Supplementary Information for: "What Goes up Must Come Down: Theory and Model Specification of Threshold Dynamics"

May 24, 2022

Contents

1 Software for Estimating Threshold Models

We use Stata's `threshold` command throughout this paper. R users can use the `tsDyn` package to estimate the TAR model. The `tsDyn` package contains functions which allow the user to compare the fit of these nonlinear models, although one cannot include multiple covariates, only lags of the dependent variable. Readers may have also come across the `TAR` command in R, which uses a Bayesian approach to modeling threshold autoregressive time series, although we do not discuss Bayesian approaches to modeling thresholds in this paper. In EViews, researchers can use the `threshold` command to specify TAR models.

2 Other Threshold-Style Models

In the main paper, we briefly mentioned a number of alternative threshold models. We covered the simpler threshold models in the main paper because they mesh well with a number of theories in political science, and because there are readily-available tests for them in regards to establishing whether a threshold exists. They are also implemented in popular software such as Stata. Below, we discuss a number of alternative threshold modeling strategies. Many of them can be estimated in specialized time series software such as EViews or RATS.

A popular set of non-linear autoregressive models are variants on the smooth transition autoregressive (or STAR) model (Terasvirta and Anderson 1992; Dijk, Teräsvirta and Franses 2002). For a single variable, y_t, with no regressors, where some y_{t-d} is thought to be the transition variable, this is given as:

$$\begin{aligned} y_t =& (\beta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_d y_{t-d})(1 - G(y_{t-d}, \gamma, c)) + \\ & (\beta_0^* + \phi_1^* y_{t-1} + \phi_2^* y_{t-2} + \cdots + \phi_d^* y_{t-d})(G(y_{t-d}, \gamma, c)) + \varepsilon_t \end{aligned} \quad (1)$$

We could include independent variables in this equation, as well as let some other variable determine the transition (i.e., substitute y_{t-d} in $G(.)$ for some exogenous variable). It is also assumed that the residual is i.i.d.: $\varepsilon \sim N(0, \sigma^2)$. The STAR model differs from the

TAR-style models discussed in the main paper in that, while there are essentially two regimes, STAR models allow the transition between the two regimes to occur smoothly.

There are two common specifications for the smooth transition function, $G(.)$. One is the logistic STAR (LSTAR) function:

$$G(y_{t-d},\gamma,c) = (1+exp[-\gamma(y_{t-d}-c)])^{-1} \tag{2}$$

where $\gamma > 0$. As is clear from Equation 2, two additional parameters are estimated, γ and c. γ can be thought of a parameter that "determines the smoothness of the change in the value of the logistic function and, thus, the smoothness of the transition from one regime to the other" (Dijk, Teräsvirta and Franses 2002, p. 3). c is akin to the threshold value in a TAR model. In fact, if γ is large enough, the smooth transition becomes effectively instantaneous, and the model becomes a two-regime TAR. Examples using LSTAR models include Bradley and Jansen (2004), who look at excess stock market returns and industrial production, and Hall, Skalin and Teräsvirta (2001), who model the weather processes of El Niño.

Another common specification involves using an exponential STAR (ESTAR) function:

$$G(y_{t-d},\gamma,c) = 1-exp[-\gamma(y_{t-d}-c)^2] \tag{3}$$

See Gregoriou and Kontonikas (2006), who use an ESTAR approach to model inflation in seven countries.

A third specification (which is far less common) uses a second-order logistic function for $G(.)$ (Dijk, Teräsvirta and Franses 2002):

$$G(y_{t-d},\gamma,c) = (1+exp[-\gamma(y_{t-d}-c_1)(y_{t-d}-c_2])^{-1} \tag{4}$$

where $\gamma > 0$, and $c_1 \leq c_2$.

More complicated specifications for STAR models are possible, including time-varying

STAR (TVSTAR) model (Lundbergh, Teräsvirta and Van Dijk 2003; Van Dijk, Strikholm and Teräsvirta 2003), which allow for non-linear dynamics across time, as well as multiple-regime STAR (MRSTAR) models (Van Dijk and Franses 1999), which allow for more than two regimes. See Holt and Craig (2006) for an example of using a TVSTAR specification to model hog-corn price dynamics, and Bradley and Jansen (2004) for a MRSTAR example.

3 Existing Approaches

Despite a substantial body of literature in political science considering dynamic relationships, non-linear dynamic models are not a part of political scientists' core methodological repertoire. Approaches to accounting for non-constant effects between covariates and the dependent variable remain disjointed if they are used at all. This is not for the absence of non-linear modeling techniques, which are frequently employed by economists.[1] What explains this lack of usage? As Richards and Doyle (2000) point out, although many political science theories posit non-linear relationships, they do not map onto simple non-linear functions in the same way that economic theories do.[2] Here, we focus on a class of non-linear models that we believe are particularly relevant to political science theories: threshold autoregressive (TAR) models.

First, we discuss two current ends of the spectrum with regard to modeling non-linearities in dynamic data in political science research. At one end of the spectrum, there may exist theoretical expectations about a structural break or regime shift in the series that we can explicitly parameterize and model. At the other end, hidden Markov processes allow us to examine the effects of some "unobservable" causes of regime shifts in a model (Hamilton 1989). In other words, we may want these processes to inform us as to where structural breaks occur rather than imposing some theoretical expectation. We briefly discuss each of these approaches below, and then introduce asymmetric non-linear models—in particular the TAR model, which falls somewhere between these two ends of

[1]See (Granger, Terasvirta et al. 1993) for a review of non-linear dynamic models used in economics.

[2]One might also note that economics typically models longer time series compared to political scientists.

the spectrum—as an alternative for modeling non-constant effects.

Interactive effects are commonplace in dynamic theories in political science, 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—e.g., the marginal effect remains constant.[3] Generally, non-linear effects have been modeled by transforming the data to ensure linearity in the parameters, such as 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. 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). However, few papers in the literature model other forms of non-linear dynamic effects explicitly, especially regarding the parameters on lagged coefficients, including the lagged dependent variable. An exception is Philips, Rutherford and Whitten (2015), who allow short- and long-run effects in their model of German party support for the four largest parties to differ based on the party in government. They do so to test whether the effect of the evaluation of the leader of the Liberal Party on party support is larger when the party is in coalition compared to when it is not. Another example is by Clarke, Ho and Stewart (2000), who find that the determinants of party support in the UK had differing long-run effects depending on whether Margaret Thatcher or John Major was Prime Minister. A third exception is Esarey and DeMeritt (2014), who model state-dependent processes using the applied example of presidential approval and economic performance. The authors model the non-constant effect of the US economy, conditional on presidential approval, by interacting the lag of presidential approval with current economic performance. They find that high unemployment hurts presidents the most among those with initially high approval ratings. In other words, how quickly presidential approval declines depends on

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

how poorly the economy is doing.

While the above articles explicitly parameterize non-linearities and asymmetries in dynamic models, a different body of literature treats these shifts as unobservable, often modeling them as a hidden Markov process (Hamilton 1989).[4] For 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.

While structural break parameterization and hidden Markov models are reasonable modeling approaches, there are models that bridge the gap between them in two ways. First, we can explicitly hypothesize about where non-linearities lie, similar to parametric approaches. Second, we can obtain parameter estimates through a data-driven approach similar to Markov modeling and the changepoint literature. This threshold modeling approach is more flexible than the fully-parameterized non-linear model, and, in some cases, more reflective of the underlying data-generating process. There are a variety of important threshold effects in autoregressive distributed lag models, which we discuss below.

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

4 Modeling Asymmetric Effects

Modeling the asymmetric effects of regressors on the dependent variable is perhaps more familiar to readers, and similar to some of the examples provided above (e.g. Soroka 2006; Lipsmeyer 2011). This involves explicitly parameterizing or setting the point at which the asymmetries in effects occur in either the independent or dependent variable.

Asymmetries in the Independent Variable

Consider a simple autoregressive distributed lag, ARDL(1,1), with a single weakly exogenous regressor, x_t:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \tag{5}$$

where $\varepsilon_t \sim N(0, \sigma^2)$. It is straightforward to interpret the effects of x_t on y_t (c.f. De Boef and Keele 2008). The contemporaneous effect of x_t on y_t is given by β_0; one period later this effect is β_1. The long-run effect is the total effect that a change in x_t has on y_t, and is given by: $\frac{\beta_0+\beta_1}{1-\alpha_1}$.

While Equation 5 is flexible in terms of dynamics, it assumes that x_t has the same effects on y_t, no matter the level of x_t. As discussed above, asymmetries may be present; for instance, if x_t falls above or below some value, its effect on y_t may differ. To incorporate this, consider instead the following data-generating process for y_t:

$$y_t = \begin{cases} \alpha_0 + \alpha_1 y_{t-1} + \beta_0^* x_t + \beta_1^* x_{t-1} + \varepsilon_t, & \text{if } x_{t-1} \leq \omega \\ \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, & \text{if } x_{t-1} > \omega \end{cases} \tag{6}$$

We could write Equation 9 in a single step by adding D_t, a dichotomous function where

$D_t = 0$ if $x_{t-1} > \omega$, and equal to one if $x_{t-1} \leq \omega$.[5] The resulting model can be written as:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0(1 + D_t\gamma_0)x_t + \beta_1(1 + D_t\gamma_1)x_{t-1} + \varepsilon_t \tag{7}$$

where $(\beta_0 + \gamma_0 = \beta_0^*)$ and $(\beta_1 + \gamma_1 = \beta_1^*)$. The short-run effect when x_t is above the threshold value ω is still β_0, and the long-run effect $\frac{\beta_0+\beta_1}{1-\alpha_1}$. However, when x_t falls on or below the value ω, the contemporaneous effect becomes $\beta_0 + \gamma_0$, and the long-run effect is now $\frac{(\beta_0+\gamma_0)+(\beta_1+\gamma_1)}{1-\alpha_1}$. In other words, the short- and long-run dynamics can change dramatically based on the value of x_t.[6] As an example, we might expect that more negative economic growth has an effect larger in magnitude on the approval rating of the incumbent than a similarly-sized increase in positive economic growth; this would imply that $\omega = 0$, which would allow the effect of positive economic growth to differ from the effect of negative economic growth.

A similar approach could also be taken for modeling first-differences (i.e., changes). For example, in examining the effect of changes in unemployment on public support for an incumbent, it is plausible that positive changes—increases in unemployment—might hurt incumbent support more than negative changes—decreases in unemployment—help the incumbent:

$$y_t = \begin{cases} \alpha_0 + \alpha_1 y_{t-1} + \beta_0^* \Delta x_t + \varepsilon_t, & \text{if } \Delta x_{t-1} \leq \omega \\ \alpha_0 + \alpha_1 y_{t-1} + \beta_0 \Delta x_t + \varepsilon_t, & \text{if } \Delta x_{t-1} > \omega \end{cases} \tag{8}$$

As an extension, analysts could also add asymmetries in the lagged dependent variable (caused by x_t falling on either side of ω), if they thought that the rate of autoregression in the dependent variable itself might change based on x_t.[7] For instance, public mood about

[5] In other words, the function is: $D_t = \begin{cases} 1, & \text{if } x_{t-1} \leq \omega \\ 0, & \text{if } x_{t-1} > \omega \end{cases}$

[6] While we consider "self-exciting" asymmetries—in other words, asymmetries in x_t that occur when its own value falls above or below a particular value—we could just as easily let the effect of x_t on y_t be asymmetric based on the value of some other variable. There may or may not be a theoretical reason to include this variable in the model, although it typically makes sense to do so. Since this falls under interactive specifications more generally, we focus mostly on these self-exciting asymmetries, although the same intuition discussed below applies.

[7] e.g.,

$$y_t = \begin{cases} \alpha_0 + \alpha_1^* y_{t-1} + \beta_0^* x_t + \beta_1^* x_{t-1} + \varepsilon_t, & \text{if } x_{t-1} \leq \omega \\ \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, & \text{if } x_{t-1} > \omega \end{cases} \tag{9}$$

the economy may be less persistent when the economy is doing poorly than when it is doing well. Regardless, all approaches above are possible ways to relax the assumption that x_t—or the change in x_t—always has a homogenous effect on y_t.

Asymmetries in the Dependent Variable

So far, we have considered asymmetric effects determined by the independent variable. It may also be the case that the level of y_t itself determines this asymmetry. That is, we might expect that "self-exciting" asymmetries characterize values of the dependent variable. Moreover, the dynamic effects of the dependent variable may also change, such that (using the notation from our example above):

$$y_t = \begin{cases} \alpha_0 + \alpha_1^* y_{t-1} + \beta_0^* x_t + \beta_1^* x_{t-1} + \varepsilon_t, & \text{if } y_{t-1} \leq \omega \\ \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, & \text{if } y_{t-1} > \omega \end{cases} \tag{10}$$

Here, the short- and long-run dynamics may change based on the value of y_t. Since y_t is typically continuous, it is a bit harder to parameterize ω, although plausible values might be zero, or the mean of the series. As discussed in the previous section, we could also model first-differences, or changes, in the the dependent variable based on asymmetries in previous changes in the dependent variable.

4.1 STAR Threshold Models

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, as well as discuss estimation strategy, in an applied example. 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. For our example below, we apply the TAR and Band-TAR, which are more frequently applicable to political science theories.

5 Testing for Non-Linearities

Before estimating any non-linear model, it is advisable to test whether a non-linear data-generating process is plausible. There are a number of portmanteau-style tests for this. The McLeod and Li (1983) test examines correlations between squared residuals—from a standard linear model—and its successive lags by forming a Lung-Box Q statistic:

$$Q = T(T+2)\sum_{i=1}^{m}\frac{\rho_i^2}{T-i} \tag{11}$$

where ρ_i is the correlation coefficient between ε_t and ε_{t-i} (Enders 2010, p. 435).[8] The test statistic Q is asymptotically distributed χ^2 with m degrees of freedom. Rejecting the null hypothesis provides evidence of non-linear effects, although it does not specify the type of non-linearity present. Using the saved residuals from the standard ARDL model in Table 1, Model 1 in the main paper, we reject the null hypothesis of linearity across all but the first lag (for up to five lags), as shown in Table 1. This suggests that a non-linear process is present.

As an additional test, we rely on the Regression Error Specification Test (RESET) to test a null hypothesis of linearity against the alternative of non-linearity, again without specifying the particular form of non-linearity. Once more using the residuals from the

[8]The McLeod and Li (1983) test can also be specified as:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1\hat{\varepsilon}_{t-1}^2 + \cdots + \alpha_n\hat{\varepsilon}_{t-n}^2 + \nu_t \tag{12}$$

which is the same as the ARCH-Lagrange Multiplier test (Enders 2010, p. 436). The null hypothesis is that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. The test statistic is distributed χ^2 with n degrees of freedom, although an F-test can be used in small samples.

Table 1: Results Suggest that a Non-Linear Process is Present

Test		Value
McLeod and Li	Lag $i = 1$	1.67
	Lag $i = 2$	6.72*
	Lag $i = 3$	10.53*
	Lag $i = 4$	12.09*
	Lag $i = 5$	14.33*
RESET	$H = 3$	3.98*
	$H = 4$	3.63*
	$H = 5$	2.72*

Note: $* \ p < 0.05$. H_0 : linear data-generating process. Q statistics shown for McLeod and Li test, and F statistics shown for the RESET.

linear regression in Table 1, Model 1 in the main paper, we estimate:

$$\hat{\varepsilon}_t = \boldsymbol{\delta}\mathbf{z}_t + \sum_{h=2}^{H} \alpha_h \hat{y}_t^h \tag{13}$$

where H is usually 3 or 4, $\hat{y}_t$ are the fitted values from the linear model, and $\mathbf{z}_t$ is a matrix of all regressors from the linear model. Since the variables in Equation 13 should have low explanatory power of the residuals, so the test consists of an F-test that $\alpha_2 = \cdots = \alpha_H = 0$, with standard critical values from the F-statistic. The alternative hypothesis is non-linearity. As shown in Table 1, we reject the null hypothesis of linearity across plausible values of H. Overall, given the results of the McLeod and Li and RESET tests, we have strong evidence to support a non-linear modeling strategy. For further discussion of non-linearity testing, see Granger, Terasvirta et al. (1993), Diks and Manzan (2002), andFrühwirth-Schnatter (2006).

References

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

Bradley, Michael D and Dennis W Jansen. 2004. "Forecasting with a nonlinear dynamic model of stock returns and industrial production." *International Journal of Forecasting* 20(2):321–342.

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.

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.

Dijk, Dick van, Timo Teräsvirta and Philip Hans Franses. 2002. "Smooth transition autoregressive models—a survey of recent developments." *Econometric reviews* 21(1):1–47.

Diks, Cees and Sebastiano Manzan. 2002. "Tests for serial independence and linearity based on correlation integrals." *Studies in Nonlinear Dynamics & Econometrics* 6(2).

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

Esarey, Justin and Jacqueline HR DeMeritt. 2014. "Defining and modeling state-dependent dynamic systems." *Political Analysis* 22(1):61–85.

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.

Frühwirth-Schnatter, Sylvia. 2006. *Finite mixture and Markov switching models*. Springer Science & Business Media.

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

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

Gregoriou, Andros and Alexandros Kontonikas. 2006. "Inflation targeting and the stationarity of inflation: New results from an ESTAR unit root test." *Bulletin of Economic Research* 58(4):309–322.

Hall, Anthony D, Joakim Skalin and Timo Teräsvirta. 2001. "A nonlinear time series model of El Nino." *Environmental Modelling & Software* 16(2):139–146.

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.

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.

Holt, Matthew T and Lee A Craig. 2006. "Nonlinear Dynamics and Structural Change in the US Hog—Corn Cycle: A Time-Varying STAR Approach." *American Journal of Agricultural Economics* 88(1):215–233.

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

Lundbergh, Stefan, Timo Teräsvirta and Dick Van Dijk. 2003. "Time-varying smooth transition autoregressive models." *Journal of Business & Economic Statistics* 21(1):104–121.

McLeod, Allan I and William K Li. 1983. "Diagnostic checking ARMA time series models using squared-residual autocorrelations." *Journal of time series analysis* 4(4):269–273.

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

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

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.

Richards, Diana Eva-Ann and Diana Richards Doyle. 2000. *Political complexity: Nonlinear models of politics*. University of Michigan Press.

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

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.

Van Dijk, Dick, Birgit Strikholm and Timo Teräsvirta. 2003. "The effects of institutional and technological change and business cycle fluctuations on seasonal patterns in quarterly industrial production series." *The Econometrics Journal* 6(1):79–98.

Van Dijk, Dick and Philip Hans Franses. 1999. "Modeling multiple regimes in the business cycle." *Macroeconomic dynamics* 3(3):311–340.