Journal of Public Policy (2018), 38:4, 553–579.
doi:10.1017/S0143814X17000186

A dynamic linear modelling approach to public policy change

MATT W. LOFTIS
Department of Political Science, Aarhus University, Denmark
E-mail: mattwloftis@ps.au.dk

PETER B. MORTENSEN
Department of Political Science, Aarhus University, Denmark
E-mail: peter@ps.au.dk

Abstract: Theories of public policy change, despite their differences, converge on one point of strong agreement: the relationship between policy and its causes can and does change over time. This consensus yields numerous empirical implications, but our standard analytical tools are inadequate for testing them. As a result, the dynamic and transformative relationships predicted by policy theories have been left largely unexplored in time series analysis of public policy. This article introduces dynamic linear modelling (DLM) as a useful statistical tool for exploring time-varying relationships in public policy. The article offers a detailed exposition of the DLM approach and illustrates its usefulness with a time series analysis of United States defense policy from 1957 to 2010. The results point the way for a new attention to dynamics in the policy process, and the article concludes with a discussion of how this research programme can profit from applying DLMs.

Key words: policy dynamics, public policy, public spending, Punctuated Equilibrium Theory

Many policy theories incorporate the idea that, in one way or the other, the causal effects driving public policy change over time. The Advocacy Coalition Framework proposed by Sabatier (1987, 1988) and later developed in collaboration with Jenkins-Smith (Sabatier and Jenkins-Smith 1993; Jenkins-Smith et al. 2014) is one. A second is the theory of social learning and policy paradigms introduced by Hall (1993), and a third theory is the Punctuated Equilibrium Theory developed by Baumgartner and Jones (1993).

As laid out in this article, the three theoretical approaches differ in their conceptualisation, description and explanation of policy change. However, they all depict a policy-making process in which decisionmakers' weighting and evaluation of what is important to policy decisions is likely to change over time (for reviews, see Weible et al. 2009; Baumgartner et al. 2014; Cairney and Heikkila 2014; Hogan and Howlett 2015). Thus, policy theory may tell us that an effect waxes or wanes from one period to another, or even that it changes direction. Such change may be slow or fast, gradual or punctuated. It may reflect change in policymakers' core beliefs, change in the paradigms dominating policymaking or change in the images of various policy areas. Nevertheless, a common implication of the dominant public policy theories is that the importance of explanatory factors varies over time.

In this article, we propose a dynamic linear modelling (DLM) approach as a means of more closely aligning our statistical modelling choices with the dynamic implications of the dominant theories of public policy. Some policy studies have attempted to take into account time-varying effects by specifying interaction terms or by making *a priori* predictions about the timing of shifts and accordingly dividing time series into various epochs. Such modelling approaches that address temporal heterogeneity in coefficient estimates come closer to the ontology of our theories of public policy, but they overlook an important theoretical gap: policy theories rarely make point predictions that tell us when changes in the relationship between x and y take place. Furthermore, even if a theoretical argument identifies a specific point in time when parameters are expected to shift, it requires a flexible estimation method to evaluate such a claim empirically. Thus, we argue that public policy researchers should favour estimation methods that are guided by the data and their chronology, rather than *a priori* information, with regard to the timing and direction of shifts.

The DLM approach is aimed at directly estimating time-varying influences on public policy. The method is familiar in political science, but has primarily been used to study public opinion and international conflict (Beck 1983; Gerber and Green 1998; Green et al. 1999; Mitchell et al. 1999; Brandt et al. 2000; Wood 2000; Mcavoy 2006; Enns and McAvoy 2011; Fariss 2014). There are other approaches to identify time-varying relationships, but as we argue, the DLM approach is particularly well-aligned with the challenges posed by dynamic public policy theories.

In this article, we describe and demonstrate the basic methodology for unfamiliar readers and argue that the tools currently available for building and estimating these models allow researchers to build the right models for their own policy data. Fundamentally, we argue that, with regard to investigating dynamic effects on public policy, the method we propose accomplishes this efficiently and in a way that approximates the processes described by the major theories of public policy.

To corroborate this claim, we conduct a DLM analysis of change in United States (US) defense policy from 1956 to 2010. The case of US defense is a particularly attractive test ground given the many previous time series studies of this policy area, which provides direction for the choice of explanatory variables, and given the availability of high-quality data spanning several decades.

Motivating a dynamic approach to modelling public policy

A growing body of public policy theories argues that politics is a terrain that is restructured over time. A prime example is the work of Baumgartner, Jones and colleagues, which claims that over time most policy areas may be subject to changes in policymakers' focus (Jones 1994; Baumgartner and Jones 2002; Jones and Baumgartner 2005; Baumgartner and Jones 2015). As a consequence, no single logic of policymaking and no single frame of reference lasts forever. Historical contingencies are fundamental to politics, and the threshold for action even within a single policy area may change over time, as the weighting of policy problems is dynamic (Jones and Baumgartner 2005, 90–91).

Other major contributions to the study of public policy have a different conceptualisation, description and explanation of policy change, but they share the important implication that the determinants of public policy may change over time. One example is the Advocacy Coalition Framework, according to which every coalition of policymakers contains secondary aspects, policy core beliefs and so-called deep core beliefs. Secondary aspects comprise a large set of narrower beliefs of the policymakers concerning, for instance, their evaluation of relevant problems and performances within a policy subsystem as well as their preferences regarding regulations or budgetary allocations (Sabatier 1998). Policy core beliefs refer to a coalition's basic normative commitments and causal perceptions across an entire policy domain, and deep core beliefs include basic ontological and normative beliefs operating across almost all policy domains (see also Weible et al. 2009, 2011). When it comes to the core beliefs, these are assumed to be very stable, though not time-invariant. Even if changes mainly involve secondary aspects, this may, for instance, have implications for the relative importance of various causal factors within a given policy domain (Sabatier 1998, 104).

Another example is Hall's (1993) seminal work on social learning and policy paradigms, according to which both the instruments and goals of public policymaking may change over time (see also Campbell 2002; Blyth 2013; Princen and 't Hart 2014). In Hall's thinking, policy change comes in three basic types. First-order changes are policy shifts that leave the larger policy regime in place, for instance, changing the minimum lending rate or fiscal stance to adjust macroeconomic policy. In Hall's analogy, first-order changes adjust the settings on the "instruments" of policymaking. Second-order changes, on the other hand, replace the instruments themselves (Hall 1993, 278–279) – in the example of macroeconomic policy, changes in the system of monetary control or adopting or abandoning strict targets for monetary growth would be second-order changes. Third-order changes alter the instruments, their settings and the hierarchy of goals behind policy. Hall's prime example of a third-order change is the paradigmatic shift in British macroeconomic policy in the 1970s and 1980s from Keynesian to monetarist modes of economic regulation.

Just as changes in deep core beliefs are assumed to be rare according to the ACF, paradigmatic third-order changes are very rare according to Hall's theory. It is important to note, however, that even second-order changes or, in the parlance of ACF, changes in policy core beliefs will most likely be reflected in time-varying relationships in time series models of public policy. A second-order change, for instance, reflects an alternation of the hierarchy of goals behind policy (Hall 1993, 282), and a change in the policy core beliefs of policymakers may involve a change of their causal perceptions of a policy domain (Sabatier 1998, 103). Put differently, if the decisionmakers' causal understanding of the relationship between problems and solutions changes within policy domains such as defense, welfare, transportation or environmental policies, then the relationship between x and y probably changes as well.[1]

Across the different theoretical approaches and different subfields of policy research there is a consensus that the dynamic implications of these theories are not adequately analysed by using standard time series regression techniques (Jones and Baumgartner 2005; Hall 2006, 90–91). The basic problem with standard regression models is simple. If relationships between x and y are time-variant then the basic regression assumption of unit homogeneity is violated. In other words, changes in the value of an explanatory variable, x, will not produce corresponding changes in the value of the policy output variable, y, of the same magnitude, or even direction, across all time (cf. Hall 2003, 382).

As we will discuss in detail below, common responses to this problem are to divide time series into epochs or to specify interaction terms conditioning

[1] In a broader perspective, dynamic causal forces have also been a major theme in comparative politics, exemplified by research on historical institutionalism (Steinmo et al. 1992; Pierson 2000; Lieberman 2002) as well as reflected in more particular discussions about the changing effect of political parties on welfare policies (Pierson 1996; Huber and Stephens 2001; Kwon and Pontusson 2010).

the relationship between x and y. The problem, however, is that such approaches require strong theories about when the relationship between x and y transforms. Often, as argued by Jones and Baumgartner (2005, 23), we have no *a priori* idea when such changes take place: "If we look at a series of policy activities, as we do for crime, social welfare, and other policy areas, we find clear evidence of self-reinforcing changes. While it is possible to model such complex series, how does one know in advance when these interactions will occur?". Similarly, in a recent review of the Advocacy Coalition Framework, Jenkins-Smith et al. (2014) identify four major pathways to policy change, each of which may be a necessary but not sufficient source of change in policy core beliefs. Finally, even if theory does contain a strong prediction of the timing of a transformative policy change, we need a flexible estimation model in order to be able to falsify or support that prediction empirically.

This discontent with standard time series approaches has led policy scholars to advocate alternative methods that are better aligned with their theoretical approaches to public policy. One example is systematic process tracing (Hall 2003), which can be a strong approach to identifying causal processes leading to a political outcome. Another example is the stochastic process approach advocated by Jones and Baumgartner (2005). A major advantage of a stochastic process approach is the focus on entire distributions of data (not only mean and variance) and the lack of assumptions required. Yet, we also set aside important information by using these approaches: giving up the ability to examine directly the relationships between changes in (potentially multiple) policy determinants and changes in policy outputs or ignoring the chronology of our data. Thus, we sympathise with the expansion of methods better aligned with the ontology of the major public policy theories, but we believe there also is a strong case for improving the statistical time-series analysis of policy dynamics. In the next section, we motivate the adoption of a flexible approach to time series policy analysis suited to uncovering time-varying relationships.

Estimating dynamic relationships in public policy research

Given the goal of examining dynamic relationships between an explanatory variable (x) and a policy output (y), we propose that the hypothetical ideal statistical model has two features. First, it would estimate variation in x's relationship to y over time (t), directly from the data. Formally, we might write,

$$y_t = x_t \beta_t + \epsilon_t \tag{1}$$

The t subscript on β_t indicates a new parameter value is estimated each time period, indicating how the relationship between x_t and y_t changes or remains constant. Such a model's estimates would be guided by the data, as they would be estimated directly.

Second, an ideal model would estimate β_t in a way that respects its chronology. That is, given the emphasis theory places on history and path dependence in public policy (cf. Jones and Baumgartner 2005, 49), we must not assume $\beta_{t'}$ is completely independent of $\beta_{t''}$ (e.g. β_t is not independent of β_{t-1} or of β_{t-4}, etc.).

The main barrier to this idealised model is identification – that is, its parameters outnumber its data points and therefore, even with infinitely long time series, one cannot estimate them without making additional assumptions. The time series regression modelling toolkit contains many second-best alternatives, however. We argue that time-varying relationships of the form presented in Equation 1 are common in policy studies, but that the tools commonly used to study them provide answers that are not guided only by the data nor do they account for the data's chronology.

In this section, we provide an accessible introduction to a well-studied method for estimating dynamic relationships that meets both criteria; we review the advantages these criteria hold for public policy research; and we consider the method's strengths relative to commonly used approaches. Namely, we propose that researchers examining policy changes featuring time-varying relationships of this sort consider a class of modelling techniques known as DLM (Kim and Nelson 2000; Mcavoy 2006; Shumway and Stoffer 2010) or flexible least squares (Kalaba and Tesfatsion 1988; Wood 2000).[2] The difference between the two, for our purposes, lies in assumptions about the distribution of errors in DLMs. We refer to DLMs throughout as we adopt this formulation to estimate uncertainty.

In what follows, we make the case for more widely adopting DLMs for studying policy change. It is worth reiterating that DLMs, a subset of the broader category of state-space models, are not unknown in political science (Gerber and Green 1998; Green et al. 1999; Mitchell et al. 1999; Brandt et al. 2000; Wood 2000; Mcavoy 2006; Enns and McAvoy 2011; Fariss 2014). Indeed, Beck (1983) argued that the approach had strong advantages for theory testing and should be added to the arsenal of those who work with time series data. Though they have been applied to study public opinion and international conflict, specifically to deal with the kind of effect heterogeneity that we are concerned with, the method has not gained traction in studies of public policy. Potential applications are,

[2] These approaches are algebraically equivalent approaches to estimating dynamic coefficients in linear models (Montana et al. 2009).

nevertheless, widespread. The goal of estimating time-varying relationships originally motivated state-space models' development in physics and, more recently, has encouraged their adoption and popularisation especially in economics and finance (for background, see: Kim and Nelson 2000; Mergner 2009).

Returning to Equation 1's hypothetical model, the DLM recovers estimates of time-varying effect coefficients, that is, β_t, provided some scaffolding in the way of assumptions about how β_t changes with time. Specifically, in a DLM β_t is a realisation of an underlying "state" that varies as a Markov chain over time. This formulation ensures identification and reflects one of the principles we proposed above: that dynamic relationships should be estimated in a way that respects the data's chronology. As we will see, DLMs can flexibly return time-constant coefficient estimates where appropriate or detect the presence of dynamics. A simple DLM built to emulate Equation 1 might be structured as follows:

$$y_t = \boldsymbol{X}_t\boldsymbol{\beta}_t + v_t \tag{2}$$

$$v_t \sim i.i.d.(0, \sigma_v^2)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{w}_t \tag{3}$$

$$\boldsymbol{w}_t \sim i.i.d.(0, Q)$$

The first line, Equation 2, is called the measurement equation. Inputs to the policy-making process in vector $\boldsymbol{X}_t$ influence policy y_t via the vector of time-varying effect coefficients $\boldsymbol{\beta}_t$. The disturbances, v_t, are independent zero-mean Gaussian noise with variance σ_v^2. The third line, Equation 3, is the state equation. Here, the effect coefficients, $\boldsymbol{\beta}_t$, are modelled as an unobserved state varying in a random walk over time. The state disturbances, in vector $\boldsymbol{w}_t$, are also mean-zero Gaussian noise with covariance matrix Q and are uncorrelated to v_t.

Equations 2 and 3 demonstrate how a DLM respects the goal of producing estimates guided by the data and accounting for chronology. Estimates of $\boldsymbol{\beta}_t$ come directly from the data and are linked over time in Equation 3. This structure allows estimates of $\boldsymbol{\beta}_t$ to be informed by the whole time series, while still ensuring them flexibility to reflect any pattern of variation indicated by the data. What value added does this bring to policy studies examining time-varying relationships? We suggest several. For one, estimates that depend only on the data limit the assumptions analysts must make. As we review below, many methods for identifying dynamic relationships must justify not only their measurements, cases and modelling choices – as in any empirical analysis – but also must justify *a priori* expectations about the timing of shifts in relationships or the use of statistical tests to pinpoint them.

Perhaps more importantly, estimates of dynamic relationships arising from the data enrich our ability to falsify hypotheses. Hypotheses about time-varying relationships may be falsified if a relationship does or does not exist, does or does not vary, or does or does not vary at a particular time. DLMs can speak to these criteria simultaneously. This has the advantage of avoiding bias from or corrections for multiple testing if the analyst wishes to consider different null hypotheses. Furthermore, since DLMs feature time-connected estimates of $\boldsymbol{\beta}_t$, they utilise all the information in the data, maximising the statistical power of the estimates and using information from the full time series to estimate each coefficient.

Consider, briefly, advantages DLMs enjoy relative to other approaches to estimating time-varying relationships. A common approach to this problem is to estimate variation in effect parameters by subdividing time series into shorter epochs and running separate regressions on each subset of the data or, similarly, including interaction terms for epochs. This encounters several weaknesses. The first relates to statistical power. Subdividing time series necessarily weakens inferences by working with smaller data subsamples, increasing risks of both Type I and Type II errors. Second, and potentially more challenging, subdividing requires analysts to identify and motivate the choice of relevant time periods *a priori*. A principled choice of time periods may be available for certain tests, but considering more than one choice ventures into multiple testing. Furthermore, subdividing still forces analysts to assume effect parameters are constant within epochs.

Low statistical power is also a difficulty for CUSUM and CUSUMSQ plots as tests for parameter stability (Wood 2000; Baltagi 2011). These plot functions of cumulative sums of residuals to identify shifts in parameters. Although this approach is estimated directly from the data, it falls short of the criterion of time-connectedness. Having identified a change in a parameter via CUSUM plotting, this method provides no guidance in estimating or accounting for temporal dependency in parameter shifts.

Difficulties in motivating *a priori* information also apply to the Chow test, which can reveal if an effect parameter shift occurs at a particular point (Beck 1983; Wood 2000) – a point the analyst identifies. Again, having identified potential change points, there is no clear method for estimating parameter changes that respect time-dependence.

Concerns with statistical inefficiency and motivating *a priori* information also extend to moving window analysis (Kwon and Pontusson 2010; Finseraas and Vernby 2011). Moving window analyses ignore large portions of data to estimate each bit of dynamics in effect coefficients. Although eventually all data are considered, observations at the beginning and end of time series are less influential than those in the middle, and each individual estimation has less statistical power than would be achieved using all data. Furthermore, decisions such as window size and which observations to include in each analysis can be consequential for the results, and yet are left to the analyst's choice.

Still more systematic methods include multivariate generalised autoregressive conditional heteroskedasticity models (MV-GARCH) or dynamic conditional correlations (DCC) (Engle 2002; Lebo and Box-Steffensmeier 2008). These dynamically estimate the covariance matrix of multivariate time series using a framework adapted from GARCH modelling. This is a powerful method for identifying a certain type of dynamics, namely in correlations. Although they satisfy our proposed criteria of being data-driven and (partially) time-connected, they require quite large data sets and address themselves to correlations and not to the explanatory modelling problem described in Equation 1.[3]

There are, of course, limitations to DLMs. As in any parametric modelling, the assumptions underlying the method matter and model reliability is better confirmed than assumed. Therefore, we provide guidance in probing the conditions for applying DLMs by using appropriate diagnostics and demonstrate how to use in-sample predictions and simulations to check the appropriateness of DLMs for a given application.[4]

Estimating a DLM

A standard approach to estimating DLMs like the one we describe above is via a combination of maximum likelihood and the Kalman filter (Kim and Nelson 2000; Commandeur and Koopman 2007; Petris et al. 2009; Shumway and Stoffer 2010), though Bayesian estimation methods are also common (Petris et al. 2009; Shumway and Stoffer 2010). We describe here how to apply the former approach to estimate a DLM. To illustrate the method, we walk through the process of estimating the model described in Equations 2 and 3 – an example of a dynamic model of the impact of various influences on some policy output in a single time series.

DLMs are estimated recursively. Referring to Equation 3, at each time t, $\boldsymbol{\beta}_{t-1}$ serves as our expectation of the new period's value of $\boldsymbol{\beta}_t$, conditional on the information observed up to time $t-1$. On the basis of this conditional expectation for $\boldsymbol{\beta}_t$, we calculate a prediction for the outcome in time t: $\hat{y}_t$. The error in this prediction, $y_t - \hat{y}_t$, is used to update our final estimate of $\boldsymbol{\beta}_t$, such that larger errors provoke larger shifts in coefficient estimates. The recursion

[3] There is also ongoing debate regarding the econometric properties of DCC/MV-GARCH estimators (Caporin and McAleer 2013).

[4] See Online Appendices D and E.

proceeds sequentially from time 1 through time n. This means that at any point, t, all past information about the underlying state is summarised in the point estimates, $\boldsymbol{\beta}_t$, and their covariance matrix at time $t-1$.

Shifts in the coefficient estimates each period are moderated by the ratio of uncertainty regarding the estimates of $\boldsymbol{\beta}_t$ to the total overall estimation uncertainty in the model (including that from $\boldsymbol{\beta}_t$). Where uncertainty about $\boldsymbol{\beta}_t$ is a larger share of overall model uncertainty, this ratio – termed Kalman gain – is closer to one. Updates to the coefficients become more responsive to that period's prediction error as the Kalman gain approaches one. A formal definition of the Kalman gain and detailed estimating equations are provided in the Online Appendix A.

The final step in the estimation is to smooth the time-varying state estimates using information from the full time series. Kalman smoothing utilises the same approach as the filter, but is run in reverse. Recall that when filtering at each time period, t, all information about the unobserved states up to period t is summarised in the point estimates and associated uncertainty of last period's state estimates: $\boldsymbol{\beta}_{t-1}$. The smoother, by running in reverse from time T back to time 1, updates each period's estimate of $\boldsymbol{\beta}_t$ conditional on all information in the model. Filtering alone is used for applications in which the goal is to forecast future outcomes, whereas smoothing is applied in applications aimed at making inferences about effect parameters. In public policy, our inferences are generally of the latter type, addressing effect estimates in fixed time series. For that goal, filtering by itself suffers several weaknesses relative to filtering plus smoothing. First, changes in filtered parameter estimates lag changes in smoothed parameter estimates (Commandeur and Koopman 2007, 85–89). This is a product of the estimation process and does not negatively impact filtering's predictive accuracy. However, the lagged filtered parameter estimates are out of sync with the best estimate of the actual pattern in the data – a situation smoothing resolves. Second, filtering overstates uncertainty in early estimates of $\boldsymbol{\beta}_t$, as chronologically earlier observations are estimated using fewer data points. Third, filtering can over- or understate the magnitude of shifts in coefficient estimates relative to smoothed estimates. In sum, smoothing provides the more accurate inference for $\boldsymbol{\beta}_t$ because it uses more information than filtering alone (Kim and Nelson 2000, 27). Furthermore, as we have argued our estimates of $\boldsymbol{\beta}_t$ must not be fully independent of one another, we prefer combined filtering and smoothing.[5]

Final effect coefficient estimates are these smoothed time-varying coefficients. The recursive estimation procedure is identified conditional on a set

[5] An example of our main results when applying filtering only can be seen in the Online Appendix I, for reference.

of starting values for $\boldsymbol{\beta}$ ($\boldsymbol{\beta}_0$) and estimates of Q, σ_v^2 and the starting value of the covariance matrix of innovations, $\sum_0$. These are estimated via maximum likelihood. The process begins with initial values for each of these latter parameters, and then each $\boldsymbol{\beta}_t$ is estimated from time 1 using the Kalman filter followed by smoothing.

As can be seen in Equation 2, the DLM also relies on assumptions about the independence, normality and homoskedasticity of errors. Violations of these assumptions can be checked straightforwardly by examining standardised residuals from an estimated model (see, for example, Commandeur and Koopman 2007, 90–96). An assumption common to standard time series techniques that is not encoded in the DLM framework, is stationarity. That is, our methods often require we assume that statistical features of time series inputs to the model, like their mean and variance, do not vary over time. Such variation cannot be captured by static coefficient estimates. However, since coefficients vary with time in the DLM data used in such applications need not exhibit stationarity (Commandeur and Koopman 2007, 134; see also, Petris et al. 2009, 34, 115).

The basic DLM described here can also be extended to accommodate a variety of alternative settings, for example, autoregressive processes, moving averages, constraining certain parameters to be constant, generalised linear models for nonnormally distributed response variables and others (Petris et al. 2009; Shumway and Stoffer 2010). As we show later, the DLM can also be extended to panel data.

Statistical software for estimating DLMs is widely available. An excellent starting point for applying them in the reader's own work is to review volume 41 of the *Journal of Statistical Software*, a special issue on state-space estimation in STATA, R, SAS, RATS and other statistical software packages. The results we present are estimated using the "dlm" library for R (Petris 2010), which is one of several options in R (Petris and Petrone 2011; Tusell 2011).

We turn now to applying DLMs to a real empirical case in policy studies to demonstrate its usefulness and potential applications.

A DLM analysis of US defense policy

We illustrate the dynamic modelling approach using real data on US defense spending, applying the technique to test for dynamic relationships between the explanatory variables and policy outputs. US defense spending offers an attractive test ground for several reasons. First, although the use of public spending as a policy indicator has been subject to some debate, the defense area is one where changes in spending have often been taken as evidence of important changes in defense policies (see e.g. Hartley and

Russett 1992; Wlezien 1996; True 2002). Second, many studies have tried to model time series of US defense spending, which provides some direction for our choice of explanatory variables. The many defense studies partly reflect a genuine interest in this policy field during the Cold War, and partly reflect the availability of rather reliable and long time series within this policy field. A broad range of drivers of US defense spending have been examined in these time series studies including election cycles (Nincic and Cusack 1979; Cusack and Ward 1981; Zuk and Woodbury 1986; Kamlet and Mowery 1987), measures of unemployment (Griffin et al. 1982; Kiewiet and McCubbins 1991; Cusack 1992; Majeski 1992; Su et al. 1993), arms races (Ostrom 1977; Ostrom 1978; Ostrom and Marra 1986; Correa and Kim 1992), inertia (Ostrom 1977; Majeski 1983; Correa and Kim 1992) and public opinion (Hartley and Russett 1992; Higgs and Kilduff 1993; Wlezien 1996). A more complete overview of variables and measures adopted in previous work is included in Table A.1 in the Online Appendix B.[6]

The US defense spending studies disagree with respect to several model assumptions, including the choice of explanatory variables and the exact specification of the dependent variable. Nevertheless, common to virtually all of them is the modelling assumption that relationships are invariant over time. The assumption is so uncontested that it is not given serious consideration in most previous defense spending studies. This is a further argument for investigating whether the approach advocated in this article is able to detect dynamic relationships within such a well-researched field as US defense spending. Only one study we know of escapes this characterisation of the field. True (2002) examined US defense spending over several decades and found that the relationship between defense spending and prior spending, international tensions and wars, and US intelligence estimates of Soviet defense spending showed some evidence of changing over time.

To reach this conclusion, True (2002) split the time series into two periods, 1966–1979 and 1980–1992. Given the limitations with this research strategy we reviewed above, we reevaluate True's findings using a DLM to detect dynamics in effect coefficients over time. We take our analysis of US defense spending in two stages. First, using the dynamic estimation approach advocated here, we analyse US spending based on similar data as True (2002), but for a longer time period, 1956–2010.[7] Second, we

[6] See Correa and Kim (1992) for a similar review containing some older material we have omitted.

[7] Lagged Soviet defense spending is only available between 1966 and 2007 (Russian defense spending is substituted following the Soviet Union's collapse). We collected data from the original US intelligence estimates since we were unable to procure True's own data. See the Online Appendix F for a replication using True's exact time period and measurement choices.

Table 1. Explanatory variables from previous research

Variables	Measure	Model 1*	Model 2†
Lag defense spending	Lagged real defense budget authority (True 2002)	✓	✓
International aid	US foreign aid spending (constant USD) (True 2002)	✓	
War/tension	Indicator taking value of 1 in periods of war (i.e. Korean War, Vietnam War, Reagan buildup, first Gulf war and 2001 onward) (True 2002)	✓	
Lag unemployment	National unemployment rate (True 2002)	✓	✓
Presidential election year	Indicator taking value of 1 in years in which an incumbent competes in presidential election (True 2002)	✓	✓
Lag Soviet spending	Lag Soviet defense spending assembled from US intelligence reports (True 2002)	✓	
Lag Congressional ideology	Lagged polarisation in the House of Representatives, i.e. mean DW-NOMINATE score on first dimension (Rosenthal and Poole 2015)		✓
Change in GDP	% change in GDP (Whitten and Williams 2011)		✓
Hostilities	Summed annual intensity of all US militarised interstate disputes (Whitten and Williams 2011; Palmer et al. 2015)		✓
Lag public opinion	% of Americans saying foreign affairs is an important problem‡ (Cusack 1992)		✓

Note: A more comprehensive list of explanatory variables from previous research can be found in the Online Appendix B. US=United States; GDP=gross domestic product.
*This model builds on True (2002).
†This model builds on the outcome of the model 1 analysis as well as the broader literature on defense spending (see also Table A.1 in the Online Appendix B).
‡Heffington, Colton, Brandon Beomseob Park, and Laron K Williams. (2017) The "Most Important Problem" Dataset (MIPD): a new dataset on American issue importance. *Conflict Management and Peace Science*. [dataset]

broaden the range of explanatory variables in light of the outcomes of the model 1 analysis and in light of the broader literature on US defense spending. We hold constant across both model specifications the response variable: the percentage of annual change in US defense budget authority.[8] Thus, we collected similar data but for an expanded period, 1956–2010.

Table 1 provides an overview of the range and specification of explanatory variables used in models 1 and 2, respectively. As argued above, the dominant theories of public policy do not generally offer guidance about when relationships change, only that they may change over time.

[8] Budget authority data were originally collected by Frank R. Baumgartner and Bryan D. Jones, with the support of National Science Foundation (NSF) grant numbers SBR 9320922 and 0111611, and were distributed through the Department of Government at the University of Texas at Austin. Neither NSF nor the original collectors of the data bear any responsibility for the analysis reported here.

For instance, partisan alignments around the defense issue may change over time, reflecting new problem developments, a new understanding of the issue and/or new linkages of defense to other issues such as terrorism, the economy and employment. Similarly, the importance of public opinion on this issue relative to other drivers of policymaking may change over time resulting in a dynamic impact of public opinion on defense spending. On the other hand, it could also be the case that such transformative change is less prevalent than claimed by theories of public policy. The previous time series analyses on defense spending do not tell us much about this question and it is therefore warranted to initiate a closer inspection of this question utilising the rich time series data on US defense spending.

Summary statistics for all variables in both models can be found in the Online Appendix C. For both models, we summarise results by plotting estimated effect coefficients over time with confidence intervals. Since the model estimates coefficients and standard errors for each year (observation) in the data, coefficient tables are impractical. Results are most easily interpreted graphically. Both models use data standardised by subtracting variable means and dividing by one standard deviation.[9] As a result, both models also exclude constants. For assessments of both models' in-sample predictive accuracy, see the Online Appendix E.

Results: model 1

Dynamic coefficient estimates from our first model specification are plotted in Figure 1. There is some initial evidence here that influences on public spending do not have constant effects over time. Periods of war and international tension are not consistently related to changes in defense spending, for example, with certain periods associated with more and certain periods with less spending. Furthermore, lagged real defense spending (in levels) is not consistently negatively associated with subsequent spending changes, which indicates a dynamic adjustment process (see De Boef and Keele 2008). Whereas a negative coefficient may be interpreted as reflecting negative feedback, with the defense budget regressing back to its "normal" level, a positive coefficient would signal positive feedback in the sense that high budgets would be followed by further increases (see also Wood and Doan 2003). Figure 1 does not show a full shift from negative to positive feedback, but the varying effects over time of the lagged spending measure may be interpreted as evidence in support of the emphasis on changing feedback processes in the punctuated equilibrium theory (Baumgartner and Jones 2002).

[9] See the Online Appendix C for plots of standardised variables.

The rest of the estimates from the model show little evidence of dynamic effects. International aid demonstrates a brief positive association with defense spending in the first decade of the 21st century. This largely reconfirms True's (2002) null finding regarding his hypothesis that international aid and military spending were determined in coordination. The period of our data also present in his analysis shows virtually no evidence in favour of this. Lagged unemployment shows statistically indistinguishable effects throughout, despite its mean estimated effect crossing the zero line during our period of analysis. The effects of both lagged Soviet defense spending and the election year indicator are basically constant and statistically indistinguishable from zero.

As the combination of relatively short time series and numerous parameters being estimated places high demands on our limited amount of data, we may improve our estimates if we can eliminate unimportant information from the model. Thus, we now turn to our revised model specification.

Results: model 2

In our second model, plotted in Figure 2, we see more suggestive evidence of dynamic effects over time. The results in this second model specification are estimated with greater certainty than the first and further sharpen some of the trends glimpsed in the first model. First, consistent with results in Figure 1, Figure 2 shows lagged defense spending is only sometimes negatively associated with percentage change in defense spending. Substantially, the dynamic estimates on lagged defense spending suggest that during the second half of the 1980s and through most of the 1990s the defense budget was regressing back to a "normal" level following a peak at the height of the Cold War. That is, the higher than average lagged spending levels in this period had a negative relationship to spending changes, and, indeed, the defense budget decreased annually from 1986 through 1998.

Second, international hostilities display less dynamics than effects for True's war and tensions indicator in the previous model, and yet it reaches negative statistical significance around the same time as the war/tension indicator – in the early 1990s. Although perhaps counter to intuition, the negative association between intensity of US participation in foreign hostilities and defense spending changes is likely meaningful here. Hostilities vary annually even during periods of international tension and are, therefore, measured more finely than a simple indicator. However, the summary statistics show it does not vary widely – US participation in militarised interstate disputes peaks several times during our period of analysis. Therefore, these estimates tell us that, despite an uptick in US participation in militarised interstate disputes in the 1990s, the period of hostilities

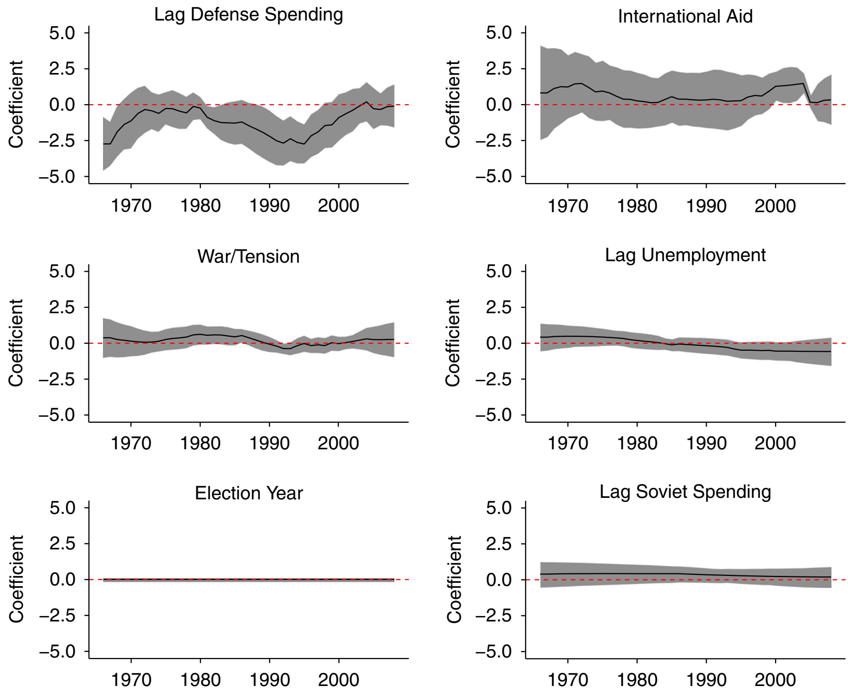

Figure 1 Estimated time-varying coefficients from model 1.
Note: Effect coefficients are plotted over time. The solid line indicates the point estimate for the coefficient each period, whereas grey areas are 95% confidence intervals. The dashed horizontal line indicates 0. Where the grey area overlaps the dashed line, estimates are not statistically significant by conventional standards.

following the end of the Cold War coincided with decreasing military spending, in contrast to the rest of the time series.

Third, the perhaps most noteworthy result in Figure 2 is that the effect of political preferences in the House of Representatives actually changes its direction over time. In contrast with classic conceptions of party positions on the defense issue, a more conservative House ideology had a negative association with changes in defense spending in the 1960s. The estimate became indistinguishable from zero until the late 1990s, and then the association reversed so that more conservative House ideology was related to greater spending through the first few years of the new century. These results indicate that partisan politics intrudes into defense spending changes only irregularly and in changing ways over time. Furthermore, the effects can be decisive. At its greatest positive point, in 2004, a 1 SD shift in House

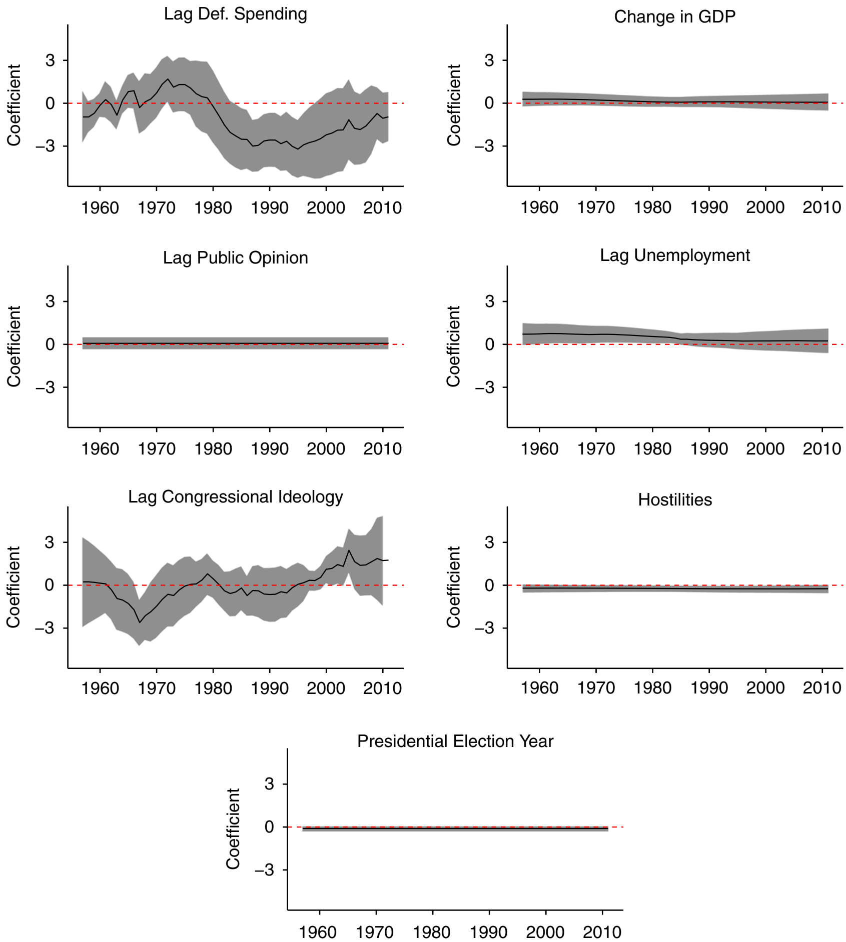

Figure 2 Estimated time-varying coefficients from model 2.
Note: Effect coefficients are plotted over time. The solid line indicates the point estimate for the coefficient each period, whereas grey areas are 95% confidence intervals. The dashed horizontal line indicates 0. Where the grey area overlaps the dashed line, estimates are not statistically significant by conventional standards. GDP = gross domestic product.

ideology towards conservatism would have an estimated effect of increasing defense spending by around 20%. The actual shift in House ideology that year was around one-fourth of a standard deviation towards the conservative side. At its most negative, in 1967, a shift towards conservatism in the House was associated with an even greater in magnitude *drop* in defense spending.

https://doi.org/10.1017/S0143814X17000186 Published online by Cambridge University Press

The timing of these effect estimates is noteworthy. The most negative estimate falls in 1967, following the major escalation of US involvement in Vietnam after the Gulf of Tonkin Resolution, whereas the most positive estimate falls in 2004, as the depth of US involvement required in Iraq was becoming clear the year after declaring victory. At these times of major foreign escalations, the partisanship of the House of Representatives is likely most important to presidents' ability to mobilise resources for increases in the defense budget. After all, the late 1960s and the early 2000s were both periods of increasing defense budgets to accommodate expanding conflicts and they were also periods in which the presidents could rely on copartisan majorities in the House. Such partisan dynamics are in line with recent qualitative, comparative studies (e.g. Mortensen et al. 2011), but have not gained much scholarly attention in time series analyses of public policy and public spending.

Fourth, the effect of the lagged national unemployment rate also merits attention. This effect estimate shows evidence of minor gradual changes in effect, going from a significant relationship to defense spending until the period of the Reagan buildup, after which it no longer shows a statistical relationship to changes in defense spending. Thus, our model indicates that the traditional expectation that increasing defense spending is a way to offset unemployment finds support throughout most of our time series up until the last year of the Reagan buildup. However, after this point, we no longer find a statistically significant relationship. This finding is a strong candidate for a case in which politicians' core beliefs about policy have evolved. By the time of the defense buildup in the first part of the new century, several trends made defense spending a less useful tool for addressing unemployment: responding to international terrorism lent itself less to massive buildups of troops and bases, and sizable portions of defense spending were outsourced to private contracts. Thus, the logic underlying the former connection between unemployment and defense spending altered in the mid-1980s and seemingly has not returned.

Finally, public sentiment that defense is an important problem shows no association to defense spending changes, being consistently statistically insignificant and near zero. Likewise, presidential election years show no relationship to the outcome either.

DLM accuracy using simulated data

The model description demonstrates the DLM is capable of estimating coefficients that change with time, and our empirical examples show how these models can help us learn about policy-making trends. Nevertheless, it is another matter to ask whether the dynamic estimates from a DLM are necessarily the "true" coefficients we would like to recover from the estimation. We examine this latter, more crucial, question by simulating a time series of 41 observations on which to test our model. The chosen number of observations reflects the length of our actual time series used to generate Figure 1. We simulate a mix of six dynamic and constant true effect parameters and use them to calculate a simulated outcome variable that is a function of six randomly simulated explanatory variables and a normally distributed error term.

The design of the simulation reflects different types of time-varying relationships implied by policy theories. While the framework of punctuated equilibrium theory claims that changes can be sudden and quite dramatic, the advocacy coalition framework depicts more gradual changes, and a similar shape of change would be expected from the social learning perspective. In addition, it is important to validate that the DLM can in fact return constant effects where appropriate. Thus, two of the six factors are related to the outcome via constants (Beta 1 and Beta 5). One is constant except for a single large shift from a negative to a positive effect (Beta 6), which is chosen in order to examine whether the approach can pick up a major and sharp punctuation in the relationship between the explanatory variable and the outcome variable. One is related to the outcome via a coefficient that varies randomly according to a normal distribution over time but has no systematic pattern (Beta 4). Finally, two of these factors are related to the outcome via effect coefficients exhibiting distinct, gradual temporal patterns (Beta 2 and Beta 3).

The factors themselves, our simulated explanatory variables, are a matrix of six random vectors drawn from standard normal distributions. The outcome variable is calculated deterministically from the data and the true effects we wish to recover, after which we add a realisation of mean-zero normally distributed random noise to each observation. We exclude an intercept term from the model and, prior to running the model, we standardise the outcome by subtracting its mean and dividing by its standard deviation. We do not transform the explanatory variables, as they are random draws from standard normal distributions.

Figure 3 shows the results of running our DLM on this simulated data. The solid lines plot the true coefficients for our model and the dashed lines surrounded by grey confidence intervals plot our estimates of them. The model is quite successful at tracking how coefficients evolve over time, in the case of the two dynamic coefficients (Beta 2 and Beta 3), and it filters out noise to estimate Beta 4 as roughly constant. Beta 1 and Beta 5 are estimated accurately as constants, though the estimate for Beta 1 is biased slightly downward. Finally, the model converges to a roughly correct estimate of Beta 6, though it estimates the last few observations somewhat too conservatively.

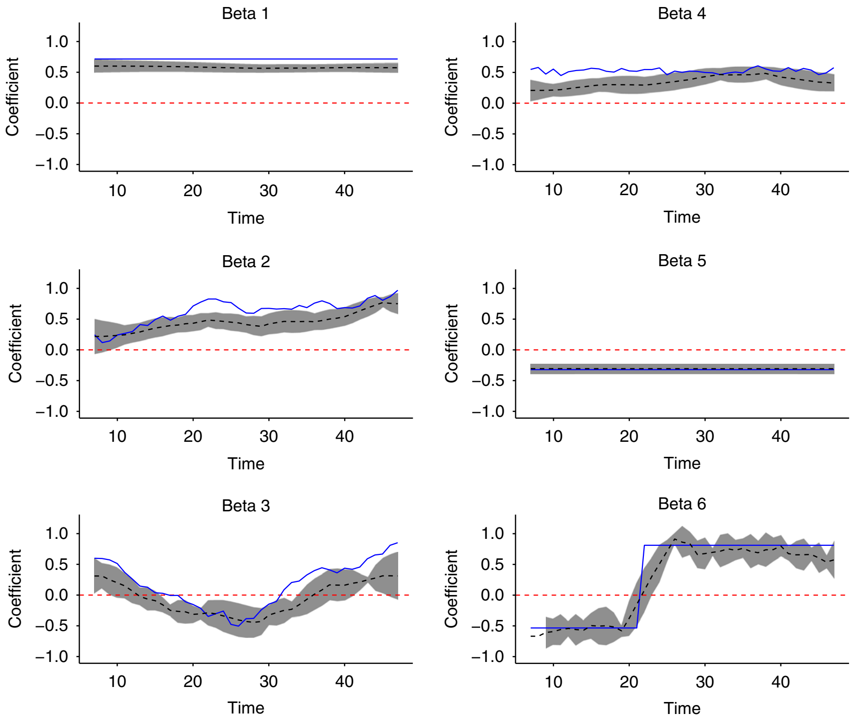

Figure 3 Estimated and true time-varying coefficients from simulated data.
Note: The solid line is the known value of the coefficient. The dashed lines plot central predictions of dynamic effect coefficients, and the grey regions are 95% confidence intervals around those predictions.

We note that repeating this test with all coefficients varying dynamically, or none varying dynamically, does not substantively change the accuracy. Maintaining the same patterns of dynamics and the same number of independent variables, we can also demonstrate that changing the length of the time series has relatively little impact on model performance. We refer the reader to the Online Appendix G for graphs of these models in which we vary the length of the time series from a low of 30 observations to a high of 100 observations. Results are virtually identical to those reported here, though confidence intervals shrink with longer time series.

These are important evidence demonstrating that the DLM modelling approach is sensitive to various types of transformative changes, both the gradual movement predicted by the social learning and advocacy coalition framework, as well as the more punctuated changes of main interest to the punctuated equilibrium theory. It should be noted, however, that these findings do not amount to conclusive evidence that the DLM will recover the true coefficients under all circumstances, only that the DLM can recover true estimates under conditions similar to our own empirical tests. We encourage the reader to perform simulations emulating the patterns of variation in their own data when applying the DLM. To aid in this, we provide example code in the Online Appendix H for creating simulations like our own using the R statistical software environment and the "dlm" package (Petris 2010).

Extending DLMs to panel data

Sometimes policy scholars utilise panel data, for instance, several countries observed over time, to evaluate dynamic policy theories. Extending DLMs using panel data can be accomplished in several ways. Two common choices are to model the panel time series in the DLM framework either as seemingly unrelated time series equations (SUTSE) or by using a hierarchical model with levels accommodating panel variation. In the former approach, the model structure is virtually the same as above except that errors in both the measurement and state equations are assumed correlated across panel units within time periods (Petris et al. 2009, 132–134). This amounts to a partial relaxation of the assumption that all errors are independent. Errors in a SUTSE DLM model take a block diagonal form, in which errors within a period across units are correlated and all other errors are independent. It can, however, become computationally demanding to estimate the large covariance matrices that result when the number of panel units is large.

In such cases, an alternative is to estimate a dynamic hierarchical model. In this framework, the state equation has (at least) two levels. The dynamic coefficient estimates (in the measurement equation) are estimated as realisations from an underlying state (first state level), and the state varies as a Markov chain over time (second state level) (Petris et al. 2009, 134–136). This introduces dependence across panel units, which also varies dynamically, without requiring estimation of large covariance matrices. In such a framework, with panel units denoted by j, a basic model would be:

$$\begin{aligned} y_{jt} &= \boldsymbol{X}_{jt}\boldsymbol{\beta}_{jt} + v_{jt}, \quad & v_t &\sim i.i.d.(0, \sigma^2_{jt}) \\ \boldsymbol{\beta}_{jt} &= \boldsymbol{\lambda}_t + \epsilon_{jt}, \quad & \epsilon_{jt} &\sim i.i.d.(0, \Sigma_t) \\ \boldsymbol{\lambda}_t &= \boldsymbol{\lambda}_{t-1} + \boldsymbol{w}_t, \quad & \boldsymbol{w}_t &\sim i.i.d.(0, Q) \end{aligned}$$

It remains, as Commandeur et al. (2011) noted, that, at the time of writing, not all software packages for estimating DLMs support every form of DLM. Readers who find their preferred software does not include commands for easily estimating dynamic hierarchical or seemingly unrelated regressions can find this functionality in the “dlm” package for R. Similarly, readers who wish to estimate dynamic nonlinear regressions will find that, at the time of writing, “dlm” does not provide commands for this while MATLAB, RATS or the “KFAS” R package do provide such tools.

Conclusion

Theories of public policy imply that the causes of public policy may not be consistent across time. These time-varying relationships represent a major challenge to standard regression techniques, as standard techniques assume that a change in the value of an explanatory variable, x, will produce a corresponding change in the value of a dependent variable, y, of the same magnitude and direction at all times. This lack of alignment between our policy theories, which are explicitly dynamic, and our methods of empirical analysis, which are largely static, has been forcefully pointed out by leading public policy scholars. These voices have advocated for increased use of alternative methods such as process tracing and stochastic process analyses (Jones and Baumgartner 2005; Hall 2006).

Although we applaud this expansion of methods within the field of public policy research, this article has advocated a third alternative: better aligning statistical time series analysis of public policy with the ontology of the major theories in the field. More particularly, we have advocated the DLM – a flexible approach to time series analysis that can uncover both time-varying and constant relationships. Compared to alternative dynamic statistical methods utilised in public policy research, DLMs require fewer assumptions from the analyst, allowing the data to speak for itself to the greatest extent regarding the timing, direction and magnitude of effect changes. This is a major advantage over alternative methods, given the lack of specificity regarding the timing and particular conditions of change in public policy theories.

Another advantage is the fact that DLMs utilise all the information in the time series data to model policy outcomes in a way that closely maps onto the data-generating process envisioned by our theories of public policy. Given the multiple causes of public policies, a further advantage of the modelling approach advocated in this article is the ability to estimate multivariate models. Furthermore, because DLMs provide information about both stable and time-varying coefficients, they can be used to investigate the general expectation derived from all of the major policy theories that relationships are time-varying. This is clearly illustrated in our example analysis of US defense spending, where a mix of constant, dynamic and nonsystematic effects are present.

Applying the DLM to policy analysis opens up a range of new avenues for public policy research. First, whereas much quantitative policy research has been focused on theorising and estimating stable relationships, applying DLMs as a tool invites a higher-order level of theorising and modelling aimed at explaining when and why the relationship between x and y may change. Inspired by Lieberman (2002), this latter type of change can be denoted *transformative change*. Given the flexibility of the DLM, it is also a tool that can aid in theory development. As we identify transformative changes in relationships between policy inputs and policy outputs, these findings – and the further puzzles they will present – will direct and inform refinements in our theoretical explanations of the policy process.

A related strand of research our work points to is the usefulness of applying DLMs on a larger scale to identify variables, or sets of variables, most prone to exhibiting time-varying relationships to policy outcomes. As DLMs are applied to more policy areas and/or to more countries, we hope this will reveal which types of variables – for instance, political, economic or institutional – exhibit more dynamic effects. Although it may remain futile to attempt to make point predictions about the exact form and timing of transformative changes, it may eventually be possible – and would certainly be useful for future theorising – to identify clusters of more or less stable causes of public policy. Developing a systematic understanding of when dynamic relationships are observed will help us refine our theories about how they arise in the first place.

A third extension would be to examine time series data that are closer to the actual policy-making process, such as agenda setting data covering congressional hearings, presidential speeches, law-making activities and other outcomes that are intermediate to the production of actual policy outputs. The fact that such data are now available for the US, as well as a number of other countries,[10] means that we can utilise DLMs to begin comparing transformative dynamics across different stages of the policy-making process and across political systems.

As our examples have illustrated, the DLM can be a powerful tool for studying dynamic relationships. We believe that applying DLMs more widely offers significant practical and theoretical advantages to students of policy change in the work of probing and expanding our leading explanations of stability and change in public policy.

Acknowledgements

The authors wish to thank Gregory McAvoy, participants at the Aarhus University, Department of Political Science, Public Administration

[10] See www.comparativeagendas.net.

workshop, and panel participants at the Comparative Agendas Conference 2015 and Midwest Political Science Association general meeting 2015.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S0143814X17000186

References

Baltagi B. (2011) *Econometrics*, 5th ed. Berlin: Springer Texts in Business and Economics.

Baumgartner F. R. and Jones B. D. (1993) *Agendas and Instability in American Politics*. Chicago, IL: University of Chicago Press.

Baumgartner F. R. and Jones B. D. (eds.) (2002) *Policy Dynamics*. Chicago: University Of Chicago Press.

Baumgartner F. R. and Jones B. D. (2015) *The Politics of Information: Problem Definition and the Course of Public Policy in America*. Chicago: University Of Chicago Press.

Baumgartner F. R., Jones B. D. and Mortensen P. B. (2014) Punctuated equilibrium theory: Explaining stability and change in public policymaking. In P. A. Sabatier and C. M. Weible (eds.), *Theories of the Policy Process*, 3rd ed. New York, NY: Westview Press, 59–103.

Beck N. (1983) Time-Varying Parameter Regression Models. *American Journal of Political Science* 27(3): 557–600.

Blyth M. (2013) Paradigms and Paradox: The Politics of Economic Ideas in Two Moments of Crisis. *Governance* 26(2): 197–215.

Brandt P. T., Williams J. T., Fordham B. O. and Pollins B. (2000) Dynamic Modeling for Persistent Event-Count Time Series. *American Journal of Political Science* 44(4): 823–843.

Cairney P. and Heikkila T. (2014) A Comparison of Theories of the Policy Process. In P. A. Sabatier and C. M. Weible (eds.), Theories of the Policy Process, 3rd ed. New York, NY: Westview Press. pp. 363–390.

Campbell J. L. (2002) Ideas, Politics, and Public Policy. *Annual Review of Sociology* 28(1): 21–38.

Caporin M. and McAleer M. (2013) Ten Things You Should Know About the Dynamic Conditional Correlation Representation. *Econometrics* 1(1): 115–126.

Commandeur J. J. F. and Koopman S. J. (2007) *An Introduction to State Space Time Series Analysis*. Oxford: Oxford University Press.

Commandeur J. J. F., Koopman S. J. and Ooms M. (2011) Statistical Software for State Space Methods. *Journal of Statistical Software* 41(1): 1–18.

Correa H. and Kim J.-W. (1992) A Causal Analysis of the Defense Expenditures of the USA and the USSR. *Journal of Peace Research* 29(2): 161–174.

Cusack T. R. (1992) On the Domestic Political-Economic Sources of American Military Spending. In A. Mintz (ed.) *The Political Economy of Military Spending in the United States*. London: Routledge. pp. 103–133.

Cusack T. R. and Ward M. D. (1981) Military Spending in the United States, Soviet Union, and the People's Republic of China. *The Journal of Conflict Resolution* 25(3): 429–469.

De Boef S. and Keele L. (2008) Taking Time Seriously. *American Journal of Political Science* 52(1): 184–200.

Engle R. (2002) Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* 20(3): 339–350.

Enns P. K. and McAvoy G. E. (2011) The Role of Partisanship in Aggregate Opinion. *Political Behavior* 34(4): 627–651.

Fariss C. J. (2014) Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability. *American Political Science Review* 108: 297–318.

Finseraas H. and Vernby K. (2011) What Parties are and What Parties Do: Partisanship and Welfare State Reform in an Era of Austerity. *Socio-Economic Review* 9(4): 613–638.

Gerber A. and Green D. P. (1998) Rational Learning and Partisan Attitudes. *American Journal of Political Science* 42(3): 794–818.

Green D. P., Gerber A. S. and de Boef S. L. (1999) Tracking Opinion Over Time: A Method for Reducing Sampling Error. *The Public Opinion Quarterly* 63(2): 178–192.

Griffin L. J., Wallace M. and Devine J. (1982) The Political Economy of Military Spending: Evidence from the United States. *Cambridge Journal of Economics* 6(1): 1–14.

Hall P. A. (1993) Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain. *Comparative Politics* 25: 275–296.

Hall P. A. (2003) Aligning Ontology and Methodology in Comparative Research. Comparative Historical Analysis in the Social Sciences. Chapter 11. Cambridge and New York: Cambridge University Press. pp. 373–404.

Hall P. A. (2006) Systematic Process Analysis: When and How to Use it. *European Management Review* 3(1): 24–31.

Hartley T. and Russett B. (1992) Public Opinion and the Common Defense: Who Governs Military Spending in the United States? *The American Political Science Review* 86(4): 905–915.

Higgs R. and Kilduff A. (1993) Public Opinion: A Powerful Predictor of U.S. Defense Spending. *Defence Economics* 4: 227–238.

Hogan J. and Howlett M. (eds.) (2015) *Policy Paradigms in Theory and Practice. Discourses, Ideas and Anomalies in Public Policy Dynamics*. London: Palgrave Macmillan.

Huber E. and Stephens J. D. (2001) Welfare State and Employment Regimes in the Era of Retrenchment. In P. Pierson (ed.) *The New Politics of the Welfare State*. Oxford: Oxford University Press, 107–145.

Jenkins-Smith H. C., Nohrstedt D., Weible C. M. and Sabatier P. A. (2014) The Advocacy Coalition Framework: Foundations, Evolution, and Ongoing Research. In P. A. Sabatier and C. M. Weible (eds.), *Theories of the Policy Process*, 3rd ed. New York, NY: Westview Press. pp. 183–223.

Jones B. D. (1994) *Reconceiving Decision-Making in Democratic Politics: Attention, Choice, and Public Policy*. Chicago: University Of Chicago Press.

Jones B. D. and Baumgartner F. R. (2005) *The Politics of Attention. How Government Prioritizes Problems*. Chicago, IL: University of Chicago Press.

Kalaba R. and Tesfatsion L. (1988) The Flexible Least Squares Approach to Time-Varying Linear Regression. *Journal of Economic Dynamics and Control* 12: 43–48.

Kamlet M. S. and Mowery D. C. (1987) Influences on Executive and Congressional Budgetary Priorities, 1955-1981. *The American Political Science Review* 81(1): 155–178.

Kiewiet D. R. and McCubbins M. D. (1991) *The Logic of Delegation: Congressional Parties and the Appropriations Process*. Chicago: University Of Chicago Press.

Kim C.-J. and Nelson C. R. (2000) *State-Space Models with Regime Switching*. London: The MIT Press.

Kwon H. Y. and Pontusson J. (2010) Globalization, Labour Power and Partisan Politics Revisited. *Socio-Economic Review* 8(2): 251–281.

Lebo M. J. and Box-Steffensmeier J. M. (2008) Dynamic Conditional Correlations in Political Science. *American Journal of Political Science* 52(3): 688–704.

Lieberman R. C. (2002) Ideas, Institutions, and Political Order: Explaining Political Change. *The American Political Science Review* 96(4): 697–712.
Majeski S. J. (1983) Mathematical Models of the U.S. Military Expenditure Decision-Making Process. *American Journal of Political Science* 27(3): 485–514.
Majeski S. J. (1992) Defense Budgeting, Fiscal Policy, and Economic Performance. In A. Mintz (ed.), *The Political Economy of Military Spending in the United States*. London: Routledge. pp. 217–237.
Mcavoy G. E. (2006) Stability and Change: The Time Varying Impact of Economic and Foreign Policy Evaluations on Presidential Approval. *Political Research Quarterly* 59(1): 71–83.
Mergner S. (2009) *Applications of State Space Models in Finance*. Göttingen, Germany: Universittsverlag Göttingen.
Mitchell S. M., Gates S. and Hegre H. (1999) Evolution in Democracy-War Dynamics. *Journal of Conflict Resolution* 43(6): 771–792.
Montana G., Triantafyllopoulos K. and Tsagaris T. (2009) Flexible Least Squares for Temporal Data Mining and Statistical Arbitrage. *Expert Systems WITH APPLICATIONS* 36(2, Part 2): 2819–2830.
Mortensen P. B., Green-Pedersen C., Breeman G., Chaqus-Bonafont L., Jennings W., John P., Palau A. M. and Timmermans A. (2011) Comparing Government Agendas: Executive Speeches in the Netherlands, United Kingdom, and Denmark. *Comparative Political Studies* 44(8): 973–1000.
Nincic M. and Cusack T. R. (1979) The Political Economy of US Military Spending. *Journal of Peace Research* 16(2): 101–115.
Ostrom C. W. Jr. (1977) Evaluating Alternative Foreign Policy Decision-Making Models: An Empirical Test Between an Arms Race Model and an Organizational Politics Model. *The Journal of Conflict Resolution* 21(2): 235–266.
Ostrom C. W. Jr. (1978) A Reactive Linkage Model of the U.S. Defense Expenditure Policy-making Process. *The American Political Science Review* 72(3): 941–957.
Ostrom C. W. Jr. and Marra R. F. (1986) U.S. Defense Spending and the Soviet Estimate. *The American Political Science Review* 80(3): 819–842.
Palmer G., D'Orazio V., Kenwick M. and Lane M. (2015) The MID4 Dataset, 20022010: Procedures, Coding Rules and Description. *Conflict Management and Peace Science* 32(2): 222–242.
Petris G. (2010) An R Package for Dynamic Linear Models. *Journal of Statistical Software* 36(12): 1–16.
Petris G. and Petrone S. (2011) State Space Models in R. *Journal of Statistical Software* 41(4): 1–25.
Petris G., Petrone S. and Campagnoli P. (2009) *Dynamic Linear Models with R*. New York, NY: Springer.
Pierson P. (1996) The New Politics of the Welfare State. *World Politics* 48: 143–179.
Pierson P. (2000) Path Dependence, Increasing Returns, and the Study of Politics. *The American Political Science Review* 94(2): 251–267.
Princen S. and 't Hart P. (2014) Putting Policy Paradigms in Their Place. *Journal of European Public Policy* 21(3): 470–474.
Rosenthal H. and Poole K. (2015) House Polarization 1st to 113th Congresses, ftp://voteview.com/house-polarization46-113.txt (accessed 13 July 2015).
Sabatier P. (1987) Knowledge, Policy-Oriented Learning, and Policy Change. *Knowledge* 8: 649–692.
Sabatier P. (1988) An Advocacy Coalition Framework of Policy Change and the Role of Policy-Oriented Learning Therein. *Policy Sciences* 21: 129–168.

Sabatier P. A. (1998) The Advocacy Coalition Framework: Revisions and Relevance for Europe. *Journal of European Public Policy* 5(1): 98–130.

Sabatier P. and Jenkins-Smith H. (eds.) (1993) *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder, CO: Westview Press.

Shumway R. H. and Stoffer D. S. (2010) *Time Series Analysis and its Applications: With R Examples*. New York: Springer.

Steinmo S., Thelen K. and Longstreth F. (eds.) (1992) *Structuring Politics: Historical Institutionalism in Comparative Analysis*. Cambridge: Cambridge University Press.

Su T.-T., Kamlet M. S. and Mowery D. C. (1993) Modeling U.S. Budgetary and Fiscal Policy Outcomes: A Disaggregated, Systemwide Perspective. *American Journal of Political Science* 37(1): 213–245.

True J. L. (2002) The Changing Focus of National Security Policy. In F. R. Baumgartner and B. D. Jones (eds.), *Policy Dynamics*. Chicago, IL: University of Chicago Press, 155–183.

Tusell F. (2011) Kalman Filtering in R. *Journal of Statistical Software* 39(2): 1–27.

Weible C. M., Sabatier P. A., Jenkins-Smith H. C., Nohrstedt D., Henry A. D. and deLeon P. (2011) A Quarter Century of the Advocacy Coalition Framework: An Introduction to the Special Issue. *The Policy Studies Journal* 39(3): 349–360.

Weible C. M., Sabatier P. A. and McQueen K. (2009) Themes and Variations: Taking Stock of the Advocacy Coalition Framework. *The Policy Studies Journal* 37(1): 121–140.

Whitten G. D. and Williams L. K. (2011) Buttery Guns and Welfare Hawks: The Politics of Defense Spending in Advanced Industrial Democracies. *American Journal of Political Science* 55(1): 117–134.

Wlezien C. (1996) Dynamics of Representation: The Case of US Spending on Defence. *British Journal of Political Science* 26(1): 81–103.

Wood B. D. (2000) Weak Theories and Parameter Instability: Using Flexible Least Squares to Take Time Varying Relationships Seriously. *American Journal of Political Science* 44(3): 603–618.

Wood B. D. and Doan A. (2003) The Politics of Problem Definition: Applying and Testing Threshold Models. *American Political Science Review* 47(4): 640–653.

Zuk G. and Woodbury N. R. (1986) U.S. Defense Spending, Electoral Cycles, and Soviet-American Relations. *The Journal of Conflict Resolution* 30(3): 445–468.