht}, $h = \{1, 2\}$, to control for the endogeneity of the threshold variable.

Given the critical values of $LR^{\sup}$, at the 5% level, the power of the test, which represents rejection frequencies of the above null hypothesis, is calculated by generating data under the alternative hypothesis the alternative hypothesis H_a: $\beta_1 \neq \beta_2$, namely model (A.1). The remaining steps of the MC exercise are as above. For each iteration, the error term ε_t, the variables x_{2t} and x_{3t}, and the threshold variable z_t are generated as before, while in estimating δ we trim out the top and bottom 10 percentiles of its distribution, while for γ we rely on the set of values $Q_\gamma = \left[\frac{1}{10\sigma_z}, \frac{100}{\sigma_z}\right]$. To see if the power of the test depends on ignoring the issue of the endogeneity of the threshold variable z_t, we also present results for the case that model (A.1) is not adjusted by the terms z^*_{ht}, $h = \{1, 2\}$ to calculate statistic $LR^{\sup}$.

The results of the power of the test are reported in Table B.1. These indicate that the power of the test statistic $LR^{\sup}$ is high and close to unity for all the cases of γ and $\rho_{\varepsilon z}$ considered. They also indicate that ignoring the endogeneity of the threshold variable leads to a version of the test statistic which has less power. However, this is more clear for the case that $\gamma = 3.5$. Another interesting result is that the power performance does not depend on the adjustment of model (5) for possible endogeneity of threshold variable (see model (14)) when $\rho_{\varepsilon z} = 0$. This result was expected. Unless there is a serious degrees of freedom problem, it means that augmentation of this auxiliary regression by the transformed variables z^*_{ht} does not affect the power of statistic $LR^{\sup}$ and, hence, it can be safely implemented, in practice, for all possible values of the correlation coefficient $\rho_{\varepsilon z}$. The above results are robust to the different values of γ and T considered.

[2] We found that extension of Q_γ to include higher values of γ does not affect our simulation results.

TABLE B.1: Power of test statistic LR^{sup} with $\beta_{22} = 1.5$

	No Endogeneity		Controlling For Endogeneity	
$\rho_{\epsilon z}$		$\gamma = 1.5$		
0.00	0.910	0.926	0.896	0.936
0.55	0.837	0.886	0.870	0.934
0.75	0.781	0.826	0.868	0.925
$\rho_{\epsilon z}$		$\gamma = 3.5$		
0.00	0.852	0.893	0.827	0.896
0.55	0.766	0.833	0.812	0.896
0.75	0.687	0.762	0.804	0.895
T	50	250	50	250

Note. The table presents the power of test statistic

$$LR^{\text{sup}} \equiv \sup_{(\gamma,\delta)\in Q_\gamma \times Q_z} LR(\gamma, \delta)$$

under the alternative hypothesis H_a: $\beta_1 \neq \beta_2$, at the 5% nominal size, for model (A.1). The critical values of the test statistic are simulated under the null H_0: $\beta_2 = \beta_2$, for alternative N and T. We use the parameter values from Appendix A but concentrate on the $\beta_{22} = 1.5$ case.

C Prediction with the Threshold Factor Model

In this section we examine the framework in light on forecasting exercises. Specifically, we evaluate the predictive properties of the dynamic OL allowing for STR effects, $\mathcal{M}^{DOL}_{STR}$, based on (18) and threshold factor variable z_{ft}. We investigate if it can improve upon the performance of the exogenous threshold case as well as the AR and dynamic OL benchmarks to predict y_t. The latter two models were those reported in Table 1. These comparator models we label, respectively, as $\mathcal{M}^{DOL}_{STR(X)}$, $\mathcal{M}^{AR}$ and $\mathcal{M}^{DOL}$.

We conduct *in-* and *out-of-sample* prediction exercises based on pair model comparisons. To evaluate the robustness of our results to possible missing sources of non-linearities in the data and structural instabilities (uncertainty), the out-of-sample prediction exercise is based on rolling regression estimates of the models. For both exercises, we report results based on the realized values of cyclical unemployment series u_t and the sample estimates of the threshold factor variable, but *also* based on predicted values of u_t and z_{ft} using sample information at time $t-1$ and the exponential smoothing method suggested by Bergmeir et al. (2016) (henceforth BHB). These predictions make comparison with the $\mathcal{M}^{AR}$ model predictions more fair, as they are based on sample information at the same time.

The BHB method constitutes an expansion of the standard exponential method of forecasting. It decomposes a time series into its trend, seasonality and remainder components, where the latter is bootstrapped and is added to the trend and seasonal components to generate a new bootstrapped series. Based on these series, we obtain one-quarter ahead predictions based on the exponential smoothing method. We take the median of these predictions to provide the resulting prediction of the model since it is less sensitive to outliers.[3] The bootstrap aggregation involved in the BHB method can improve the accuracy of predictions, by addressing issues of data irregularities and parameter uncertainty.

C.1 In-sample prediction

Table C.1 presents the in-sample exercise. Panel A presents the case that realized values of u_t and estimates of z_{ft} are used to predict y_t, and Panel B for the case that predicted values of u_t and z_{ft} are used, based on sample information at time $t-1$ and the BHB method, described above. We report the root mean square and mean absolute error (respectively, $RMSE$ and MAE) and fit tests based, respectively, on the Diebold and Mariano (1995) statistic (DM), and Vuong's (1989) statistic (see also Rivers and Vuong (2002)).[4] These are one-sided test statistics calculated for all model pairs. Under the null, they assume that the models compared are equivalent, and under the alternative that the forecasts of the $\mathcal{M}^{DOL}_{STR}$ dominate those of the others ($\mathcal{M}^{DOL}$, or $\mathcal{M}^{AR}$, or $\mathcal{M}^{DOL}_{STR(X)}$ has superior forecasting properties compared to the $\mathcal{M}^{DOL}$. The long-run variance of the DM test is calculated based on the HAC estimator (see Harvey et al. (1997)).

$\mathcal{M}^{DOL}_{STR}$ has the globally smallest $RMSE$ and MAE and its predictive superiority can be formally justified by all the tests reported at the 5% level (or below). Most tests select model $\mathcal{M}^{DOL}_{STR}$. Its superiority means that capturing the instability of the OL relationship is an important step in improving its predictive performance. These results hold for both the A and B cases. For the

[3] Note that, to stabilize the variance of time series, the BHH method relies also on the Box-Cox transformations of the original series, requiring an inversion of the bootstrapped series to obtain the ones used for forecasting.

[4] Note that the Voung test statistic is more appropriate for non-nested models. We present its values for completeness and to appraise the fitness of the models, given that it is based on an information criterion.

Table C.1: In-Sample Forecast Exercises

		A: Sample values of u_t and estimates of z_{ft}		B: Predicted values of u_t and z_{ft}	
		$RMSE$	MAE	$RMSE$	MAE
$\mathcal{M}^{AR}$		0.00512	0.00386	0.00512	0.00386
$\mathcal{M}^{DOL}$		0.00469	0.00355	0.00464	0.00356
$\mathcal{M}^{DOL}_{STR}$		0.00432	0.00332	0.00437	0.00335
$\mathcal{M}^{DOL}_{STR(X)}$		0.00446	0.00342	0.00443	0.00396
		Diebold–Mariano test			
$\mathcal{M}^{DOL}$	*vs* $\mathcal{M}^{DOL}_{STR}$	4.074 [0.0001]		3.953 [0.0001]	
$\mathcal{M}^{AR}$	*vs* $\mathcal{M}^{DOL}_{STR}$	2.636 [0.0044]		2.477 [0.0070]	
$\mathcal{M}^{DOL}_{STR(X)}$	*vs* $\mathcal{M}^{DOL}_{STR}$	1.8222 [0.0348]		1.409 [0.0800]	
		Voung test			
$\mathcal{M}^{DOL}$	*vs* $\mathcal{M}^{DOL}_{STR}$	-4.640 [0.0001]		-4.142 [0.0010]	
$\mathcal{M}^{AR}$	*vs* $\mathcal{M}^{DOL}_{STR}$	-2.263 [0.0035]		-2.554 [0.0050]	
$\mathcal{M}^{DOL}_{STR(X)}$	*vs* $\mathcal{M}^{DOL}_{STR}$	-1.833 [0.0344]		-1.424 [0.0773]	

Note. This table shows the in-sample forecasting metrics for the PC threshold model under cases (A) Sample values of u_t and estimates of z_{ft} and (B) Predicted values of u_t and z_{ft}. Numbers in brackets represent probability values.

latter, the superiority of the $\mathcal{M}^{DOL}_{STR}$ model is stronger in terms of the reported p-values. This may be attributed to the ability of the BHB method to reduce the effects of parameter and data uncertainty on the predictions of the models.

The $\mathcal{M}^{DOL}_{STR}$ model also notably improves upon the performance of its non-endogenous threshold counterpart $\mathcal{M}^{DOL}_{STR(X)}$. The close-to-5%, or less, p-values of the test statistics reported in Panel A reject null that $\mathcal{M}^{DOL}_{STR}$ and $\mathcal{M}^{DOL}_{STR(X)}$ have equal predictive ability, at 3.5% level. The former also has lower $RMSE$ and MAE.

Table C.2: Out-of-Sample Forecast Exercises

		A: Sample values of u_t and estimates of z_{ft}		B: Predicted values of u_t and z_{ft}	
		$RMSE$	MAE	$RMSE$	MAE
$\mathcal{M}^{AR}$		0.00298	0.00183	0.00440	0.00340
$\mathcal{M}^{DOL}$		0.00259	0.00188	0.00360	0.00232
$\mathcal{M}^{DOL}_{STR}$		0.00258	0.00174	0.00330	0.00237
$\mathcal{M}^{DOL}_{STR(X)}$		0.00263	0.00177	0.00380	0.00229
		$Diebold–Mariano\ Out\ of\ Sample$ test			
$\mathcal{M}^{DOL}$	$vs\,\mathcal{M}^{STR}_{DOL}$	4.018 [0.0411]		7.068 [0.0009]	
$\mathcal{M}^{AR}$	$vs\,\mathcal{M}^{DOL}_{STR}$	5.506 [0.0144]		4.419 [0.0250]	
$\mathcal{M}^{DOL}_{STR(X)}$	$vs\,\mathcal{M}^{DOL}_{STR}$	3.481 [0.0660]		5.905 [0.0005]	
		$Mixed–Window\ Out\ of\ Sample$ test			
$\mathcal{M}^{DOL}$	$vs\,\mathcal{M}^{STR}_{DOL}$	1.794 [0.0400]		2.957 [0.0015]	
$\mathcal{M}^{AR}$	$vs\,\mathcal{M}^{DOL}_{STR}$	2.214 [0.0130]		1.865 [0.0245]	
$\mathcal{M}^{DOL}_{STR(X)}$	$vs\,\mathcal{M}^{DOL}_{STR}$	1.462 [0.0718]		2.276 [0.0110]	

Note. This table shows the out-of-sample forecasting metrics for the PC threshold model under cases (A) Sample values of u_t and estimates of z_{ft} and (B) Predicted values of u_t and z_{ft}. Numbers in brackets represent probability values.

C.2 Out-of-sample prediction

Table C.2 presents the out-of-sample exercises. Panel A presents results based on the out-of-sample window realized values of u_t and values of z_{ft}. Panel B shows the case that predicted values of the one-quarter ahead are used based on the rolling window information and the BHB method. In addition to $RMSE$, MAE we show the test statistic of Clark and West (2007), CW and the mixed estimation window statistic of Calhoun (2016) (denoted CW and CMW, respectively).[5]

To calculate the values of the above metrics and statistics, our exercises are based on a fixed-length rolling window estimates of the models consisting of 232 observations, implying that we use a number of 80% observations to calculate the prediction errors. For each window, we update the estimates of the models and calculate the one-period ahead, out-of-sample prediction of y_t and their associated prediction errors based on the sample window estimates of the threshold factor z_t and u_t. We carry out similar pair comparisons of the models to those of the in-sample exercise. Under the null hypothesis, both the CW and CMW statistics are asymptotically normally distributed. For CMW, normality is achieved by estimating the benchmark model recursively

[5] These statistics have good size and power properties for one-period ahead out-of-sample forecasts (see Clark and McCracken (2011), for a survey).

(using an expanding window). In brackets, we give probability values for the CMW test based on the standard normal distribution and bootstrapped for the CW, as it is approximately normal in small samples (Clark and West (2007)). The results show that model $\mathcal{M}^{STR}_{DOL}$ outperforms the other models in terms of out of sample predictive accuracy. All tests select this model.

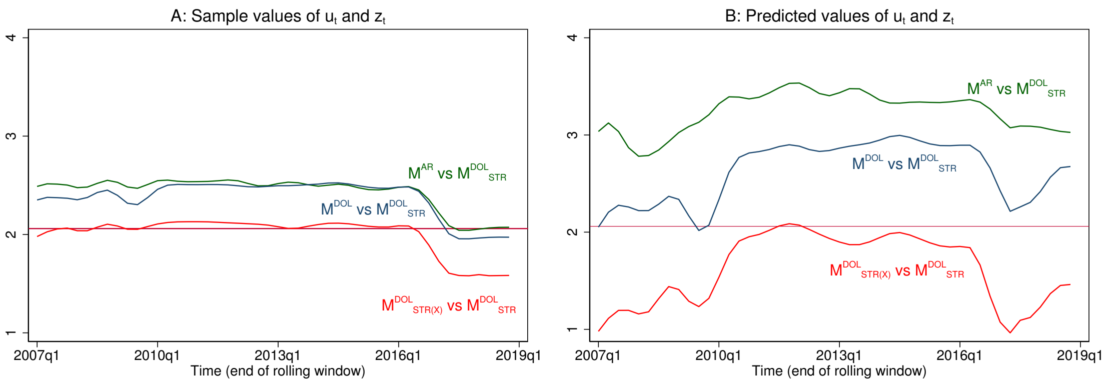
In addition Figure C.1 graphically reports the values of Giacomini's and Rossi (2010) fluctuation test statistic (GR) against the 5% critical value. This is based on the quadratic loss function (like the DM), and can examine the relative superiority that a model holds at each point in the out-of-sample period.[6] Conclusions echo those of the in-sample exercise. The $\mathcal{M}^{DOL}_{STR}$ model improves significantly upon the predictive performance of the $\mathcal{M}^{AR}$ (at a below 5% level), upon the predictive performance of the $\mathcal{M}^{DOL}$ model, and the version ignoring the endogeneity of z_{ft} at close to the 5% level.

The inspection of the graphs of the GR test reveals that the superiority of the $\mathcal{M}^{DOL}_{STR}$ model relative to the $\mathcal{M}^{AR}$ holds across all time points of the out-of-sample interval. For this pair of models, the values of the GR test statistic shows that model $\mathcal{M}^{DOL}_{STR}$ forecasts better than model $\mathcal{M}^{AR}$ as the null hypothesis of equal predictive ability at each point in time is rejected against the alternative at 5% critical level that $\mathcal{M}^{DOL}_{STR}$ forecasts the best at least one point in time.

These results hold based on both sets of the GR test values reported in the table; i.e., the sample values of u_t and the estimates z_{ft}, and their predicted out-of-sample values. Compared to the $\mathcal{M}^{DOL}$ model, the values of the GR test favor the $\mathcal{M}^{DOL}_{STR}$ model for both sets of results, at 5%. Finally, the model seems to perform better than that ignoring the endogeneity of the threshold variable for most tests reported in the tables and the figure, at the 10% level or below.

Summing up, our results provide interesting evidence in supporting the use of the $\mathcal{M}^{DOL}_{STR}$ model to provide short-term (one-quarter ahead) predictions of the output gap. The cyclical regime-dependent sensitivities of unemployment can help in this direction, as it can capture the instability of the OL relationship.[7]

FIGURE C.1: Out of Sample Forecasting One-Sided t Tests



Note. The figure plots the values of the GR (Giacomini-Rossi) fluctuation tests of the out-of-sample exercise. Panel **A** presents results for the case that the sample (realized) values of u_t and sample estimates of threshold factor zs_t are used to predict output gap, while panel **B** for the case that the predicted values of these variables one-period ahead are employed. The horizontal line represents the one-sided critical value.

[6] As with the DM, the long-run variance of all the above tests is calculated using the HAC estimator.

[7] We also examined the models through the Rossi and Sekhposyan (2016) forecast rationality test, without any indication that any of the models reject forecast rationality. Details available.

References

Bergmeir, C., R. J. Hyndman and J. M. Benitez (2016). Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation, *International Journal of Forecasting*, 32, 303-312.

Calhoun, G. (2016) An asymptotically normal out-of-sample test based on mixed estimation windows, mimeo, Iowa State University.

Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1): 291–311.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy, *Journal of Business & Economic Statistics*, 13, 253-263.

Giacomini, R. and B. Rossi (2010) Forecast comparisons in unstable environments, *Journal of Applied Econometrics*, 25(4), 595-620.

Harvey, D. L., Leybourne, S. J. and P. Newbold (1997). Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, 13, 281-291.

Rivers, D., and Q. Vuong (2002). Model Selection Tests for Nonlinear Dynamic Models, *Econometrics Journal*, 5, 1–39

Rossi B. and T. Sekhposyan (2016). Forecast Rationality Tests in the Presence of Instabilities, with Applications to Federal Reserve and Survey Forecastst, *Journal of Applied Econometrics*, 31(3), 507-532.

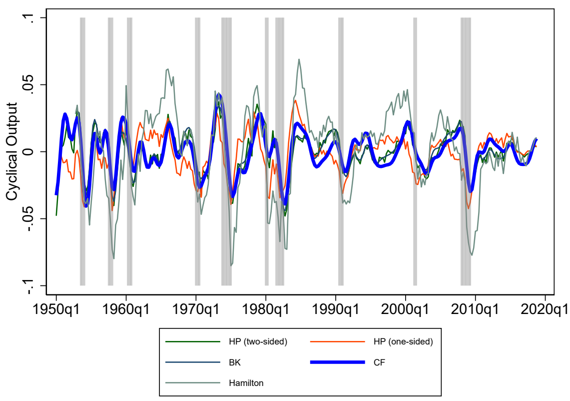
D Additional Material

D.1 Filtered Series

In Figure D.1 we plot various measures of the cyclical series of log output and unemployment. Specifically we show the one and two sided HP filter, the Christiano-Fitzgerald series (our preferred series marked in thick blue) and the Baxter-King and the Hamilton series. To a simple approximation their cyclical behaviors are quite similar, although typically the Hamilton filter is more volatile (see also the discussion in Schüler, 2018).

In Figure D.2 and Figure D.3, we plot the spectral density of the various filtered series in the order Series are ordered as HP 1 sided then 2-sided filter, Baxter-King Filter, Christiano-Fitzgerald, and Hamilton filter. The vertical lines are consistent with the business-cycle frequencies of 6 and 32 quarters. Thus, $32^{-1} = 0.031$ and $6^{-1} = 0.167$. It can be seen that the bulk of the frequency representation falls within the bands most obviously for the CF filters.

FIGURE D.1: Filtered Cyclical Series



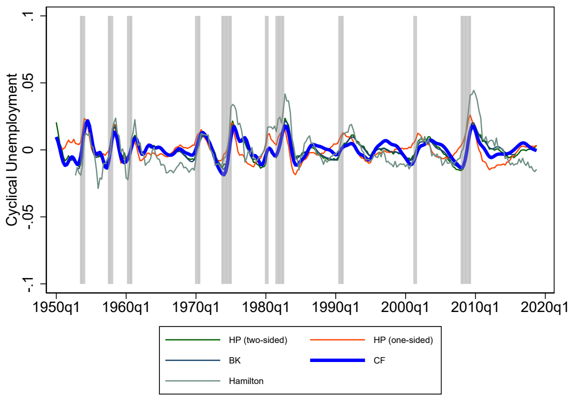
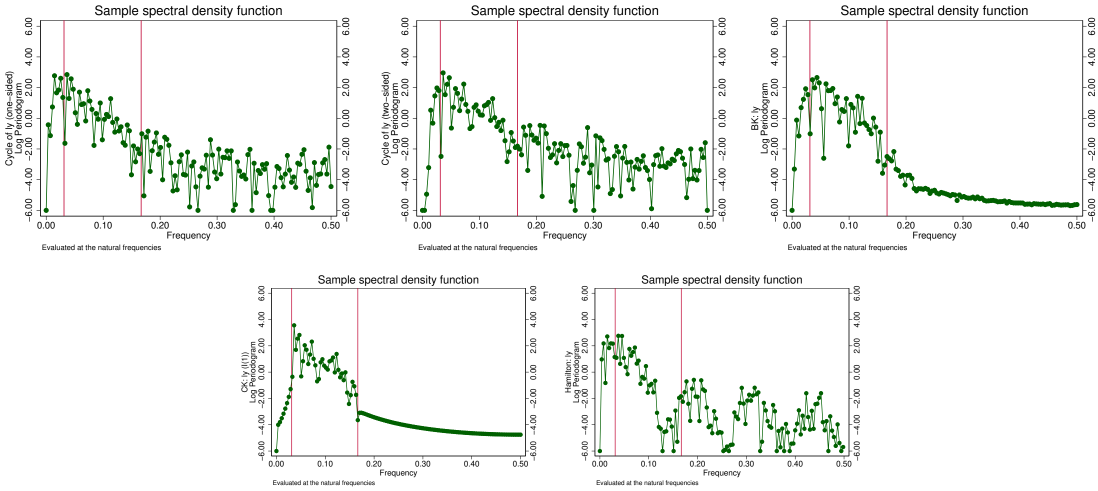
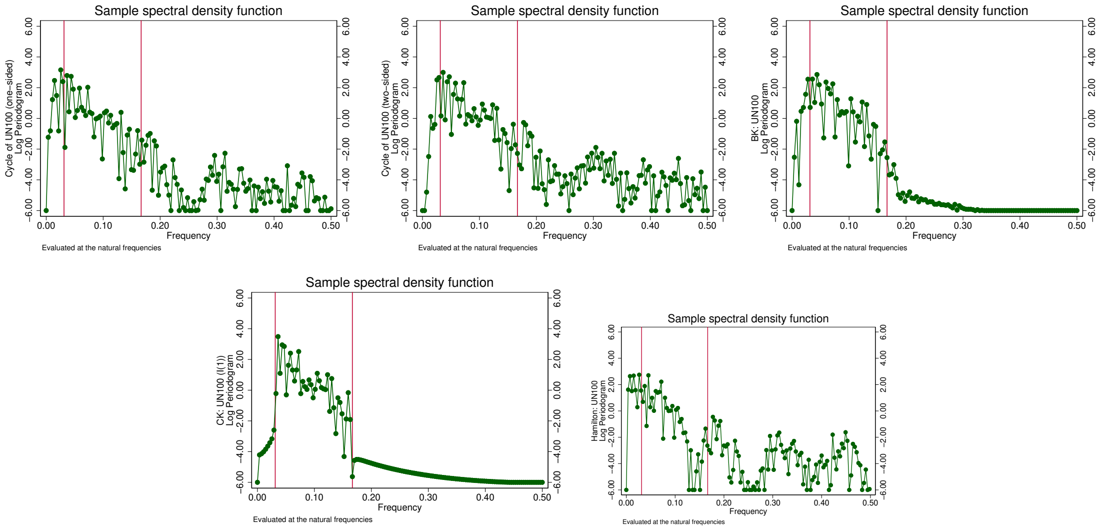


Figure D.2: Spectral Density of Filtered Output Series



Note. Series are ordered as HP 1 sided then 2-sided filter, Baxter-King Filter, Christiano-Fitzgerald, and Hamilton filter.

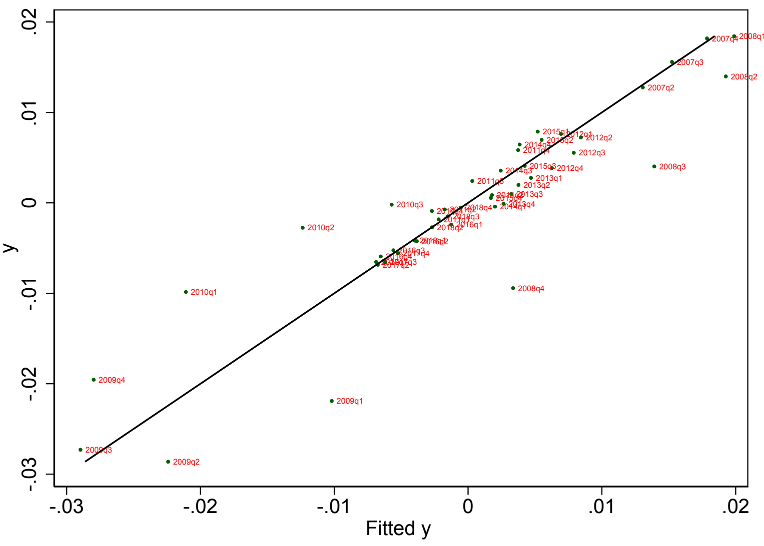
Figure D.3: Spectral Density of Filtered Unemployment Series



Note. Series are ordered as HP 1 sided then 2-sided filter, Baxter-King Filter, Christiano-Fitzgerald, and Hamilton filter. By "natural frequencies", we mean the standard frequencies divided by 2π. The vertical lines are consistent with the business-cycle frequencies of 6 and 32 quarters. Thus, $32^{-1} = 0.031$ and $6^{-1} = 0.167$.

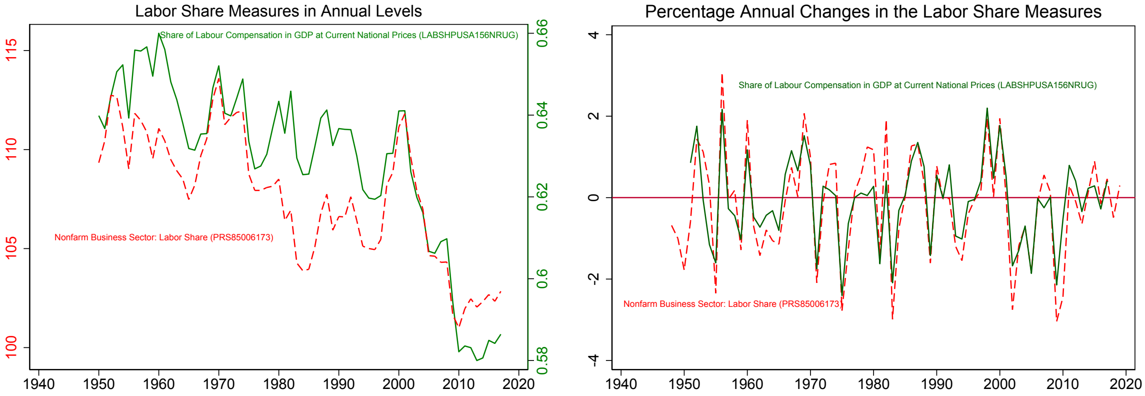
D.2 Miscellaneous Additional Material

Figure D.4: Actual vs. Fitted (2007-2018)



Note. This figure takes the full sample implied fit of the dynamic Okun's law and scatter plots these predictions with realized out turns from 2007q1-2018q4. The predictions are mostly below the actual values suggesting that the Okun coefficient is lower (more negative) than would be required over this period.

FIGURE D.5: Comparisons of FRED Labor Income Share Series



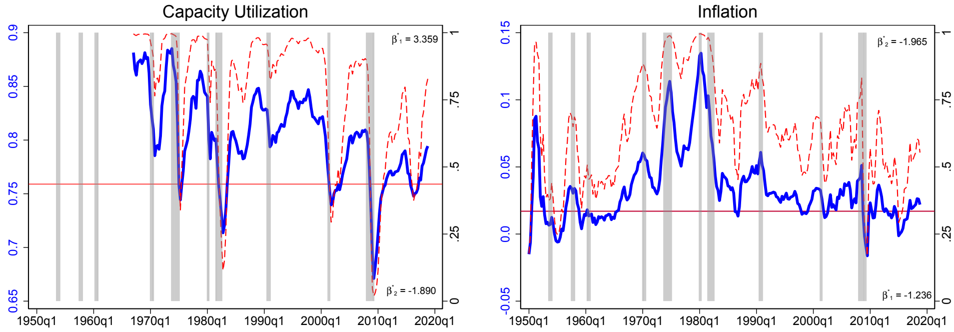
Note. These figures compare the level and percentage change characteristics of two measures of the US labor income share.

TABLE D.1: Threshold Okun Law Results: Ignoring Endogeneity

para./z_t	$labsh^{\dagger}$	u_l	π	$i_{s,xw}$	i_s	cu	i_l
ρ	0.541^{***} (0.080)	0.559^{***} (0.073)	0.586^{***} (0.075)	0.686^{***} (0.070)	0.687^{***} (0.070)	0.865^{***} (0.085)	0.766^{***} (0.073)
β	-0.509^{***} (0.138)	-0.967^{***} (0.195)	-0.623^{***} (0.153)	-0.273^{**} (0.138)	-0.272^{**} (0.138)	-0.183 (0.163)	0.023 (0.184)
b	-0.346^{***} (0.127)	0.861^{***} (0.259)	-0.171 (0.125)	-0.319^{***} (0.120)	-0.318^{***} (0.114)	0.123 (0.156)	-0.943^{**} (0.444)
δ	107.642 {107.610-107.645}	0.0279 {0.0268-0.0288}	0.0404 {0.0106-0.060}	4.100 {4.02-4.150}	4.100 {4.04-4.140}	0.835 {0.807-0.860}	9.420 {9.110-9.652}
γ	5.500 {4.10-6.90}	0.100 {0.010-0.4}	732.500 {731.099-733.500}	6.600 {4.90-8.00}	8.900 {7.100-10.800}	461.400 {460.00-463.00}	0.800 {0.091-0.097}
β_1^*	-1.109	-2.193	-1.505	-0.869	-0.869	-1.356	0.098
β_2^*	-1.863	-0.240	-1.918	-1.885	-1.885	-0.444	-3.932
AIC	-1028.876	-1030.348	-1026.544	-975.474	-975.510	-802.645	-912.298
$\mathcal{D}_2$	70.100	10.000	27.540	53.125	53.750	25.490	10.000

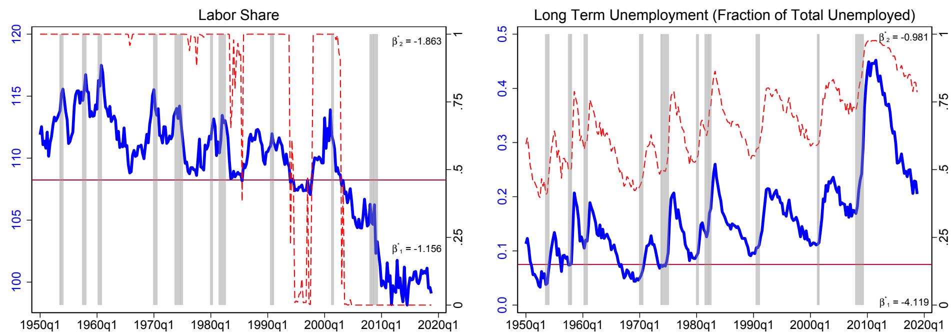
Note. This table shows the results from estimation of (18) for the threshold candidates that rejected a linear specification. We show the case where there is no endogeneity modelled, $\omega_1 = \omega_2 = 0$. Numbers in parentheses below the coefficients represent bootstrapped standard errors, and those in braces represent bootstrapped 95% confidence intervals. The terms β_1^* and β_2^* are as defined in (19). $\mathcal{D}_2$ is the sample percentage residing in the second regime. † The labor share measure used is the quarterly indexed one provided by FRED where $2012 = 100$.

FIGURE D.6: Endogenous Threshold Variables and Transition Probabilities with NBER reference dates (Demand)



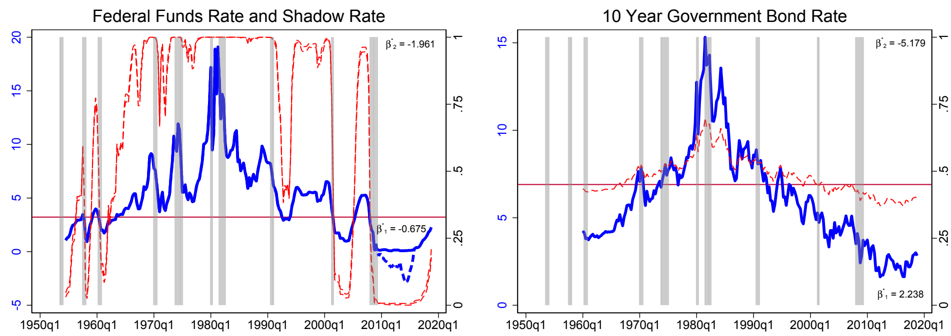
Note. Blue solid line represent the threshold variable, z, red dashed line represent the $\{0,1\}$ transition probabilities. The horizontal solid line represents the estimated threshold parameter, $\hat{\delta}$. The β^* values on the rhs edge are the Okun parameters in the respective regimes.

FIGURE D.7: Endogenous Threshold Variables and Transition Probabilities with NBER reference dates (Structural)



Note. See notes to Figure D.6.

FIGURE D.8: Endogenous Threshold Variables and Transition Probabilities with NBER reference dates (Policy/Financial)



Note. See also notes to Figure D.6.

References

Schüler, Yves S. (2018) On the cyclical properties of Hamilton's regression filter, Discussion Papers 03/2018, Deutsche Bundesbank.

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2020

01-20 Incubated Early Stage Startuppers in Athens and their Initiatives during the Crisis (2010-2016), Ioannis Besis and Ioanna Sapfo Pepelasis
02-20 Alternative Monetary Policy Rules in an Imperfectly Competitive Dynamic Stochastic General Equilibrium Model, George Alogoskoufis and Stelios Giannoulakis
03-20 Quantifying Qualitative Survey Data: New Insights on the (Ir)Rationality of Firms' Forecasts, Alexandros Botsis, Christoph Görtz and Plutarchos Sakellaris
04-20 Can Country-Specific Interest Rate Factors Explain the Forward Premium Anomaly?, Efthymios Argyropoulos, Nikolaos Elias, Dimitris Smyrnakis and Elias Tzavalis
05-20 Signaling product quality through prices in an oligopoly model with costly consumer search, Minghua Cheny, Konstantinos Serfes and Eleftherios Zacharias
06-20 Thinking ahead of the next big crash: Clues from Athens in classical times, George C. Bitros
07-20 Financial crises, firm-level shocks, and large downturns: Evidence from Greece, Stelios Giannoulakis and Plutarchos Sakellaris
08-20 Notes on the Demand Side Secular Stagnation, George D. Demopoulos and Nicholas A. Yannacopoulos
09-20 Institutions and Macroeconomic Performance: Core vs Periphery Countries in the Eurozone, Tryfonas Christou, Apostolis Philippopoulos and Vanghelis Vassilatos
10-20 Macroeconomic Policy Lessons for Greece, George Economides, Dimitris Papageorgiou and Apostolis Philippopoulos
11-20 Energy Transitions and the role of the EU ETS: The case of Greece, Andriana Vlachou and Georgios Pantelias
12-20 Measuring the Default Risk of Small Business Loans: Improved Credit Risk Prediction using Deep Learning, Yiannis Dendramis, Elias Tzavalis and Aikaterini Cheimarioti
14-20 Saving Democracy by Switching to Corporate-like Governance, George C. Bitros
15-20 The impact of the lockdown on the Greek economy and the role of the Recovery Fund, George Economides, Apostolis Philippopoulos and Vanghelis Vassilatos

2021

01-21 Historical Cycles of the Economy of Modern Greece From 1821 to the Present, George Alogoskoufis
02-21 Greece Before and After the Euro: Macroeconomics, Politics and the Quest for Reforms, George Alogoskoufis
03-21 Commodity money and the price level, George C. Bitros
04-21 Destabilizing asymmetries in central banking: With some enlightenment from money in classical Athens, George C. Bitros
05-21 Exploring the Long-Term Impact of Maximum Markup Deregulation, Athanasios Dimas and Christos Genakos

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

06-21 A regularization approach for estimation and variable selection in high dimensional regression models, Y. Dendramis, L. Giraitis, G. Kapetanios

07-21 Tax Competition in the Presence of Environmental Spillovers, Fabio Antoniou, Panos Hatzipanayotou, Michael S. Michael, Nikos Tsakiris

08-21 Firm Dynamics by Age and Size Classes and the Choice of Size Measure, Stelios Giannoulakis and Plutarchos Sakellaris

09-21 Measuring the Systemic Importance of Banks, Georgios Moratis, Plutarchos Sakellaris

10-21 Firms' Financing Dynamics Around Lumpy Capacity Adjustments, Christoph Görtz, Plutarchos Sakellaris, John D. Tsoukalas

11-21 On the provision of excludable public goods General taxes or user prices? George Economides and Apostolis Philippopoulos

12-21 Asymmetries of Financial Openness in an Optimal Growth Model, George Alogoskoufis

13-21 Evaluating the impact of labour market reforms in Greece during 2010-2018 , Georgios Gatopoulos, Alexandros Louka, Ioannis Polycarpou, Nikolaos Vettas

14-21 From the Athenian silver to the bitcoin standard: Private money in a state-enforced free banking model, George C. Bitros

15-21 Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities. Stelios Arvanitis and Thierry Post

16-21 Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka

17-21 Competition and Pass-Through: Evidence from Isolated Markets, Christos Genakos and Mario Pagliero

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5