What Goes up Must Come Down: Theory and Model Specification of Threshold Dynamics*

Hannah Paul†
Andrew Q. Philips‡

June 20, 2022
Forthcoming at *Social Science Quarterly*

Abstract

Objectives: Despite the frequent use of time series models in the social sciences, they have often remained within the confines of assuming purely linear dynamic effects. We contend that many theories involve relationships that are inherently non-linear. *Methods:* We discuss several approaches to modeling a variety of these types of non-linear autoregressive data-generating processes, specifically threshold effects. *Results:* We replicate and extend a recent analysis, and show evidence of threshold processes. *Conclusion:* In doing so, we show that threshold models allow us to test richer, more complex theoretical implications about dynamic effects.
Word Count: 93 (abstract); 7932 (manuscript)
Keywords: time series; threshold modeling; asymmetric effects; dynamic models

*We would like to thank Neil Beck, Patrick Brandt, Matt Lebo, Mark Pickup, Vera Troeger, Matt Yarbrough, Joe Zamadics, the anonymous reviewers and the editors for their helpful comments and suggestions. Any remaining errors and omissions are those of the authors.

†hannah.paul@colorado.edu. Assistant Professor, Truman School of Government and Public Affairs, University of Missouri, 101 Middlebush Hall, Columbia, MO 65211

‡andrew.philips@colorado.edu. Associate Professor, Department of Political Science, University of Colorado Boulder, UCB 333, Boulder, CO 80309.

Many phenomena in political science are inherently dynamic. From public opinion to political economy, a characteristic of time series data is that current values are often related to prior ones. While early approaches often centered on modeling out this temporal dependence through the error process, a substantial body of recent work in applied time series analysis aims to model these dynamic processes as something of theoretical interest.

Although scholars have adapted complex time series models for applied research, they have often remained within the confines of linear (i.e., constant) dynamic effects. That is to say, it is almost always assumed that dynamics in the dependent variable (and sometimes in the independent variable) have a constant effect, no matter the level—or direction of change—of the series. While these linear models are easy to interpret, they are not always reflective of our theories, which often focus on asymmetric or non-linear effects. This is true for many different bodies of literature in the social sciences. For instance, economic voting only occurs in developing countries once a threshold of political development has been met (Wimpy and Whitten 2017). In the literature on public policy, governments may react differently to an economic boom versus a recession (Lipsmeyer 2011), and government response to public opinion may not always be characterized by small, stable shifts (Baumgartner, Green-Pedersen and Jones 2006; Jennings and John 2009; Baumgartner and Jones 2010). In the public opinion literature, support for a political system changes depending on whether or not population size is above or below particular thresholds (Matsubayashi 2007), while shifts in opinion and voting behavior are thought to be larger in response to negative economic news (or actual economic performance) compared to good news (Soroka 2006; Kappe 2018; Park 2019). In the urban studies literature, after a certain threshold in disadvantages to a neighborhood is met, diminishing political participation rates occur (Levine et al. 2018), while a neighborhood's urban, rural and suburban setting affects its level of resource deprivation in different ways depending on whether or not its above or below some threshold of poverty (Murphy and Wallace 2010). All of these theories, while diverse in nature, have at their core a notion that effects are not always constant.

In this paper we discuss a variety of strategies for testing and modeling a specific type of non-linear relationship in time series, that of threshold effects. Threshold models have existed for decades in fields such as economics (Tong and Lim 1980; Hansen 2011; Tong 2012). However, with few exceptions they have been underutilized in other social science fields.[1] We find this puzzling for two reasons. First, thresholds often form the core of theories in the social sciences; thresholds exist anytime we expect that effects may change only after a certain point has been surpassed. For instance, women's representation in politics may only play a meaningful role after a certain 'critical mass' has been reached (Funk, Paul and Philips 2021), or government ideology may only lead to policy change when the economy is performing strongly (Lipsmeyer 2011). Second, other non-linear approaches, such as Markov-switching (Brandt, Freeman and Schrodt 2011) or even asymmetries (e.g., Soroka 2006; Philips, Rutherford and Whitten 2015), have frequently been used in the literature. We compliment these studies by showing an additional type of dynamic structure to model in the form of threshold effects. Although threshold models can be used in static applications (c.f., Kappe 2018), as we show below, this approach is well-suited to studies of dynamic processes. By incorporating these tools into their modeling strategy, scholars can better reflect the underlying theory they seek to understand.

The rest of this paper proceeds as follows. We begin with a discussion of threshold effects and discuss how to test for and model this characteristic, allowing for threshold effects in stationary autoregressive distributed lag (ARDL) models. We demonstrate this strategy and how to link it to theory with an applied time series example using data from Soroka, Stecula and Wlezien (2015). We provide guidance as to how researchers can choose the appropriate modeling approach for their theories, and conclude with recommendations for the application of threshold models to future research.

[1] Exceptions include Kappe (2018) and Zhang et al. (2021).

Modeling Threshold Effects

Approaches to accounting for non-constant effects between covariates and the dependent variable remain disjointed, if they are used at all.[2] Interactive effects are commonplace in dynamic theories in the social sciences, whereby one or more variables condition the effect of another. While these interactions moderate a variable's effect, they often remain linear in how they are specified—i.e., changes in the marginal effects remain constant.[3] Generally speaking, non-linear effects have been modeled by transforming the data to ensure linearity in the parameters, such as by employing a squared term or taking the log of a variable. However, this strategy has limited use if the non-linearity cannot be modeled by changing the functional form through common transformations.

Other articles model asymmetries based on expectations about non-constant effects on either side of a cut-point, which we refer to as structural break parameterization (e.g., Soroka 2006; Lipsmeyer 2011; Philips, Rutherford and Whitten 2015). While technically another form of interaction, structural break parameterization is better seen as conditioning a variable's effect by whether its own values are above or below a certain point. These cut-points are typically specified based on theory, although they may be supported by empirical testing such as Chow tests (e.g., Clarke, Ho and Stewart 2000).[4] While such approaches to modeling a theoretical threshold are useful, they remain bound to the hypothesized threshold value. Sometimes the researcher might not know exactly where the cut point is, and specifying a particular cut point in this case could lead to incorrect inferences. In other situations, the location of the cut point itself might be a substantively interesting hypothesis to test on its own.

A different body of literature treats shifts between regimes on either side of a cut-point as unobservable, often modeling them as a hidden Markov process (Hamilton 1989).[5] For

[2] This is not for the absence of non-linear modeling techniques, which are frequently employed by econometricians (c.f., Granger, Terasvirta et al. 1993), but their use outside this field is much less common.

[3] Of course, models with non-linear marginal effects are possible (Berry, Golder and Milton 2012, pp. 669-671).

[4] See Soroka (2006) and Lipsmeyer (2011) for examples.

[5] For a review of the Markov modeling approach to modeling non-linear dynamics, see (Quandt 1972; Goldfeld and Quandt 1973; Hamilton 1989, 1993, 1995).

instance, there may be high and low conflict regimes in intra-state conflict, for which dynamic processes may differ across the two regimes (Brandt, Freeman and Schrodt 2011). Or, the use of force by US presidents may have undergone a structural change before and after the Second World War (Park 2010). Non-linearity might characterize the regime transitions themselves, requiring a 'multistate survival model' that allows for both recursive and sequential regime changes for a single country (Metzger and Jones 2016). Freeman, Hays and Stix (2000) and Hays, Freeman and Nesseth (2003) use Markov switching models to predict transition probabilities between currency market equilibria as a function of different types of political institutions and information; both articles show that how politics shapes currency market equilibration depends on how democratic a country is and how transparent its policymaking is.

We focus on threshold models, which let the cut-point, or "threshold" value, remain unspecified (Gonzalo and Pitarakis 2002; Hansen 2011; Tong 2012). Instead, the data identify the most likely threshold value, typically by minimizing the sum of the squared residuals (Hansen 2000). This is beneficial to scholars for two reasons. First, it allows us to better map theories about non-linear relationships onto our statistical models. Second, this approach allows us not only to find the threshold value, but also to form a statistical test which, as we show below, can be used to support a theoretical hypothesis. Outside of economics, the use of threshold models is rare. Even those who do employ threshold models (for instance, Kappe 2018; Zhang et al. 2021), do not examine them in the context of dynamic models.[6] For brevity, we discuss threshold models below, leaving a more in-depth discussion of asymmetries with fixed cut-points (i.e., the literature on structural break parameterization mentioned above) for interested readers in the Supplemental Information (SI). In the conclusion (see Table 2), we also provide further guidance as to how researchers can think about mapping theories about thresholds onto other types of model specifications.

[6] Kappe (2018) uses a threshold approach on the independent variable to test prospect theory in the context of government popularity in the UK, but does not employ an autoregressive model, which is the focus of our paper. Zhang et al. (2021) use a fixed-effects panel threshold model, although similarly have no dynamics in their models (or thresholds).

Consider an autoregressive distributed lag series, y_t, that is a function of a constant, its own lag, y_{t-1}, and a white-noise error process. A threshold data-generating process (DGP) takes the following form:

$$y_t = \begin{cases} \beta_0^* + \phi^* y_{t-1} + \varepsilon_t, & \text{if } y_{t-1} \leq \omega \\ \beta_0 + \phi y_{t-1} + \varepsilon_t, & \text{if } y_{t-1} > \omega \end{cases} \tag{1}$$

In Equation 1, ω is some unknown threshold value, and $|\phi, \phi^*| < 1$, thus satisfying the stationarity condition. This is known as a "self-exciting" threshold autoregressive model (SETAR), since the threshold comes from lagged values of the dependent variable (Tong and Lim 1980; Tong 2012). In other words, if past values of y_t are above the threshold point ω, the autoregressive process in y_t can be completely different from that where past values of y_t are at or below the threshold point.[7]

A stylized example of a SETAR process is shown in Figure 1. The autoregressive series shown at the top of Figure 1 is a linear AR(1) process where $\phi = 0.75$. Note that this process is analogous to a SETAR DGP with a single regime for all y_t. In the two-regime SETAR shown below it, the series is autoregressive where $\phi = 0.75$ when $y_t \leq 0.5$, above which the series is autoregressive where $\phi = 0.10$. It is clear that this series looks quite different from the constant-autoregressive process, especially when the series is above the threshold. Note also how few observations lie above the threshold, a consequence of the large decrease in autoregression when series is above the threshold. The bottom plot in Figure 1 shows a three-regime SETAR DGP, where $\phi = 0.75$ if y_t lies between -1 and 1, and $\phi = 0.10$ if the series exceeds these bounds. Unlike the other two series, this one lies mostly within the two thresholds; when y_t falls outside of the thresholds, it quickly tends to move back into the middle regime.

In addition to generalizing up to any number of threshold regimes, there are a few smaller points worth discussing about threshold models.[8] First, although in Equation

[7] While non-self-exciting models are possible (i.e., the data-generating process of y_t changes when some z_t crosses some threshold, they are uncommon since they should be included in the empirical model.

[8] Applied work deals mostly with two or three regimes due to computational difficulty, although in theory there could be any number of regimes.

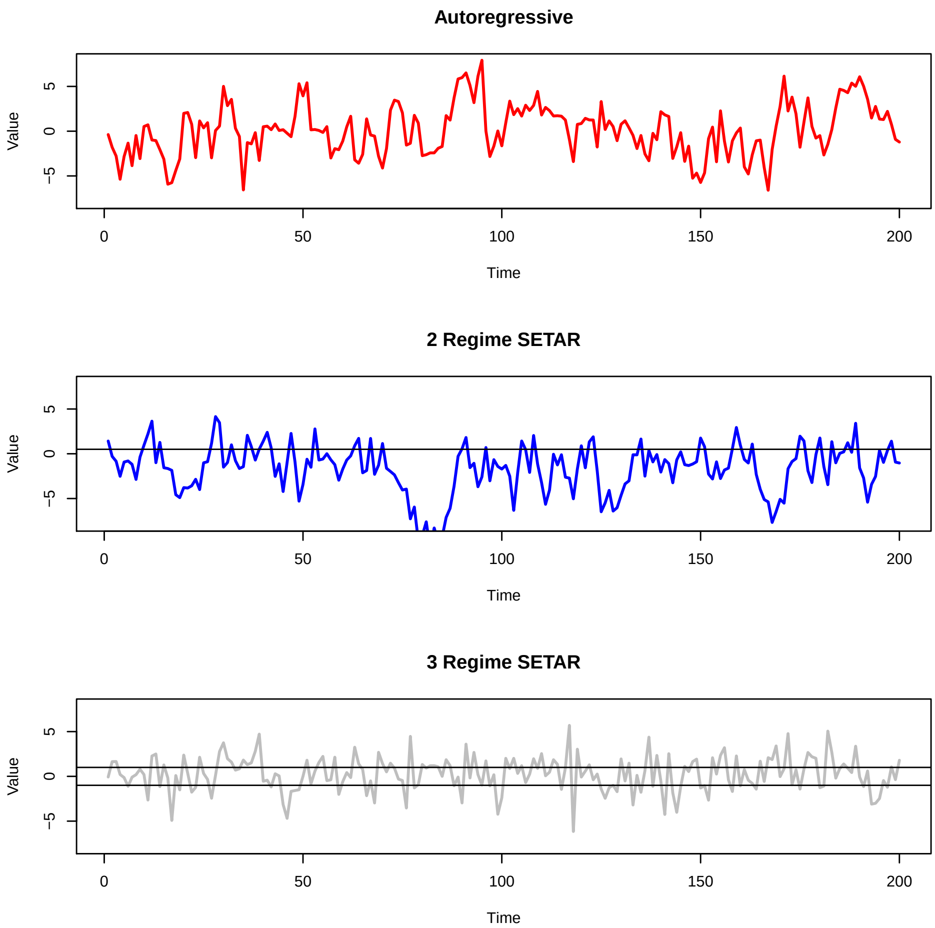

Figure 1: Autoregressive, 2- and 3-Regime SETAR Series

Note: Autoregressive series' DGP is $y_t = 0.75y_{t-1} + \varepsilon_t$ (where $\varepsilon_t \sim N(0,2)$, 2-regime SETAR DGP is $y_t = 0.75y_{t-1} \cdot D_t + 0.10y_{t-1} \cdot (1 - D_t) + \varepsilon_t$ (where D_t is a dichotomous indicator function equal to one if $y_{t-1} \leq 0.5$), and the 3-regime SETAR DGP is $y_t = 0.75y_{t-1} \cdot D_t + 0.10y_{t-1} \cdot (1 - D_t) + \varepsilon_t$ (where $D_t = 1$ if $|y_{t-1}| \leq 1$. Thresholds shown by the horizontal lines.

3 there is a one-period lag on the threshold, which determines the regime the data-generating process is in, there could be a delay of up to d time points (i.e., if $y_{t-d} > \omega$) (c.f., Balke and Fomby 1997). While $d = 1$ is common, analysts should always use theory and information criterion (e.g., AIC/BIC) to test whether the threshold appears at further lags. Second, the data-generating process of the three-regime SETAR model shown in Figure 1 is equivalent (in terms of its effects) on either side of the threshold, in what is known as a band-TAR (Balke and Fomby 1997). Asymmetric thresholds are also possible, such that the autoregression is larger or smaller on the upper and lower thresholds. Third, a variety of other models exist that allow for a richer set of dynamic effects to take place on either side of the threshold, such as the smooth transition autoregressive (STAR) (Terasvirta and Anderson 1992), exponential or logistic STAR (ESTAR and LSTAR) (Tong 2012). Below, we show the most commonly used models and discuss estimation strategy using an applied example. We apply the TAR and Band-TAR, which we believe are more directly applicable to social science theories. For brevity, and because models such as the E-STAR and L-STAR are less common, we discuss them more fully in the Supplemental Information.

Applied Example: Asymmetries in Media Tone

Below we use an applied example to show readers how to estimate a TAR model. After presenting the hypotheses and results of this example, we discuss the nuts and bolts of TAR estimation and related diagnostic testing. Our applied example comes from Soroka, Stecula and Wlezien (2015, henceforth SSW), who test whether media and the public respond to past, present, or future expectations about the economy. Using monthly data for the US from 1980 to 2011, they conclude that media coverage tends to focus on future expectations about the economy, and find evidence for an endogenous relationship between the media and public opinion.

Novel Hypotheses Using the TAR Approach

We were able to replicate the results of SSW exactly. Focusing specifically on Model 6 in Table 3 in their article, they use an error-correction model to examine how media respond (in changes in media tone) to past (which they refer to as "lagged"), contemporaneous ("coincident") and future ("leading") predictions about the economy. Their results suggest that leading-indicators of the economy are the largest drivers of changes in media tone.

However, there may be reasons to suspect a threshold relationship between the economy and media tone. Their key dependent variable, media tone—whether the media is net positive or negative—has an intuitive threshold around zero; when tone is positive, changes in the economy might lead to different effects than when it is negative. That is, the nature of previous media tone can lead to differing effects in the relationship under examination. If predictions about the economy are made when media tone was negative in the previous month, then we expect that the magnitude of the effect on current media tone will be different compared to when media tone in the previous month was positive. By testing for a threshold at previous media tone = 0, we are able to evaluate our theory that the quality of previous media tone – net positive or negative – is important for explaining how future predictions about the economy shape current media tone. Thus, we expect that:

H1: *The threshold for media tone is zero.*

As we will show below, threshold autoregressive models allow us to evaluate a statistical test of **H1**.

While a threshold at zero is perhaps the most intuitive theoretical location for many applications of threshold models—since it suggests that effects may differ based off whether lagged values of a variable are positive versus negative—it is not the only location where we might expect a threshold to lie. For instance, in sociological theories of collective behavior, Granovetter (1978, p. 1423) notes that women in South Korean villages, "may be wary of adopting birth control devices and wait to do so until some proportion of their fellow villagers do." In the literature on women's representation, it is thought that increasing the percentage of women's representation (in the legislature, on a company board, etc.) will have little effect on women's descriptive representation until a certain 'critical mass' is reached (c.f., Funk, Paul and Philips 2021). Bruch and Mare (2006, p. 672) theorize that the large amount of neighborhood racial segregation in the US is caused by individual-level choices that follow a threshold function; "the majority of whites will not tolerate neighborhoods that are more than 20% black. In contrast, most blacks prefer a neighborhood that is at least 50% black." Last, Choi (1999) finds that interest rate and inflationary responses to shocks to the money supply differ based on whether monetary policy is in one of three regimes (thus, two thresholds): tight, neutral and loose. All of these examples across various fields hypothesize a non-zero threshold.

SSW show that media coverage responds to indicators about the future economy. While this finding shows us the type of economic indicators that media are more responsive to, we are left wondering if media coverage *always* responds in this way to expectations about the economy. We expect that the quality of previous media tone—positive or negative—will shape how current media coverage responds to future economic indicators because a previous negative news cycle sets a very different stage for the next news cycle compared to a positive one. In terms of magnitudes of the effects, we anticipate that when previous media tone is negative, future economic indicators will have a greater effect on current media tone compared to when past media tone is positive. Our reasoning is as follows: if previous media coverage was negative, then the effect of current indicators about the future economy are likely to have an outsized effect on the tone of current media coverage compared to if media tone were previously positive, because negative news coverage is likely to make media more sensitive to information about the future. Negative media tone is likely to be associated with some level of anxiety about any aspect of society, such as the economy, and information about how the economy is going to be is just the type of information that anxiety craves. This leads to an additional testable hypothesis that we can leverage using our threshold approach:

H2: *The effect of leading economic indicators on changes in media tone will be greater in magnitude when lagged media tone is below a threshold of zero.*

To emphasize, standard dynamic models allow us to test only a limited subset of potential hypotheses. For Soroka, Stecula and Wlezien (2015), this meant examining whether short- and/or long run economic indicators mattered for media tone. Our approach of incorporating threshold autoregressive models allows us to test a richer set of additional hypotheses that we cannot obtain using standard modeling approaches. In our case, this is the existence of a threshold in media tone (and an expectation that it is near-zero), as well as that leading economic indicators will only matter when lagged media tone is negative.

TAR Modeling and Unit Root Testing

To examine whether a threshold model is appropriate, we first replicated Soroka, Stecula and Wlezien (2015)'s model described above.[9] The linear autoregressive distributed lag model—or ARDL, meaning one lag of the dependent variable and each independent variable appears both at time t as well as its one-period lag—is shown in Table 1, Model 1. The coefficient on the lag of media tone is positive and statistically significant, and suggests that about 38 percent of current levels of media tone are directly attributable to previous levels of media tone. None of the economic indicators are statistically significant, with the exception of leading economic indicators (Leading EI_t, Leading EI_{t-1}). Improving leading economic indicators appears to increase media tone at time t, but much of that positive increase completely disappears by time $t+1$, as evidenced by the coefficient on Leading EI_{t-1}; in fact, the calculated long-run effect of a one-unit increase in Leading EI is only about 0.027 (although this effect is statistically significant). Thus, improvements in leading economic indicators have a short-run—but very little long-run—positive impact on media tone. To reiterate, these are the same as SSW's findings.

Like SSW, we conclude that all series are stationary.[10] However, since we believe

[9]As the authors mention in an erratum (Soroka, Stecula and Wlezien 2016), all series are stationary. For ease of presenting our threshold approach, we therefore estimated an autoregressive distributed lag model instead of an error-correction model, although it should be noted that the results from the all-stationary ARDL model are identical to the all-stationary error-correction model used by the authors (Philips 2021).

[10]Covariance stationarity implies that a series has a constant mean, variance, and covariance (Enders

there to be a threshold in media tone, such threshold series complicate testing. Standard unit root tests—which are needed to ensure that stationarity conditions have not been violated—have been shown to have low power if a threshold process is present (Balke and Fomby 1997; Enders 2010). Instead, we use the testing strategy proposed by Enders and Granger (1998). This involves testing the null hypothesis that media tone contains a unit root against the alternative that the series is stationary *yet threshold autoregressive*.

To test for a threshold unit root, we first regress Tone_t on a constant, and save the residuals, thus creating a de-meaned series we call Tone^*_t. Next, estimate:

$$\Delta\text{Tone}^*_t = D_t\rho_1(\text{Tone}^*_{t-1}) + (1 - D_t)\rho_2(\text{Tone}^*_{t-1}) + \varepsilon_t \tag{2}$$

where D_t is the following dichotomous indicator function:

$$D_t = \begin{cases} 1, & \text{if } \text{Tone}_{t-1} \leq \omega \\ 0, & \text{if } \text{Tone}_{t-1} > \omega \end{cases} \tag{3}$$

Since ω is not identified, we use the approach suggested by Chan (1993)—discussed in detail in next section—of estimating regressions along all possible values of ω, and choosing the value that minimizes the sum of the squared residuals. We find that $\omega = -0.44$, and estimate the two autoregressive parameters in Equation 2 as $\hat{\rho}_1 = -.29$ and $\hat{\rho}_2 = -.63$. We then test the null hypothesis that the two coefficients are jointly equal to zero; $\rho_1 = \rho_2 = 0$ (i.e., the series contains a unit root) using an F-test. Using the non-standard critical values provided by Enders and Granger (e.g., 1998, p. 306), we easily reject the null hypothesis at conventional levels.[11] The results from the threshold unit root test in Equation 2 suggests that the series is stationary and, importantly for our theoretical claims, may contain a threshold autoregressive process.[12]

2010). Non-stationary series tend to have extremely high Type I error rates (i.e., spurious conclusions of relationships).

[11]The F-statistic is 71.43 compared against critical values (assuming $T \approx 250$) of 4.56 (95%) and 6.47 (99%).

[12]It is also advisable to test whether a non-linear data-generating process is possible. Both the McLeod and Li and RESET tests—which for brevity we describe in the SI—provide strong evidence of a threshold effect.

Table 1: Threshold Effects of the Economy on Media Tone

	(1) Standard ARDL	(2) Threshold TAR		(3) Band-TAR	
		$\text{Tone}_{t-1} \leq \omega$	$\text{Tone}_{t-1} > \omega$	Tone_{t-1} in band	Tone_{t-1} out of band
Tone_{t-1}	0.380***	0.719***	0.278***	0.336***	0.563***
	(0.047)	(0.155)	(0.072)	(0.124)	(0.101)
Lagging EI_t	0.0442	0.127	0.011	0.017	0.044
	(0.056)	(0.136)	(0.060)	(0.058)	(0.110)
Lagging EI_{t-1}	-0.034	-0.087	-0.014	-0.023	-0.015
	(0.051)	(0.129)	(0.055)	(0.053)	(0.106)
Coincident EI_t	0.021	-0.112	0.069	0.067	-0.033
	(0.078)	(0.171)	(0.086)	(0.125)	(0.127)
Coincident EI_{t-1}	-0.073	-0.041	-0.079	-0.045	-0.094
	(0.082)	(0.174)	(0.091)	(0.123)	(0.132)
Leading EI_t	0.124***	0.062	0.144***	0.201***	0.027
	(0.034)	(0.066)	(0.039)	(0.046)	(0.044)
Leading EI_{t-1}	-0.107**	-0.014	-0.137***	-0.203***	0.013
	(0.034)	(0.067)	(0.039)	(0.047)	(0.044)
Constant	0.139***	0.255***	0.174***	0.157***	-0.007
	(0.018)	(0.051)	(0.029)	(0.037)	(0.070)
Threshold (ω)	*NA*	-0.026		-0.026 and 0.519	
N	383	383		383	
SBIC	-952.78	-930.12		-884.87	
AIC	-984.36	-993.29		-979.62	

Note: Standard errors in parentheses. Tone_{t-1} determines the threshold. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Estimating a Threshold Model

While estimating a threshold autoregressive model may be theoretically justifiable—and even if the above unit root tests support this—there are a number of difficulties in estimation that make it more difficult than standard linear regression. As advice for readers, we discuss three steps that users should take when estimating TAR models:

1. Test a theoretically-informed hypothesis about the location of the threshold
2. Test whether or not modeling a threshold autoregressive process improves on a linear model
3. Construct confidence intervals around the threshold

First, how do we find the threshold value to test a theoretically informed hypothesis about a threshold? While the threshold is a parameter, it cannot be simultaneously estimated with the coefficients using OLS. Instead, the threshold parameter is commonly chosen by minimizing the sum of the squared residuals using a method proposed by Chan (1993). In other words, to find the best value of the threshold, we order all values of the threshold variable from lowest to highest. In our example, we expect that lagged media tone determines the threshold in current media tone, so it is our threshold variable. Next, we pick a candidate threshold (i.e., the value of lagged media tone upon which to split the sample), $\tilde{\boldsymbol{\omega}}$, and estimate the following regression:

$$\text{Tone}_t = \mathbf{x}_t \boldsymbol{\beta}_1 D_t + \mathbf{x}_t \boldsymbol{\beta}_2 (1 - D_t) + \varepsilon_t \tag{4}$$

where $\mathbf{x}_t$ is a vector of all variables (including the lagged dependent variable), $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are vectors of coefficients on either side of candidate threshold value $\tilde{\boldsymbol{\omega}}$, determined by D_t, where:

$$D_t = \begin{cases} 1, & \text{if } \text{Tone}_{t-1} \leq \tilde{\boldsymbol{\omega}} \\ 0, & \text{if } \text{Tone}_{t-1} > \tilde{\boldsymbol{\omega}} \end{cases} \tag{5}$$

In other words, given our candidate threshold, we code D_t, then estimate the regression shown in Equation 4, saving the sum of the squared residuals (SSR). We then pick a new candidate threshold, and repeat the process.[13] The preferred model is the threshold value that minimizes SSR, which has been shown to be a "super-consistent" estimate of ω (Chan 1993). Threshold models can be easily estimated in programs such as Stata or EViews; we discuss software packages more in the SI.

A graphical depiction of this process is shown in Figure 2. The model SSR is shown on the vertical axis, while various choices of thresholds along the threshold variable, lagged media tone, are on the horizontal axis. As Enders (2010, p. 447) notes, if there is no evidence of a threshold, "there should be no clear relationship between the sum of the squared residuals and the potential thresholds. However, if there is a single threshold, there should be a single trough in the graph." In Figure 2, the value of the threshold that minimizes SSR is $\tilde{\omega} = \omega = -0.026$, which produces a minimum SSR of about 26.4. For comparison, the SSR in the linear model is 28.1.[14]

Given Figure 2, one might be tempted to search for additional or alternative threshold 'troughs'. We caution against reading too much into Figure 2 however; the SSR may vary simply due to aberrations in the data, not some underlying threshold relationship, and we risk overfitting the data by including more thresholds (although keep in mind we do not need to specify *where* the threshold is, only the number that we think exist). Instead, we stress that in order to avoid ad-hoc threshold searches, researchers should first create a theoretically-informed hypothesis about the general location where they think the threshold may lie. As we show below, our approach allows for a hypothesis test of whether a threshold actually exists in this region. And, we emphasize that unlike fully parameterized/interactive models where researchers specify the cutpoint (and thus may get that specification wrong), our approach does allow for somewhat of a data-driven solution, although we stress the importance of creating theoretical expectations as to

[13] To avoid degree of freedom problems, thresholds are commonly set so that they cannot occur in the upper- and lower-five percentiles of the distribution.

[14] Using plots such as Figure 2 is also a way to visually detect the presence of more than one threshold; such a data-generating process will be characterized by more than one localized trough that minimizes SSR (Enders 2010).

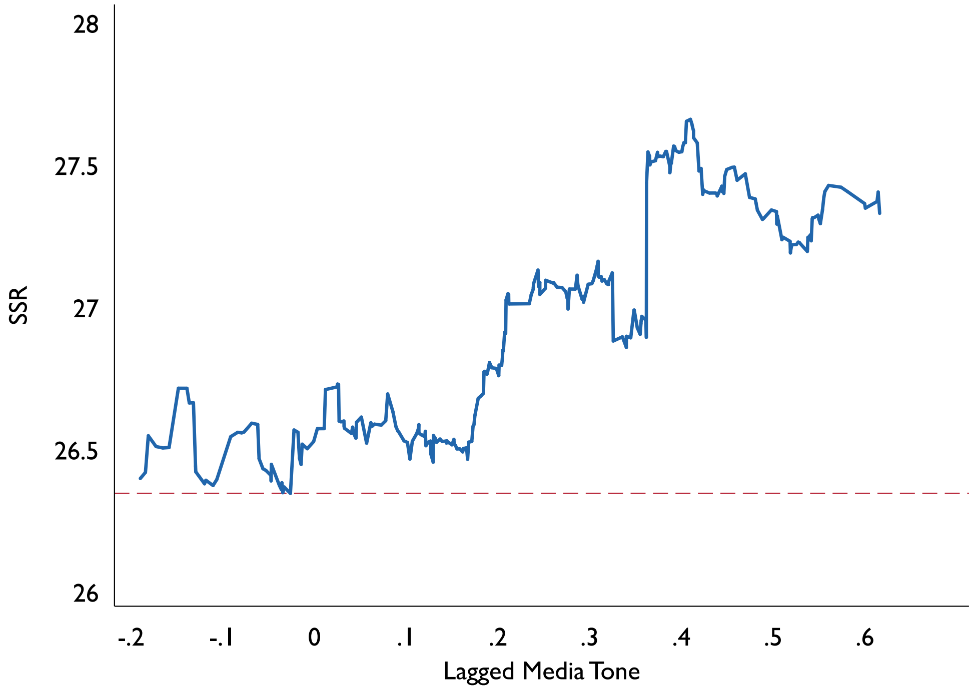

Figure 2: A Threshold of $\omega = -0.026$ Minimizes SSR

Note: Figure shows the sum of the squared residuals resulting from estimating Equation 4 across plausible candidate $\tilde{\omega}$ values, as shown on the horizontal axis.'

where the threshold lies.

The second step in estimating a TAR model is to test whether the threshold model is an improvement over the linear model. As shown in Figure 2, by construction, the selection procedure for choosing the threshold value will *always* find a threshold value that minimizes SSR, which leaves open the question of whether we need to estimate such a model in the first place rather than a linear model. Unfortunately, testing a threshold versus a linear model is not straightforward, since under the null hypothesis of no threshold effect, the threshold parameter ω is not identified. Instead, we suggest two approaches. First, Hansen (2000) recommends using a likelihood ratio test under the null hypothesis of no threshold. The distribution is non-standard, so special critical values from Hansen (2000) must be used, with p-values calculated using a bootstrap procedure. Using Hansen's approach, we find that the value of the likelihood ratio at $\omega = -0.026$ is 21.83, which exceeds the critical value of 20.22. Thus, by rejecting the null hypothesis, we conclude that the threshold model is warranted. Were we to instead fail to reject the null hypothesis, we may conclude that the purely linear model is adequate. Of course, failing to reject the null hypothesis of the Hansen test does not necessarily mean that researchers should throw out their threshold theory altogether. Rather, they should report the test result as a caveat to their theory and findings.

As a robustness check for a researcher's threshold theory, we recommend that they should evaluate whether such a threshold model actually constitutes an improvement over a linear model. One way is to use information criterion to select the preferred model, given potential finite-sample issues in these testing procedures and relying on bootstrapped p-values (Enders 2010). According to Table 1, we get somewhat mixed results regarding the preferred model; SBIC indicates that the standard ARDL approach is best, while AIC suggests that the threshold model is preferred. Overall, information criteria results—combined with the Hansen (2000) test as well as the portmanteau non-linearity tests shown in the SI—provide evidence that a threshold process exists in media tone.

Last, since the threshold can be thought of as a parameter, we may also want to construct confidence intervals around the estimate for hypothesis testing. For our example, can we reject the null hypothesis that the threshold is $H_0 : \omega = 0$? Hansen (2000) suggests using the 95th percentile of the asymptotic distribution of the likelihood ratio statistic. Any likelihood ratio value falling below the critical value can be thought of as falling within the 95 percent confidence interval. As shown in Figure 3, given a critical value of 7.35 (as indicated by the dashed horizontal line), the constructed confidence intervals around $\omega = -0.026$ are [-0.270, 0.206], which encompasses zero, as hypothesized. Thus, it appears media tone does have a threshold around zero, providing evidence that the data-generating process when tone is negative is different than when it is positive.

Now that we have shown that the threshold model is appropriate, we proceed to interpret the results in Table 1. The results in Model 2 suggest that very different processes are taking place in the two regimes; when media tone is below the threshold (i.e., negative), media tone appears to be very slow-moving, as evidenced by the positive and large (near-one) coefficient on the lagged dependent variable. Moreover, when media tone is

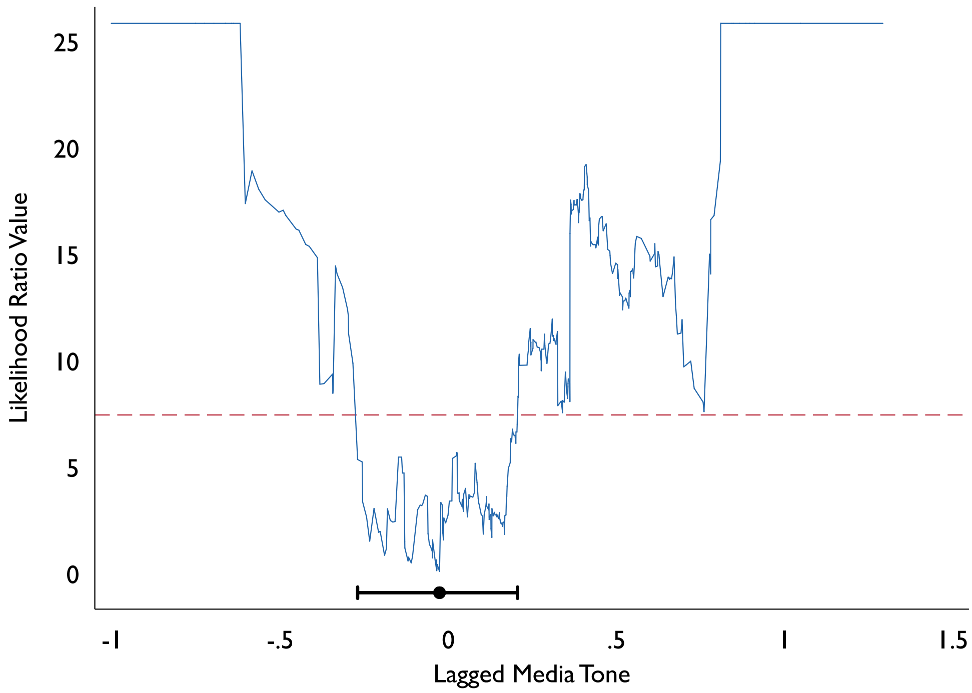

Figure 3: 95 Percent Confidence Interval Around the Threshold Encompasses 0

Note: Any likelihood ratio value falling below the 95% critical value (horizontal dashed line) comprises the estimated 95% confidence interval around ω. Values calculated from Table 1, Model 2.

negative, none of the economic indicators are statistically significantly different from zero. This contrasts sharply for when tone is above the threshold, in which the results are quite similar to the original findings. Substantively, this means that leading economic indicators only affect media tone when it is positive. When it is negative, leading economic indicators appear to have no effect on the tone of the media. Our threshold model shows us the somewhat counterintuitive effect that when net media tone is positive, rather than negative, expectations about the future economy are what shapes current media tone. While our theoretical intuition tells us that "bad" news drives anxiety around the future of the economy and thereby influences current media tone, the evidence shows that it's actually the "good" news that brings future economic expectations into focus when explaining current media tone.

Another plausible scenario is that the effect of leading economic indicators exist only around the mean of the series and not on the extreme ends. In other words, within a central part of the distribution of media tone, leading economic indicators may matter for media tone. Instead, when the media is either extremely positive or negative, economic indicators may have no effect, at least not until media tone has returned to the "normal" region. An approach we can use to test this utilizes a band-TAR. This model allows for one dynamic process to take place within a "band"—again determined by a threshold variable—and another to take place outside of this band.

The results from the band-TAR are shown in Model 3 in Table 1. We estimated thresholds of $\omega_1 = -0.026$ and $\omega_2 = 0.519$, meaning that lagged values of media tone between the two thresholds fall within the band, and any below ω_1 or above ω_2 fall outside it. This is depicted in Figure 4, which shows both ω_1 and ω_2 (as dashed horizontal lines). Values of media tone in the "out-of-band" region in Figure 4 are allowed to behave different dynamically than those in-band. Unfortunately, there is no way to test a band-TAR against the non-threshold alternative, nor can we calculate confidence intervals around the thresholds. While some caution is warranted since both AIC and SBIC indicate that the standard ARDL or TAR specifications are preferred, results in Model 3 suggest that leading economic indicators not only effect media tone when it is positive—as we saw in Model 2 with the standard TAR—but that when media tone is positive enough (outside the upper-limit of the band given by ω_2) there is also no evidence that economic indicators affect media tone. In other words, it is only within the band of approximately zero to 0.519 that future predictions about the economy have any effect on the tone of the media.

To see whether our threshold modeling approach has any substantive implications for the effect of economic indicators on media tone, we plot the short- and long-run effects of the Leading EI variable—the key finding in Model 1—in Figure 5 for all three models. For short-run effects, leading economic indicators are positive but not statistically significantly different from zero when lagged media tone is negative (Model 2, TAR), or when lagged media tone is outside the band (Model 3, B-TAR). Instead, only when lagged media tone is positive (or inside the band) do we see results similar to those in SSW; positive increases in leading economic indicators lead to positive increases in tone of just over 0.1 in the contemporaneous period.

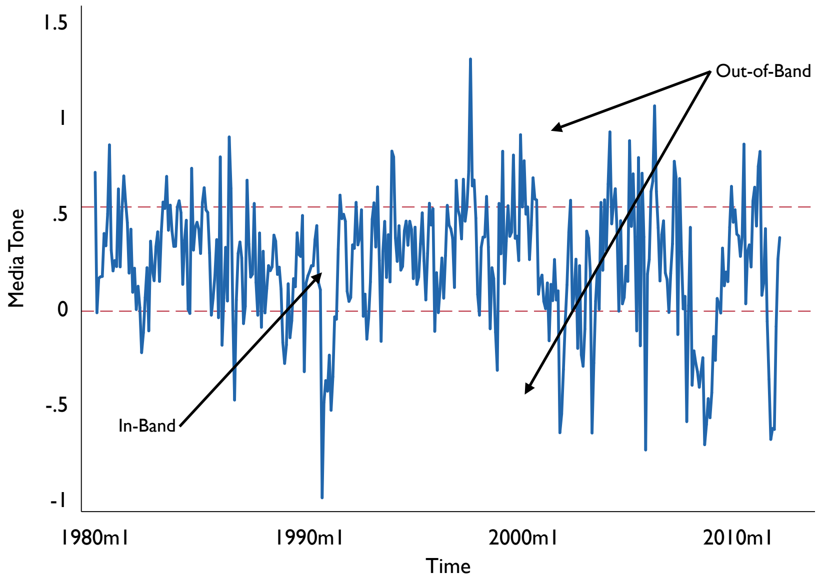

Figure 4: Leading Economic Indicators Only Matter Between the Band

Note: Dashed horizontal lines show the estimated thresholds from Table 1, Model 3, of $\omega = -0.026, 0.519$.

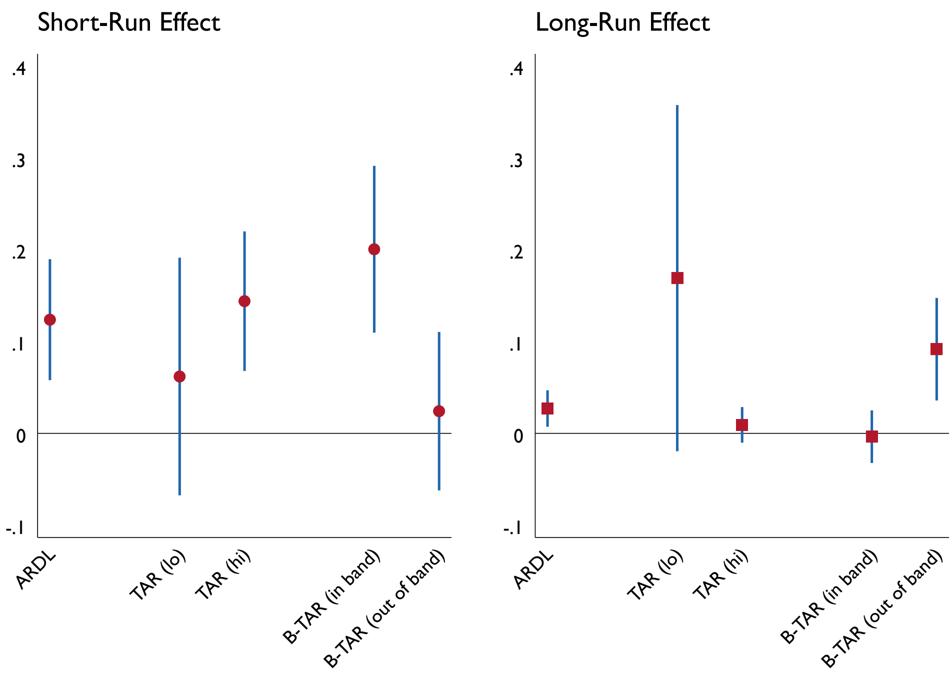

Figure 5: Leading Economic Indicators: Short- and Long-Run Effects Across the Two Models

Note: Long-run effects calculated by dividing the coefficients on the contemporaneous and lagged leading economic indicators by one minus the coefficient on lagged tone. Standard errors approximated using the delta method. 95% confidence intervals shown.

For long-run effects, SSW find that the long-run effect is positive and statistically significant, albeit quite small. Instead, using our TAR model we find that when previous media tone is negative, the long-run effect of an increase in leading economic indicator is large and positive in magnitude, but is not statistically significant. Only when previous tone is positive is there no evidence of a significant long-run effect. The band-TAR results tell a slightly different story. While we find virtually no long-run effect of leading economic indicators when media tone is inside the band, when it is outside there appears to be a positive and statistically significant long-run effect of leading economic indicators on the tone of the media.[15]

Our findings add substantial nuance to the argument put forth by SSW. We find, as hypothesized, that a natural threshold exists in media tone around zero, and that dynamic effects differ on either side of this threshold. Moreover, media tone changes slowly when it is negative, but quickly when it is positive. We also investigated whether a band might exist in media tone. Substantively, our threshold models lead us to conclude that the short-run positive relationship between leading economic indicators and media tone only exists when media tone is positive, or positive but not above 0.519 (as the B-TAR results indicate). In other words, SSW's findings regarding short-run effects only hold in these instances. Positive long-run effects—larger than those found by SSW—between leading economic indicators and media tone occur when the tone of the media is outside of the threshold band.

Mapping Theory to Model Specification

We have covered several models throughout this paper. To better illustrate where our proposed modeling strategy fits into existing model specifications, in Table 2 we provide a list of model specifications that match up with different types (linear vs. threshold)

[15]Readers may notice that while the out-of-band short-run effects are not statistically significant, the long-run effects are. While some suggest that researchers should only explore long-run effects if there is a statistically significant short-run effect (Enns, Moehlecke and Wlezien 2021), others show that long-run effects can occur even when there is no statistically significant short-run effect (De Boef and Keele 2008; Philips 2021), for instance in dead-start models.

of theoretical expectations. We use the relationship between economic growth rates and incumbent vote share as an example to illustrate possible hypotheses and model specifications a researcher may choose to use in their own work.

In Table 2, Example 1 shows a "naive" model—often the standard in the discipline—where the dynamic relationships remain constant, no matter the direction or level of Vote or Growth. If our researcher suspects that this is indeed the data generating process, a simple ARDL model is advisable.[16] In Example 2, the researcher suspects that a one-unit increase in positive growth is not the same as a one-unit increase when growth is negative; indeed, as prospect theory suggests (Kahneman and Tversky 1979), the coefficient for positive growth should be smaller in magnitude than the coefficient when growth is negative. Therefore, a model to test this expectation might create one variable equal to *Growth* when *Growth* is positive, and another equal to *Growth* when *Growth* is negative.[17] This fits with many current models of asymmetric economic voting (c.f., Park 2019).

In Example 3 in Table 2, the researcher has an expectation that the effect of lagged incumbent party *Vote* on *Vote* share in the next election differs based on whether previous incumbent party *Vote* share was above 55 percent. This expectation might come about since such a high vote share is relatively uncommon; thus, we might expect that $\phi^{\geq 55}$ is smaller than $\phi^{<55}$ in magnitude since, on average, *Vote* tends to hover around some mean value of slightly greater than 50 percent. Similar to Example 2, to test this assumption the researcher can parameterize $Vote_{t-1}$ by splitting it based on whether vote share was above or below 55 percent, and testing whether $\phi^{<55} = \phi^{\geq 55}$.

In Example 4, the researcher theorizes that *Growth* has threshold effects at some positive value; for instance, increases in *Growth* may increase incumbent *Vote* share, but at some point further increases may have a much smaller positive effect. Since it is not clear exactly where this discontinuity lies, the researcher can estimate a threshold autoregressive model, with past values of *Growth* determining the threshold.[18] Example

[16] Assuming all stationarity conditions are met.

[17] Equivalently, one could simply interact *Growth* with a dichotomous variable equal to one if *Growth* is positive and 0 otherwise.

[18] We keep the effect of $Vote_{t-1}$ constant, although this assumption could be relaxed if we suspected that this threshold also affects dynamics in *Vote* itself.

Table 2: Summary of Potential Model Specifications: An Economic Voting Example

Example	Expectation	Non-Linear Analogy	Model	Hypothesis Test, e.g.,
1	Growth/Vote have constant effects	Standard ARDL (no non-linearities)	$\text{Vote}_t = \alpha + \phi \text{Vote}_{t-1} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t$	$\beta_1 = 0$
2	Positive Growth $\neq$ negative Growth	Asymmetry in the independent variable	$\text{Vote}_t = \alpha + \phi \text{Vote}_{t-1} + \beta_1^+ \text{Growth}_t^+ + \beta_2^+ \text{Growth}_{t-1}^+ + \beta_1^- \text{Growth}_t^- + \beta_2^- \text{Growth}_{t-1}^- + \varepsilon_t$	$\beta_1^+ = \beta_1^-$
3	Dependence in Vote differs when $\text{Vote}_{t-1} \geq 55$	Asymmetry in the dependent variable	$\text{Vote}_t = \alpha + \phi^{\geq 55} \text{Vote}_{t-1}^{\geq 55} + \phi^{<55} \text{Vote}_{t-1}^{<55} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t$	$\phi^{\geq 55} = \phi^{<55}$
4	Increasing growth may have decreasing returns	TAR (Growth_{t-1} determines threshold)	$\text{Vote}_t = \begin{cases} \alpha + \phi \text{Vote}_{t-1} + \beta_1^* \text{Growth}_t + \beta_2^* \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Growth}_{t-1} \leq \omega \\ \alpha + \phi \text{Vote}_{t-1} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Growth}_{t-1} > \omega \end{cases}$	Existence of ω; $\beta_1^* = \beta_1$
5	Dependence in Vote differs when previous Vote was high	TAR (Vote_{t-1} determines threshold)	$\text{Vote}_t = \begin{cases} \alpha^* + \phi^* \text{Vote}_{t-1} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Vote}_{t-1} \leq \omega \\ \alpha + \phi \text{Vote}_{t-1} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Vote}_{t-1} > \omega \end{cases}$	Existence of ω; $\phi^* = \phi$
6	Growth/Vote effects differ when previous Vote was high	TAR (Vote_{t-1} determines threshold)	$\text{Vote}_t = \begin{cases} \alpha^* + \phi^* \text{Vote}_{t-1} + \beta_1^* \text{Growth}_t + \beta_2^* \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Vote}_{t-1} \leq \omega \\ \alpha + \phi \text{Vote}_{t-1} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Vote}_{t-1} > \omega \end{cases}$	Existence of ω; $\phi^* = \phi$
7	Growth/Vote effects differ when $45 \leq \text{Vote}_{t-1} \leq 55$	Band TAR (Vote_{t-1} determines threshold)	$\text{Vote}_t = \begin{cases} \alpha^* + \phi^* \text{Vote}_{t-1} + \beta_1^* \text{Growth}_t + \beta_2^* \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \omega_1 \leq \text{Vote}_{t-1} \leq \omega_2 \\ \alpha + \phi \text{Vote}_{t-1} + \beta_1 \text{Growth}_t + \beta_2 \text{Growth}_{t-1} + \varepsilon_t, & \text{if } \text{Vote}_{t-1} < \omega_1, \text{ or } \text{Vote}_{t-1} > \omega_2 \end{cases}$	$\phi^* = \phi$; $\beta_1^* = \beta_1$

5 shows a similar expectation to Example 4, although now the assumption is that while the effect of *Growth* remains constant, dependence in past values of *Vote* might change on either side of a threshold.[19] Example 6 also has the expectation that lagged values of incumbent party *Vote* determines the threshold, but allows *Growth* to also take on potential non-linearities. Such an expectation is consistent if we think that economic growth has differing effects based on whether the prior vote share was particularly large or small; for instance, if the previous vote share was quite large, we may expect economic conditions to play less of a role (the incumbent is likely to only lose vote share from this peak) than if vote share was previously small. Last, in Example 7, our researcher might have an expectation that the economic voting model works well under "normal" political conditions, but outside of these bounds—for instance, during periods of war when the incumbent enjoys greater support despite economic performance—this relationship may be weaker. To test this, the researcher can estimate a band-TAR model where $Vote_{t-1}$ determines the threshold, and test whether the effects differ inside versus outside of this band.

Conclusion

While social scientists have readily adopted dynamic models into their methodological toolkits, they have mostly relied on models that assume constant dynamic effects. We have shown how incorporating threshold effects are one way to relax this restriction. While not new to all fields, these models appear to be underutilized, if not absent, from many social science applications. The TAR modeling approach is flexible and can be used in a number of ways for social science theories and data. Researchers can model a threshold in either the dependent variable, as we did in our applied example, or an independent variable. While we used an applied example with time series data, researchers can use a TAR modeling approach when working with data that are not time series. Moreover, this approach is especially useful to scholars whose theories have already proposed asymmetries

[19] Since the threshold variable is the dependent variable itself, we might also unrestrict the constant on either side of the threshold, as shown in Table 2.

or non-linear effects.

Replicating a prominent article, we have shown how the TAR approach allows scholars to create testable hypotheses about the threshold, as well as create directional hypotheses for effects that occur on either side of the threshold. To better summarize the models discussed in this paper for users, in the SI we discuss several more advanced threshold models as well as specification testing. Threshold modeling strategies allow for a richer set of dynamic processes that can, at times, be a better reflection of researchers' theoretical expectations.

As is clear from the examples in Table 2, the approaches discussed in this paper allow us to model a variety of non-linear processes, as well as test our theoretical expectations about whether or not such asymmetries exist. Which model is 'correct'? We stress that researchers should choose models that best fit their theoretical expectations. If a theory suggests that asymmetries are present in either the independent or dependent variable, the model should be adjusted to account for this. Specification testing, which we have discussed above, is an additional complimentary tool that can be used in tandem with theory. By incorporating threshold models into their methodological toolkits, scholars will be able to better understand current theories, as well as create new theory.

References

Balke, Nathan S and Thomas B Fomby. 1997. "Threshold cointegration." *International Economic Review* pp. 627–645.

Baumgartner, Frank R and Bryan D Jones. 2010. *Agendas and instability in American politics.* University of Chicago Press.

Baumgartner, Frank R, Christoffer Green-Pedersen and Bryan D Jones. 2006. "Comparative studies of policy agendas." *Journal of European Public Policy* 13(7):959–974.

Berry, William D, Matt Golder and Daniel Milton. 2012. "Improving tests of theories positing interaction." *The Journal of Politics* 74(3):653–671.

Brandt, Patrick T, John R Freeman and Philip A Schrodt. 2011. "Real time, time series forecasting of inter-and intra-state political conflict." *Conflict Management and Peace Science* 28(1):41–64.

Bruch, Elizabeth E and Robert D Mare. 2006. "Neighborhood choice and neighborhood change." *American Journal of sociology* 112(3):667–709.

Chan, Kung-Sik. 1993. "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model." *The annals of statistics* pp. 520–533.

Choi, Woon Gyu. 1999. "Asymmetric monetary effects on interest rates across monetary policy stances." *Journal of Money, Credit and Banking* pp. 386–416.

Clarke, Harold D, Karl Ho and Marianne C Stewart. 2000. "Major's lesser (not minor) effects: prime ministerial approval and governing party support in Britain since 1979." *Electoral Studies* 19(2):255–273.

De Boef, Suzanna and Luke Keele. 2008. "Taking time seriously." *American Journal of Political Science* 52(1):184–200.

Enders, Walter. 2010. *Applied econometric time series.* 3 ed. John Wiley and Sons.

Enders, Walter and Clive William John Granger. 1998. "Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates." *Journal of Business & Economic Statistics* 16(3):304–311.

Enns, Peter K, Carolina Moehlecke and Christopher Wlezien. 2021. "Detecting true relationships in time series data with different orders of integration." *Political Science Research and Methods* pp. 1–13.

Freeman, John R, Jude C Hays and Helmut Stix. 2000. "Democracy and markets: The case of exchange rates." *American Journal of Political Science* pp. 449–468.

Funk, Kendall D, Hannah L Paul and Andrew Q Philips. 2021. "Point break: using machine learning to uncover a critical mass in women's representation." *Political Science Research and Methods* pp. 1–19.

Goldfeld, Stephen M and Richard E Quandt. 1973. "A Markov model for switching regressions." *Journal of econometrics* 1(1):3–15.

Gonzalo, Jesus and Jean-Yves Pitarakis. 2002. "Estimation and model selection based inference in single and multiple threshold models." *Journal of Econometrics* 110(2):319–352.

Granger, Clive WJ, Timo Terasvirta et al. 1993. "Modelling non-linear economic relationships." *OUP Catalogue* .

Granovetter, Mark. 1978. "Threshold models of collective behavior." *American journal of sociology* 83(6):1420–1443.

Hamilton, James D. 1989. "A new approach to the economic analysis of nonstationary time series and the business cycle." *Econometrica: Journal of the Econometric Society* pp. 357–384.

Hamilton, James D. 1993. "Estimation, inference, and forecasting of time series subject to changes in regime." *Handbook of statistics* 11:231–260.

Hamilton, James D. 1995. "Time series analysis." *Economic Theory. II, Princeton University Press, USA* pp. 625–630.

Hansen, Bruce E. 2000. "Sample splitting and threshold estimation." *Econometrica* 68(3):575–603.

Hansen, Bruce E. 2011. "Threshold autoregression in economics." *Statistics and its Interface* 4(2):123–127.

Hays, Jude C, John R Freeman and Hans Nesseth. 2003. "Exchange rate volatility and democratization in emerging market countries." *International Studies Quarterly* 47(2):203–228.

Jennings, Will and Peter John. 2009. "The dynamics of political attention: Public opinion and the Queen's Speech in the United Kingdom." *American Journal of Political Science* 53(4):838–854.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2):263–292.

Kappe, Roland. 2018. "Asymmetric evaluations: Government popularity and economic performance in the United Kingdom." *Electoral Studies* 53:133–138.

Levine, Jeremy R, Theodore S Leenman, Carl Gershenson and David M Hureau. 2018. "Political places: Neighborhood social organization and the ecology of political behaviors." *Social Science Quarterly* 99(1):201–215.

Lipsmeyer, Christine S. 2011. "Booms and busts: how parliamentary governments and economic context influence welfare policy." *International Studies Quarterly* 55(4):959–980.

Matsubayashi, Tetsuya. 2007. "Population size, local autonomy, and support for the political system." *Social science quarterly* 88(3):830–849.

Metzger, Shawna K and Benjamin T Jones. 2016. "Surviving phases: Introducing multistate survival models." *Political analysis* 24(4):457–477.

Murphy, Alexandra K and Danielle Wallace. 2010. "Opportunities for making ends meet and upward mobility: Differences in organizational deprivation across urban and suburban poor neighborhoods." *Social Science Quarterly* 91(5):1164–1186.

Park, Jong Hee. 2010. "Structural change in US presidents' use of force." *American Journal of Political Science* 54(3):766–782.

Park, Ju Yeon. 2019. "Punishing without rewards? A comprehensive examination of the asymmetry in economic voting." *Electoral Studies* 57:1–18.

Philips, Andrew Q. 2021. "How to avoid incorrect inferences (while gaining correct ones) in dynamic models." *Political Science Research and Methods* pp. 1–11.

Philips, Andrew Q, Amanda Rutherford and Guy D Whitten. 2015. "The dynamic battle for pieces of pie—Modeling party support in multi-party nations." *Electoral Studies* 39:264–274.

Quandt, Richard E. 1972. "A new approach to estimating switching regressions." *Journal of the American statistical association* 67(338):306–310.

Soroka, Stuart N. 2006. "Good news and bad news: Asymmetric responses to economic information." *Journal of Politics* 68(2):372–385.

Soroka, Stuart N, Dominik A Stecula and Christopher Wlezien. 2015. "It's (change in) the (future) economy, stupid: economic indicators, the media, and public opinion." *American Journal of Political Science* 59(2):457–474.

Soroka, Stuart N, Dominik A Stecula and Christopher Wlezien. 2016. "Erratum to "It's (Change in) the (Future) Economy, Stupid: Economic Indicators, the Media and Public Opinion" American Journal of Political Science, Vol. 59, No. 2, April 2015, Pp. 457–474." *American Journal of Political Science* 60(3):803–804.

Terasvirta, Timo and Heather M Anderson. 1992. "Characterizing nonlinearities in business cycles using smooth transition autoregressive models." *Journal of Applied Econometrics* 7(S1).

Tong, Howell. 2012. *Threshold models in non-linear time series analysis*. Vol. 21 Springer Science & Business Media.

Tong, Howell and Keng S Lim. 1980. "Threshold autoregression, limit cycles and cyclical data." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 245–292.

Wimpy, Cameron and Guy D Whitten. 2017. "What Is and What May Never Be: Economic Voting in Developing Democracies." *Social Science Quarterly* 98(3):1099–1111.

Zhang, Songlin, Haoqing Zheng, Hongyong Zhou, Qing Shao and Qun Wu. 2021. "Sustainable land urbanization, urban amenities, and population urbanization: Evidence from city-level data in China." *Social Science Quarterly* 102(4):1686–1698.