

Can the Federal Budget Process Promote Fiscal Sustainability? Evidence from Synthetic Control

By Denvil R. Duncan, Justin M. Ross, and John L. Mikesell
The Paul H. O'Neill School of Public & Environmental Affairs, Indiana University

Working Paper
November 26, 2019

Abstract

Can the federal budget be constrained using congressionally self-imposed rules? The Budgetary Enforcement Act of 1990 (BEA90), in effect from 1992 to 2002, is frequently held up as a political control that largely succeeded in doing so and its expiration is considered a watershed event. This paper finds evidence supportive evidence of BEA90 as a constraint on non-defense discretionary outlays from a synthetic control using Lasso regression methodology. The effect of BEA90 is estimated to be around 10 percent, but the sampling variation suggests that the probability these results arose from chance to be around 30 percent. Confronted with a common applied empirical problem of having a large number of potential predictors (over 2,000) and a relatively short time-series (12 years), we propose novel additional checks that successfully cross-validate the synthetic control's counterfactual model.

Keywords: Federal budget, budget process

JEL Code: H6, H1

Acknowledgments: We would like to thank Phil Joyce, Bob Kravchuk, James Savage, Coady Wing, Alex Hollingsworth, Bob Kravchuk, participants at the 2018 Association of Public Budgeting and Finance meetings, and participants of the Public Finance Conference on "Federal, State, and Local Budgets in Crisis" at Indiana University's School of Public and Environmental Affairs.

1. Introduction

The American federal budget process is generally regarded as a collective negotiation between the major political parties in both congressional chambers and the executive branch with an aim towards prioritizing individual public programs within aggregate fiscal constraints. To be successful, an institutional process like budgeting must be capable of coordinating diverse and changing actors in a way that is robust to environmental conditions in a manner that affects relevant outcomes. The long-term fiscal sustainability of America's federal government will ultimately depend on some degree of fiscal restraint, even if not balanced budgets, and observers of the budget process have been critical of federal budgeting and budgeting theory in just about every conceivable respect. For decades scholars have raised concerns that the budget process, despite several reforms and adaptations, is increasingly deficient by way of systematic deficits, shrinking control of appropriate levers for correction, and increased political polarization. This long-running criticism is perhaps underscored most prominently by the widely decried fact that in recent years Congress has regularly failed to pass its own budget resolution, failed to pass appropriation laws before the beginning of fiscal years, and cannot avoid deficits even in a favorable economic climate.

In the history of various revisions and amendments to the federal budget process made by congress, one effort has generally stood out as an example where congress succeeded in the design principles.¹ The Budget Enforcement Act of 1990 (BEA90), which was in effect during the fiscal years 1992 to 2002, required the executive and legislative branches to agree on a cap on discretionary spending growth and allowed for only deficit-neutral or deficit-reducing rule changes to entitlements and taxes. **Table 1** provides the spending limits under the BEA90 reign, which while exceeded more often than not, BEA90 is generally regarded as having aided in bringing about budget surpluses in fiscal 1998 through 2001 by encouraging more restraint than there might have otherwise been. Its expiration permitted the passage of the tax cuts under the Economic Growth and Tax Relief Reconciliation Act of 2001 and the mandatory spending growth in the Medicare Modernization Act of 2003, each of which (good policy or not) were

¹ Savage and Verdun (2008) argue that BEA90 reflected lessons learned by Congress from the experience under the Balanced Budget and Emergency Deficit Control Act, also known as Gramm-Rudman-Hollings (GRH), which created deep partisan distrust. Congress used BEA90 to control federal spending rather than the budget deficit, close loopholes in GRH, and establishing a process to control mandatory entitlement growth.

significant contributors to the deficit.² **Figure 1** illustrates some visual support for this perspective, where it can be observed that real discretionary outlays actually decreased during BEA90's era. **Figure 2** further demonstrates the break in real discretionary federal outlays by overlaying the pre-BEA90 trend over the BEA90's effective and expiration range, demonstrating a significant decline in discretionary outlays.

What is unclear, however, is how causal the BEA90 actually was to congressional restraint. Since congress passed BEA90, renewed it, and ultimately allowed it to expire, it is more than plausible that the act was simply a symptom of a preference for budget control and that the actual features of the policy did not matter. Nevertheless, it is also entirely possible that congress can devise means in the budget process that would raise the political cost of discretionary spending. It is this subject of this paper to investigate the case of BEA90 as a causal source of fiscal restraint.

To investigate the effect of BEA90, we adopt the synthetic control method to determine if the expiration of BEA90 led to an increase in non-defense discretionary expenditures from 2003 to 2006. Specifically, we employ the Synthetic Control Using Lasso (SCUL) regression that allows us to form a synthetic United States even when there are more candidate predictors than there are observations using a data driven approach to model selection (Hollingsworth and Wing, 2018; Mikesell and Ross, 2019). Under this causal investigative technique, it would seem BEA90's expiration led to an immediate and persistent leveling up of federal discretionary expenditures. However, conventional placebo tests for treatment randomization error suggest a large amount of uncertainty over this effect. To the best of our knowledge, this paper represents the first application of synthetic control methods in the study of federal public budgeting. Furthermore, because our paper has thousands of candidate donors to predict a relatively short time horizon (i.e. the ten years in of BEA90 before expiration), our paper performs a novel test of model selection performance to support the model's finding in the presence of a large scale data mining exercise.

After introducing the federal budget process and major historical acts that have influenced it, the next section discusses how this paper ties together the disparate elements of

² The Medicare Modernization Act of 2003 included Medicare Part D, which expanded entitlements in the form of subsidies to drug prescriptions.

budgeting theory and fiscal politics under the overarching public choice view of politics as exchange. Section 3 introduces the empirical strategy, while section 4 provides the results before concluding.

2. The Federal Budget: Process and Theory

Scholarship on the American federal budget process faces two problems which has directed empirical research away from consideration of fiscal sustainability; 1) the uniqueness of America and its federal budgeting system; 2) the limited predictive power of most budget theory - when it is applied to the American federal government in recent decades - as well as the difficulty of providing a broader theory of federal budgeting. To provide the necessary context, this section will proceed with an overview of the process and its major changes over time then proceed to the discussion of previous scholarship in budget theory.

Overview of the Process and its Evolution

The budget process provides the framework for the President and Congress to formulate, approve, and execute the expenditure programs of the federal government. While revenue issues are sometimes involved in the process, the primary focus of the process is on spending because major revenue proposals and their adoption operate through a different decision structure, i.e., different Congressional committees and not according to a regular cycle. The budget process establishes how proposals for expenditure will be developed and approved and verifies that spending has occurred according to the approved law. The process intends to provide a discipline that keeps overall spending within the bounds of available revenue (a fiscal sustainability concern), that allocates resources among possible uses to services most important to the nation, and that encourages operating units of government to manage resources they have been provided in the most efficient means possible (Mikesell 2014: 55 – 59). It requires the President and Congress to perform certain defined budgetary tasks in a regular sequence if it is to function properly.

The most basic element is in regard to control of expenditure. The federal Constitution states that “No Money shall be drawn from the Treasury, but in Consequence of Appropriations made by Law...” (Article 1, Section 9), so expenditures must be approved in a law passed by Congress and signed by the President. The process for making these appropriations is generally

the same as required for passage of any law: approval of an identical bill by both houses of Congress and signature of the bill into law by the President. This is the flow identified as starting in the appropriation committees in **Figure 3**.

The appropriation process, however, represents approval of agency expenditure programs. It does not directly involve the development of fiscal plans that budgeting entails. Two important laws have created the current federal budget process that creates integrated fiscal plans: the Budget and Accounting Act of 1921 and the Congressional Budget and Impoundment Control Act of 1974. Although enactment of each law was driven by multiple intentions, one element in each involved a desire to get budgetary deficits under control.

The Budget and Accounting Act of 1921 created the formal Presidential budget process. It required the President to present to Congress early in the calendar year an intended expenditure program for all operations of the federal government for the fiscal year that will begin several months in the future. No more would agencies bring their requests for appropriations directly to Congress when more resources were required. The President, assisted by a Bureau of the Budget created by that Act (now Office of Management and Budget), would present a single, government-wide request to Congress (a flow shown on **Figure 3**, developed according to his policy intentions and within the resources he believed to be available for that upcoming year. While the Presidential budget system was an element in an overall management agenda, it was also a product of concern about fiscal discipline. The expenses of World War I had caused deficits far in excess of experience in the early years of the twentieth century (deficits of 44 percent of outlays in 1917, 71 percent of outlays in 1918, and 72 percent of outlays in 1919, compared with surpluses in twelve of the prior years of the century) and there was much concern that there be a return to disciplined finances. An executive budget was seen as a device for re-establishing control – and the first nine years covered by the Presidential budget requirement (1922 to 1930) showed a surplus in excess of twenty percent of total outlays. The Presidential budget, as a unified and organized proposition for federal outlays developed from constrained agency requests, continues as an entrenched element of the federal fiscal process. As an executive plan, the Presidential budget provided no actual spending authority to federal agencies.

The process functioned with only a Presidential budget proposal for almost fifty years, through the fiscal stress coming from the Great Depression, World War II, and the Cold War. In non-shooting-war periods, there were years of surplus and years of deficit with deficits frequently in recession years. But from 1961, the pattern changed: from 1961 through 1973, there was only one surplus year, even though the federal government faced no extraordinary fiscal pressures in that time (the Vietnam War was not associated with the substantial run-up in defense spending traditionally associated with military action). This history was one of the major forces behind passage of the Congressional Budget and Impoundment Control Act of 1974 (Public Law 93 – 344).

The Congressional Budget and Impoundment Control Act, among other things, required Congress to pass a Congressional Budget Resolution, developed in Budget Committees (shown in Figure 1), that establishes the fiscal program of Congress.³ Key elements to be set forth in the resolution for the upcoming fiscal year included appropriate levels of outlays and new budget authority, an estimate of budget outlays and new budget authority for each major functional category of government operations, the surplus or deficit appropriate in light of economic conditions and other factors, recommended level of revenues and the aggregate amount by which revenue should increase or decrease, and the appropriate level of public debt and any appropriate change in the statutory debt limit. Hence, the Congressional Budget Resolution provided the Congressional guideline for the budget process in much the same way as the Presidential budget provided the executive branch view. The Resolution would be passed by Congress but would not be signed by the President, would guide the fiscal decisions by Congressional committees and the full Congress made on expenditures (and revenues) (shown in **Figure 3**), but would provide no obligation authority to agencies. Although Congress regularly passed its Budget Resolution in the early years of the law, as **Table 1** shows, a Congressional Budget has been passed for only eight fiscal years since 2000.

The two budget process laws created Presidential and Congressional responsibility for development of fiscally constrained spending programs. However, the fiscal discipline record continued problematic. From fiscal 1977 (the first year with both Presidential and Congressional

³ The Act also provided Congressional committees to consider the Budget Resolution, a Congressional Budget Office to provide non-partisan fiscal and economic analysis for Congress, and constrained the ability of the President to fail to spend all money provided in an appropriation act.

budgets) through the first half of the 1980s, the deficit was less than ten percent of total outlays in only one year and was over twenty percent in 1983, 1984, 1985, and 1986. At this point, lawmakers moved away from efforts to improve fiscal discipline through a strengthened budget process and proceeded to craft direct control mechanisms with triggers and consequences. Several control acts were passed to impose direct control over deficits or to cap outlays. These are included as part of **Table 1** and include the following:

- i) Balanced Budget and Emergency Deficit Control Act of 1985 (Public Law 99 – 177) [Gramm-Rudman-Hollings]. The Act established deficit targets for future years, ultimately leading to budgetary balance, with a formula sequestration process if the appropriation process appeared to be producing a fiscal outcome that would violate the target for a year. The law was applicable for fiscal 1986 through 1990, although its application was sometimes suspended.
- ii) Budget Enforcement Act of 1990 (Public Law 101 – 508). The Act divided spending into discretionary (spending controlled by the annual appropriation process) and mandatory (spending controlled by a formula) and controlled each. It established firm ceilings on discretionary spending and a PAYGO requirement on mandatory spending and revenue provisions (PAYGO meaning that any legislative change that would increase mandatory spending or reduce revenue would have to be accompanied by a provision to pay for the change in the deficit). The law and its extensions applied to fiscal year 1992 – 2002.
- iii) Statutory Pay-As-You-Go Act of 2010 (Public Law 111-139) The Act requires that new legislation changing taxes, fees, or mandatory spending, taken together, may not increase the projected deficit. OMB is the scorekeeper and applies across-the-board (with exceptions) sequestration of mandatory spending if there is a violation. According to the scoring rules of the law, no sequestration has yet been found to be necessary. The law applies to all legislation enacted after February 12, 2010.
- iv) Budget Control Act of 2011 (Public Law 112 – 25) The Act establishes discretionary spending caps through fiscal 2021. Sequestration across-the-board applied by OMB if the caps are violated. The law applies to fiscal 2012 through fiscal 2021 and has not been repealed.

- v) Bipartisan Budget Act of 2013 (Public Law 113 – 67). In response to fears that cuts from the prior act (which remained in place) did not accommodate national security, inflation, an aging population, and infrastructure spending, the Bipartisan Budget Act of 2013 produced a deal to increase spending caps for fiscal years 2014 and 2015.
- vi) Bipartisan Budget Act of 2015 (Public Law 114 – 74). This law permitted a similar agreement to increase spending beyond the limits of the 2011 law (which remained on the books) for fiscal years 2016 and 2017.

In terms of added fiscal discipline, these extraordinary targets and controls do not show a strong record of success. From fiscal 1986 through 2017, only four fiscal years – 1998, 1999, 2000, and 2001 -- have shown a surplus, with only one of those years with a surplus above ten percent of outlays. This record of deficits is particularly remarkable in that almost all years during that period were in the expansion phase of the business cycle, an environment in which revenues should be robust and demands for federal spending somewhat constrained. The rest of the years have shown deficits, twenty with deficits above ten percent of outlays. It is apparent in this record that it is easier to pass laws purported to control a budget deficit than it is to pass laws that actually reduce a deficit by increasing revenue or reducing spending. And, when there are controls in place, Congress can override those controls to spend more, as in the 2013 and 2015 acts.

Previous Scholarship

Much of the scholarly literature in public budgeting discusses the federal budget process since the 1970's as a departure from an incremental method. Incrementalism is the method of comparing proposed budgets to those of the previous period, with participants giving special attention to those components that are changing (Wildavsky, 1964; Wildavsky and Feno, 1966). Davis et. al. (1966), for instance, were able to show that from 1947 to 1963 congressional appropriations to various federal agencies were highly predictable from simple linear trend models.⁴ Though it was never without its critics (e.g. Schultze 1968; Natchez and Bupp, 1973; Berry, 1990), incrementalism is still the dominant theory of budget evaluation and processing in the sense that alternative theories are compared for how they departure from it or are otherwise

⁴ A large literature testing the incrementalism hypothesis during this era exists, a critical methodological summary and review of which can be found in Dezhbakhsh et al. (2003).

seeking to explain periods in which incrementalism is not followed. Like incrementalism, subsequent budgeting theories would emphasize the perceived new realities of the budget process by ways of its features (e.g. Bozeman and Straussman, 1982; Caiden, 1984; Pitsvada and Draper, 1984; Joyce and Reischauer, 1992; Rubin, 2007).⁵

By and large, these theories of budgeting are intended to describe the processes for the evaluation of programs. That is, these theories seek to provide observers with a means to understand the evaluation process as it progresses across political cycles, in the form of considering “program A” against “program B” whilst seemingly ignoring “program C”. Incrementalism, for instance, theorizes that programs will seldom be re-evaluated and political battles on programs will not be reopened.⁶ Bozeman and Straussman (1982) addressed inadequacies of incrementalism by arguing that a theory of federal budgeting revised from incrementalism must incorporate the existence of two processes, one that is “top-down” that sets the parameters as well as a “bottom-up” process of negotiation.

It is important to recognize the uniqueness of the American federal budget system as a contributing factor to a relatively small number of studies that consider fiscal sustainability. In terms of budgetary process, the federal government has a number of features that make budgetary and other fiscal comparisons with other countries difficult (Savage and Verdun, 2008). On top of the usual challenges of geography and economic size of the country that plagues cross-country research, the structure of government reflected in the budget process represents an important point of departure: The politically and fiscally independent tiers of government associated with the federal system; The separation of powers between legislative and executive branches that distinguish the American system from that of parliamentary governments and make the executive budget only the starting point for fiscal debate; The existence of two legislative assemblies (Senate and House) that must approve appropriations rather than only one; legislative freedom to add expenditures to the budget presented by the executive; The single-year emphasis in both budgeting and appropriating; Absence of a distinct capital / development

⁵ This shift in the literature’s attention is predicted by a literature review by Berry (1990: 167): “I argue that the most productive course would be to banish the term incrementalism from new scholarly literature, and instead, focus research on more specific characteristics of the budgetary process.”

⁶ This is a reasonable common-denominator view of what “incrementalism” means, but as Berry (1990) points out the term “incrementalism” has taken on many different theoretical meanings. \

budget, Absence of over-arching deficit-ratio control rules (e.g. the 3% of GDP EU control), and the substantial amount of expenditure that goes unchecked by the appropriation structure via the operation of entitlement formulae or through the tax expenditure system. Gathering all these features together makes panel data on the national budget outcomes as controls for operations of the U. S. federal government a difficult challenge. Furthermore, as documented by Calcagno and Lopez (2017), informal rules also play a role in congressional deliberations over federal finances.

Given the data challenges required for theory driven tests that relate the budget process to fiscal outcomes, fiscal sustainability of process has instead given way to a large literature of fiscal politics that offered more data and clearer path toward hypotheses testing. For instance, Albouy (2013) finds in a regression discontinuity design that districts represented by members of the majoritarian party with greater proposal power will receive a larger allocation of federal grants, and Berry et al. (2010) find similar evidence using data from the president's proposed budget; Larcinese et al. 2006 provide evidence that states receive greater federal funds when they were heavy supporters of the incumbent president in the last election, and when the governor of their state is from the same political affiliation; Anderson and Woon (2014) finds that, consistent with bargaining theory, delays in the delivery of the presidential budget advantages in the form of greater concessions from congress to the executive. In addition to showing that political elections encourage deficit spending, Shi and Svensson (2006) show that the size of this effect depends on politicians rents of remaining in power and in the share of uninformed voters in the electorate. Internationally, Roubini and Sachs (1989) and Edin and Ohlsson (1991) find that majority coalition governments have greater public debt to GDP ratios, while Woo (2003) provides a number of pooled cross-sectional regressions between public debt and various political policy process variables.

3. Empirical Strategy

To investigate the effectiveness of BEA90 as a fiscal constraint, we will focus on the period of 1992 to 2006 and on non-defense discretionary spending. That is, we will investigate whether BEA90's expiration in 2002 led to an increase in non-defense discretionary federal spending. The choice of this fiscal outcome is selected for two reasons. First, in 2002 US military efforts escalated with major foreign campaigns in the wake of the September 11, 2001 attacks, and so we regard the post-BEA period as inseparable from the post 9/11 period in terms of defense

spending. Indeed, the timing of 9/11 is arguably a threat to inference validity for non-discretionary spending, but we are on much sturdier ground for our inference. Secondly, as already discussed, BEA90 was partially allowed to expire in order to make it possible to pass the Medicare Modernization Act of 2003, which expanded mandatory spending. Thus it is not particularly interesting or informative to demonstrate that mandatory spending increased after BEA90 was passed.⁷

Historically, case studies like BEA90 have limited options for causal investigation in the empirical social sciences. Without comparable units and processes to the US federal government, the traditional approach is to perform a time-series econometrics modeling exercise, which is often considered a “descriptive” approach because it is a low-credibility design that requires strong assumptions over model specification and exogeneity. The more contemporary approach is to employ a synthetic control method, considered “state-of-the-art” in causal investigation of time-series data because it ultimately requires fewer assumptions that are basically analogous to the familiar assumptions of a difference-in-difference regression on panel data. In the context of this study, the synthetic control approach is to identify a set of alternative variables for other countries for which a linear combination is comparable to the US non-defense discretionary spending. This linear combination represents a “synthetic” United States (SUSA) which behaves similarly in a given policy era that can then serve the purpose of constructing a counterfactual in the new policy era. These other countries do not need to be similar to the U.S. in their federal government nor have a causal connection in the economic data used to predict the fiscal outcome of interest. The synthetic control is instead something akin to a data mining operation, where signals of US nondefense discretionary spending are separated from noise in the data so that the linear combination satisfies pre-trend assumptions in comparing the actual US trajectory to those of the synthetic US.

Lastly, we note that the 1992 to 2006 time period is selected because the synthetic control method is a pre-post design that relies on having enough pre-period data to generate a post-period counterfactual. Most pre-post designs examine a policy being implemented rather than disabled, but as described in section 2 the decade leading up to BEA90 included the enabling and expiration of the Budget and Emergency Deficit Control Act of 1985 (also known as Gramm-

⁷ However, we can demonstrate results to that effect with the Synthetic Control Method, available upon request.

Rudman-Hollings) to say nothing of the collapse of the Soviet Union and decline of the Cold War. There is no obvious stopping point to exploring the post period, but with each passing year there is reduced confidence that the pre-trends assumptions hold up for causal identification. Because 2007 saw the official beginning of the Great Recession associated with the housing and financial market crisis, we set the post-period counterfactual to end in 2007.

Synthetic Control Using Lasso Regression

The synthetic control method first pioneered by Abadie et. al. (2003, 2008) provided a strategy for estimating weights such that the weighted combination of countries creates a synthetic counterfactual history of the US. In the context of this paper, the intention of the synthetic control method is to establish a “synthetic USA” (SUSA) that can replicate the counterfactual path of the USA’s non-defense discretionary expenditures that would have been realized if BEA90 had been renewed rather than permitted to expire. This synthetic US is then compared to the actual US in order to identify the effect of BEA90’s expiration in 2002, with the key identification assumption being that this specific treatment effect was specific to the US. Again, the 9/11 attacks continue to provide a timing problem in that there is only one fiscal year of spending with BEA90 policy in effect after the attacks. To circumvent that concern, we (1) focus only on non-defense discretionary expenditures, and (2) estimate our synthetic control method to fit on the pre-treatment period one year earlier than the actual ending year so as to monitor the performance of the model before the actual treatment period. We follow the post-treatment period through 2006, allowing for four years of post-period analysis but ending before the Great Recession.

Abadie et al. (2010) provide a proof that if a synthetic counterpart can match the pre-treatment outcome trajectory of the actual unit, then the size of the bias in the post-treatment period caused by time varying unobserved confounders decreases to zero as the length of the pre-treatment period increases. The basic outline of a synthetic control method in this context is to define a set of weights on donor country covariates that creates a synthetic USA’s pre-treatment non-defense discretionary expenditures as precisely as possible. Whereas a traditional regression based approach seeks to include theory motivated controls that avoid bias from non-random assignment, the synthetic control approach uses a data driven process to determine what covariates are good at producing state weights that accurately predict outcomes across the pre-

treatment period. The synthetic control approach also carries the advantage of being a more generalizable approach to causal inference than a fixed-effect estimator because it allows for the unobserved fixed effects to have a time-varying influence on the outcome and nevertheless produce unbiased estimates of the treatment effect. The identification assumption is that the various weights that define the relationship between these predictors and the outcome of interest in the pre-treatment period remain constant in the post-treatment period, so that significant differences that emerge can be attributed to the newly introduced treatment (e.g., expiration of BEA90). This is not substantively different from the identification assumptions in a difference-in-difference framework, but carries the advantage that it can be performed with a small group of treated groups, even just one. This has been its primary appeal in causal investigative social science research, where previously only simple time series regressions were employable.⁸

We employ the synthetic control method first developed in Abadie et al. (2003, 2008) used to define a set of weights on covariates from donor states to produce a “synthetic USA.” Let y_{ijt} represent an observed fiscal indicator i in donor country j in year t and define USA nondefense discretionary spending to be y_{10t} . The donor pool provides a weighted sum of control series that will represent the synthetic USA. Using a $1 \times (1 + JP)$ vector of candidate donor pool indicators represented as $x_t = (1, y_{11t}, y_{21t}, \dots, y_{I1t}, y_{12t}, y_{22t}, \dots, y_{I1t}, \dots, y_{IJt}, \dots)$ in year t and $\beta = (\alpha, \theta_{11}, \theta_{21}, \dots, \theta_{I1}, \dots, \theta_{12}, \dots, \theta_{P2}, \dots, \theta_{PJ})^T$ be a $(1 + JP) \times 1$ vector of weights, then the synthetic USA time series is defined as:

$$(1) y_{10t}^* = x_t \beta = \alpha + \sum_{i=1}^I \sum_{j=1}^J y_{ijt} \theta_{ij}.$$

The synthetic control assumption is that if a combination of weights for donor units can be discovered which provide a close match between y_{10t} and y_{10t}^* for all pre-treatment periods, then this should provide an unbiased counterfactual set of outcomes over the post-treatment period for the treated unit (i.e. USA after BEA90 expires). Significant differences can be attributed to the treatment under the identification assumption that the various weights which define the

⁸ See Doudchenko and Imbens (2017) for a recent contrasting of Synthetic Control with other comparable regression techniques.

relationship between these predictors and the outcome of interest in the pre-treatment period remain constant in the post-treatment period.

In order to select a combination of weights there must be some computationally defined limit on their range of values. There are many recent extensions (e.g. Doudchenko and Imbens (2016), Xu (2017), and Dube and Zipperer (2015)) of the original synthetic control method set forward by Abadie et al. (2003, 2010), and this paper adopts the variant employed in Hollingsworth and Wing (2018).⁹ Using T_0 to represent the last year of the pre-treatment period, a Lasso regression is used to select weights which solve the following optimization problem:

$$(2) \arg \min(\alpha, \theta_{is}) \{ \sum_{t=1}^{T_0} (y_{00t} - x_t \beta)^2 + \lambda(|\beta|_1) \}$$

Equation (2) includes the usual regression squared prediction error, but adds to it an additional penalty in λ that increases with the sum of the absolute values of the weights defined by the expression $|\beta|_1 = \sum_{p=1}^P \sum_{s=1}^S |\theta_{ps}| + |\alpha|$. The lambda penalty rewards models that shrink coefficients to zero on predictors that offer little to no advantage for reducing squared prediction error, and allows the Lasso to estimate coefficients even when the number of independent variables exceeds the number of observations.

Whereas the original Abadie (2003, 2008) formulations restricted weights on donor predictors to non-negative values that summed to unity, the Hollingsworth and Wing (2018) synthetic control using Lasso (SCUL) approach is among the recent extensions (e.g. Doudchenko and Imbens, 2016) which allows weights be negative and does not impose the restriction that weights sum to unity.¹⁰ Consequently, instead of attempting different combinations of donors and judging the pre-treatment fit, the Lasso approach provides a systematic, automated, and data driven means of choosing predictors. This is desirable where reducing researcher judgement is valued (Belloni et al., 2014), but is further helpful to the case at hand because there is an abundance of annual federal public finance and economic data for many countries. Our donor predictor controls include nearly 3,000 potential predictors from IMF's Government Financial

⁹ The synthetic control using Lasso regression has been employed in the study of state taxation (see Mikesell and Ross, 2019) but we know of no other study using synthetic control of any variant for the study of the federal budget.

¹⁰ Since the size of the U.S. federal budget is an outlier in size at trillions of dollars, those restrictions would force a data driven process to select only those that are similar values rather than those that necessarily predict the fluctuations well.

Statistics data and the World Development Index. We also follow the Hollingsworth and Wing's (2018) recommendation to replicate this full exercise on the donor predictors for the common placebo test in determining a treatment randomized p-value, which differs from other approaches that apply the same user-selected model to the placebo units that was chosen for the case study of interest. Finally, as in Hollingsworth and Wing (2018) and Xu (2017) we adopt their recommendations for cross-validation, using the median lambda penalty arising from sub-sample pre-treatment tests with the best out-of-sample (but before the treatment period) prediction so as to better avoid overfitting the pre-treatment data.¹¹

To recap, the SCUL approach solves the problem of having many potential donor controls, many of which are likely not to be suitable informative controls. Rather than select from this list, using a Lasso regression to incorporate a penalty on the sum of the absolute values of the estimated regression shrinks large coefficients toward zero and may shrink smaller coefficients all the way to zero, providing a data-driven and automated means of discarding controls. Other approaches have users select co-variables and have the regression pick the donor countries that predict the pre-treatment period causing reliance on researcher judgment without the benefit of theory (Belloni et al., 2014), so the Lasso regression takes modeling flexibility out of the hands of the researcher. Furthermore, a placebo approach where the model provided by the synthetic control is employed on the non-treated units to provide information about the post-treatment prediction error is common practice but typically confronts the researcher with the problem that the best pool of candidate controls differs across placebo units. The SCUL approach, by contrast, selects new donor predictors for each placebo unit, which more closely replicates the exercise for donor selection in the original treated unit. Finally, we reduce concerns of overfitting historical data at the expense of out-of-sample performance by using a subsample cross validation procedure that chooses the penalty parameter controlling variable selection.

Donor Data and Calibration of SCUL Parameters

From IMF's Government Financial Statistics and the World Development Indicators, we have data available from 1992 to 2006 on 2,955 unique country-covariates. These co-variants

¹¹ To prevent overfitting over the pre-treatment period, the λ penalty parameter is chosen through a rolling K-fold cross-validation procedure where the Lasso regression is trained on subsets of the pre-treatment period, and the lambda penalty with the best average out-of-sample predictions across the datasets is adopted for the main results.

include different measures of public expenditures (e.g. general government consumption, military expenditures, tax revenue, etc.) and economic indicators (GDP, industry value-added, trade, unemployment, risk premium on lending, etc.). To employ cross-validation, we set the length of the first training run data to 1992 and 1993 followed by 5 years of test data. This resulted in four cross validation runs, where each run added a year of training data and scoring it over the subsequent five years of test data, with the fourth run training the model on 1992 to 1996 then scoring the model fit for its performance in predicting from 1997 to 2001. The model with the penalty from the Lasso regression with the best median mean-squared-error from across the four cross validation runs is then used in the Lasso regression to select country covariates to that predict the pre-treatment period of 1992 to 2001, stopping one year prior to the actual expiration of BEA90.

Ultimately, the Lambda penalty parameter adopted was 0.44 and the Lasso regression picked 16 country covariates to predict non-defense spending. **Table 2** provides the most significant predictors as measured by the partial r-square contribution. The largest contributing variable in 2006 was Jamaica's Central Government Debt, explaining 17.4 percent of the US nondefense discretionary spending.¹² Most of the remaining variation is explained by the top eight predictors whose share exceeds one percent, and interestingly it is not significant American trading partners like Canada or the UK selected by the Lasso regression.

4. Results

The main results appear in **Figure 4**, where the synthetic prediction of US federal non-defense discretionary spending from SCUL is plotted against the actual data, and numerical tabulations are reported in **Table 3**. The synthetic United States is close to ideal in that it is visually indistinguishable in **Figure 4** for the 1992 to 2001 period used by the Lasso regression to predict the actual data. If the pre-treatment period had fit poorly, we'd have little confidence in considering any deviations in the post-period. Further supportive is that the last year of BEA90 enforcement and first out-of-sample prediction from the HWSC method is fiscal year 2002. During this out-of-sample pre-trend prediction, there is a break in the trend where the actual US spending increased more rapidly than its immediate previous years trends, but this is caught by

¹² There is not much change across years in the rank ordering of these variables.

the synthetic control projection too as it missed by just 1.4 percent, lending support to the post-period predictive power of the synthetic projection. After BEA expires, there is a larger divergence between the synthetic and actual spending measures for 2003 and 2004 at 7.1 and 9.3 percent, respectively. However, the growth in the projected difference thereafter slows, rising to 11.5 percent in 2005 and 12.6 percent in 2006, suggesting that the two series were returning to the same trend pre-BEA90 trend for the last two years. This suggests that the expiration of BEA90 mostly ratcheted up the level initially but only slightly altered the trajectory of spending thereafter.

The placebo test, where for each potential donor covariate a new lambda penalty and set of predictors chosen by the Lasso regression is used to provide information on the uncertainty arising from a kind of randomization error. Intuitively, the placebo test gives us a sense, from the data at hand, how often we could have selected a unit-covariate and found a similarly sized treatment effect. The case study of interest, U.S. non-defense discretionary expenditures, had a pre-treatment Cohen D score of 0.011, so the placebo group was cut down to those whose Cohen D scores were higher than 0.02 so that we only considered cases where pre-treatment fit was similarly good; there are 1,944 cases that meet this criteria. While the exercise is not a perfect analogy to the p-values from t-tests used in regression models, they do provide a sense of the magnitude of the sampling variability of the data used to produce the results, and the frequency with which these cases provide an estimated treatment effect larger than the case-of-interest is referred to as the p-value.

Figure 6 plots the percent difference between the actual and synthetic prediction for all cases, as well as the U.S. case of interest and it can be seen that the U.S. runs fairly narrowly through the middle of the distribution. To get a better sense of this, each of the treatment years is examined in a kernel density plot in **Figure 7**, with a vertical dashed line to indicate the U.S. case of interest position. Of course, the highest density is centered around zero and generally the effect size drifts away but only very slightly, each case can be regarded as part of the main center of the distribution. Returning to **Table 4**, these percentage deviations for actual and synthetic are provided along with an accompanying p-value.¹³ The p-value in the treatment years was in the

¹³ P-values from placebo tests in synthetic control are constructed by finding the share of the placebo tests whose rank by absolute value in the distribution exceeds the case-of-interest.

range of 0.27 to 0.29, suggesting that more than a quarter of the placebo distribution yielded a percentage difference between synthetic and actual that was at least as large as the case of interest.

Model Diagnostics: Exploration of Signal & Noise Discriminatory Power

In this section, we contribute an additional test of model validity. The Hollingsworth and Wing (2018) SCUL method makes it possible to consider extremely large datasets. Our case is representative of a fairly common case where the pre-treatment time period is relatively small, which motivated the cross validation exercises to help avoid overfitting the testing data at the expense of out-of-sample prediction. Ideally, what occurs in that exercise is that the Lasso regression is successfully mining the data and picking out variables that signal information content. However, as the number of potential donors increases, the more possible it becomes to find combinations of information-free donors that can be used in combination to accurately predict our real case-of-interest. Again, cross-validation diminishes this concern, but with very large datasets and limited subsamples of pre-treatment period, we offer another complementary exercise. We investigate this concern by simulating a new potential donor covariate data set with the same mean, standard deviation, and dimensions as our real potential donor matrix. We will first examine the how well the HWSC method performs with this random data to gauge the information value-added of the real donor matrix. In a follow-up, we append into our random data the 16 predictors used by the Lasso regression in generating our main results to get a sense to which the Lasso mixes them with the pure-noise donors.

Figure 8 illustrates the result of attempting to use the Monte Carlo data in the HWSC method, which can be contrasted with the results of **Figure 5**. The pre-treatment period is not a great fit with a mean Cohen D score of 0.41 when compared to 0.11 for the real data, and the Monte Carlo data entirely fails to predict the first out-of-sample year by diverging in the opposite direction. If we add the 16 predictors to the Monte Carlo data, the Lasso regression rescues 15 of the original predictors and adds none of the Monte Carlo predictors. These findings should increase the prior belief that the dataset is not sufficiently large to make it possible for the model to rely on noise, and increase the prior belief that the model is pulling variables with informative signals.

Summary of Synthetic Control Findings

The largest incremental effect with the sharpest precision in the synthetic control results came from 2003, the first treatment year. The estimated treatment effect of \$29.6 million represented a 7.1 percent increase from synthetic United States in 2003. By fiscal year 2006, the treatment effect grew to \$62.7 million, representing a 12.6 percent increase. The sampling variability of the donor pool, consisting of various government finance and economic data from other countries, suggests that this result could have emerged from random chance almost 30 percent of the time.

At best, it seems that BEA90 had only a small, immediate effect on the base level of federal nondefense discretionary spending.

5. Conclusions

The Budget Enforcement Act of 1990 is well-regarded as a demonstration of congress's potential to regulate itself and learn from previous attempts. Discretionary federal expenditures slowed, and even declined, at times during its era from 1992 to 2002. However, this general observation does not easily withstand scrutiny from the econometrics. In a synthetic control method using Lasso regression from 1992 to 2006, the expiration of BEA90 at best provides an initial increase in non-defense discretionary spending but does not clearly alter the trajectory of these expenditures.

This paper is the first to our knowledge to empirically study the federal budget process using the synthetic control method for causal case study analysis. As discussed in the paper, we focus on nondefense discretionary spending because BEA90 expired shortly after the 9/11 attacks and plausibly contaminated any effects on defense spending. Similarly, arguably part of the motivation for allowing BEA90 to expire was to pass a program that would substantially increase mandatory spending, making those results rather uninteresting. Our inference is based partially then on the idea that nondefense discretionary is uncontaminated in any significant way and the effects attributable to BEA90 alone. Since we find rather muted, or perhaps temporary effects only, this is a less concerning bias because it plausibly should have increased the magnitude of the effect. Nevertheless, it is a limitation of the analysis, and future research might provide further insights or data that could improve on the necessary inference assumptions.

REFERENCES

- Abadie, Alberto and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* (March): 113-132.
- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2008. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493-505.
- Albouy, David. 2013. "Partisan Representation in Congress and the Geographic Distribution of Federal Funds." *The Review of Economics and Statistics* 95(1): 127-41.
- Anderson, Sarah E. and Jonathan Woon. 2014. "Delaying the Buck: Timing and Strategic Advantages in Executive-Legislative Bargaining over Appropriations." *Congress & The Presidency* 41(1): 25-48.
- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives* 28(2): 29-50.
- Berry, Christopher R., Barry C. Burden, and William G. Howell. 2010. "The President and the Distribution of Federal Spending." *American Political Science Review* 104(4): 2010.
- Berry, William D. 1990. "The Confusing Case of Budgetary Incrementalism: Too Many Meanings for a Single Concept." *The Journal of Politics* 52(1): 167-96.
- Bozeman, Barry and Jeffrey D. Straussman. 1982. "Shrinking Budgets and the Shrinkage of Budget Theory." *Public Administration Review* 42(6): 509-15.
- Buchanan, James M. 1987. "Tax Reform as Political Choice." *Journal of Economic Perspectives* 1(1): 29-35.
- Caiden, Naomi. (1984). "The New Rules of the Federal Budget Game." *Public Administration Review* 44(2): 109-118.
- Calcagno, Peter T. and Edward J. Lopez. (2017.) "Informal Norms Trump Formal Constraints: The Evolution of Fiscal Policy Institutions in the United States." *Journal of Institutional Economics* 13(1): 211-42.
- Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. 2009. "Measuring Bias and Uncertainty in DW-Nominate Ideal Point Estimates via the Parametric Bootstrap." *Political Analysis* 17: 261-75.
- Congressional Budget Office. 2003. *The Budget and Economic Outlook: Fiscal Years 2004-2013*. A Report to the Senate and House Committees on the Budget. Retrieved from https://www.cbo.gov/sites/default/files/108th-congress-2003-2004/reports/entirereport_witherrata.pdf.

- Davis, Otto A., M. A. H. Dempster, and Aaron Wildavsky. (1966). "A Theory of the Budgetary Process." *The American Political Science Review* LX(3): 529-547.
- Dezhbakhsh, Hashem, Soumaya M. Tohamy, and Peter H. Aranson. 2003. "A New Approach for Testing Budgetary Incrementalism." *The Journal of Politics* 65(2): 532-58.
- Doudchenko, Nikolay and Guido W. Imbens. 2016. "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis." NBER Working Paper No. 22791.
- Dube, Arindrajit, and Ben Zipperer. 2015. "Pooling multiple Case Studies Using Synthetic Controls: An Application to Minimum Wage Policies." IZA Discussion Paper No. 8944. Institute of Labor Economics, Bonn, Germany.
- Edin, Per-Anders and Henry Ohlsson. 1991. "Political Determinants of Budget Deficits: Coalition Effects Versus Minority Effects." *European Economic Review* 35: 1597-603.
- Hollingsworth, Alex and Coady Wing. 2019. "Marijuana Legalization and Substitution of Psychoactive Substances: An Application of Synthetic Control Using Lasso." SPEA-IU Working Paper.
- Joyce, Philip G. and Robert D. Reischauer. 1992. "Deficit Budgeting: The Federal Budget Process and Budget Reform." *Harvard Journal on Legislation* 29: 429-53.
- Larcinese, Valentino, Leonzio Rizzo, and Cecillia Testa. 2006. "Allocating the U.S. Federal Budget to the States: The Impact of the President." *The Journal of Politics* 68(2): 447-56.
- Mikesell, John L. 2014. *Fiscal Administration, Analysis and Applications for the Public Sector*. Ninth Edition. Boston: Wadsworth Cengage Learning.
- Mikesell, John L. and Justin M. Ross. 2019. "After Wayfair: What Are State Use Taxes Worth?" *National Tax Journal* 72(4): 801-20.
- Natchez, Peter B. and Irwin C. Bupp. 1973. "Policy and Priority in the Budgeting Process." *American Political Science Review* 67(3): 951-63.
- Office of Management and Budget. 2015. *The President's Budget for Fiscal Year 2016: Historical Tables*. Washington: OMB. (<https://www.whitehouse.gov/omb/budget/Historicals>)
- Pitsova, Bernard T. and Frank D. Draper. 1984. "Making Sense of the Federal Budget the Old Fashioned Way-Incrementally." *Public Administration Review* 44(5): 401-407.
- Penner, Rudolph G. and C. Eugene Steuerle. 2004. "Budget Rules." *National Tax Journal* 57(3): 547-57.
- Roubini, Nouriel and Jeffrey D. Sachs. 1989. "Political and Economic Determinants of Budget Deficits in the Industrial Countries." *European Economic Review* 33: 903-38.
- Rubin, Irene. 2007. "The Great Unraveling: Federal Budgeting, 1998-2006." *Public Administration Review* 67(4): 608-17.

- Savage, James D. and Amy Verdun. 2007. "Reforming Europe's Stability and Growth Pact: Lessons from the American Experience in Macrobudgeting." *Review of International Political Economy* 14(5): 842-67.
- Schultze, Charles. 1968. *The Politics and Economics of Public Spending*. Washington, D.C.: The Brookings Institution.
- Shi, Min and Jakob Svensson. 2006. "Political Budget Cycles: Do They Differ Across Countries and Why?" *Journal of Public Economics* 90: 1367-89.
- Suskind, Ron. 2004. *The Price of Loyalty: George W. Bush, the White House, and the Education of Paul O'Neill*. Simon & Schuster
- Wildavsky, Aaron. 1964. *The Politics of the Budgeting Process*. Boston: Little, Brown.
- Wildavsky, Aaron and Richard Fenno. 1966. *The Power of the Purse*. Boston: Little, Brown & Co.
- Woo, Jaejoon. 2003. "Economic, Political, and Institutional Determinants of Public Deficits." *Journal of Public Economics* 87: 387-426.
- Xu, Yiqing, 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models." *Political Analysis* 25, 57-76.

FIGURE 1

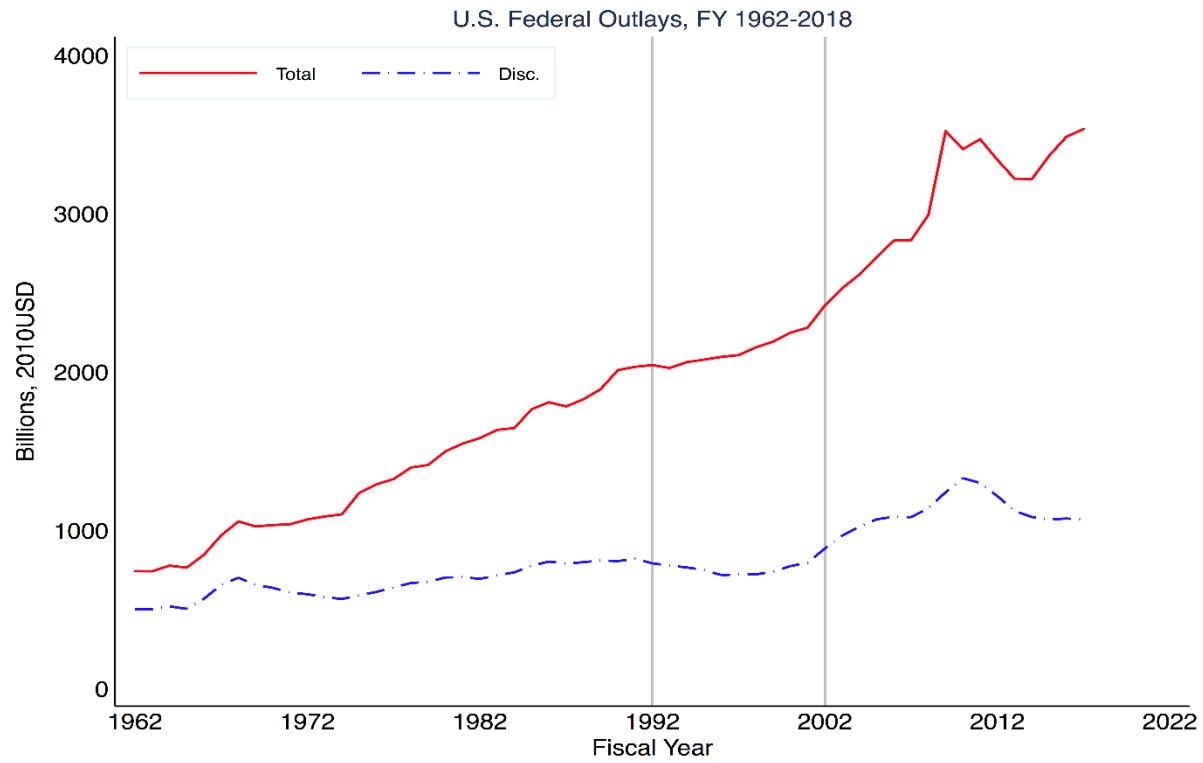


FIGURE 2

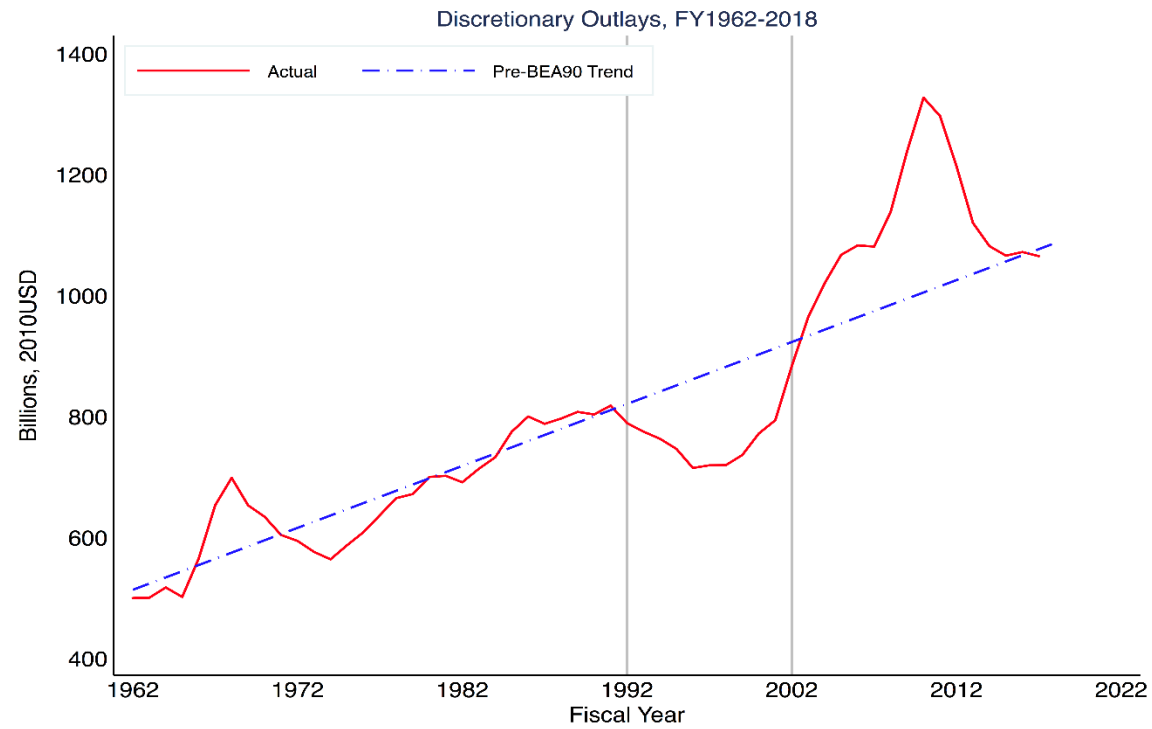


Figure 3: Federal Budget Process

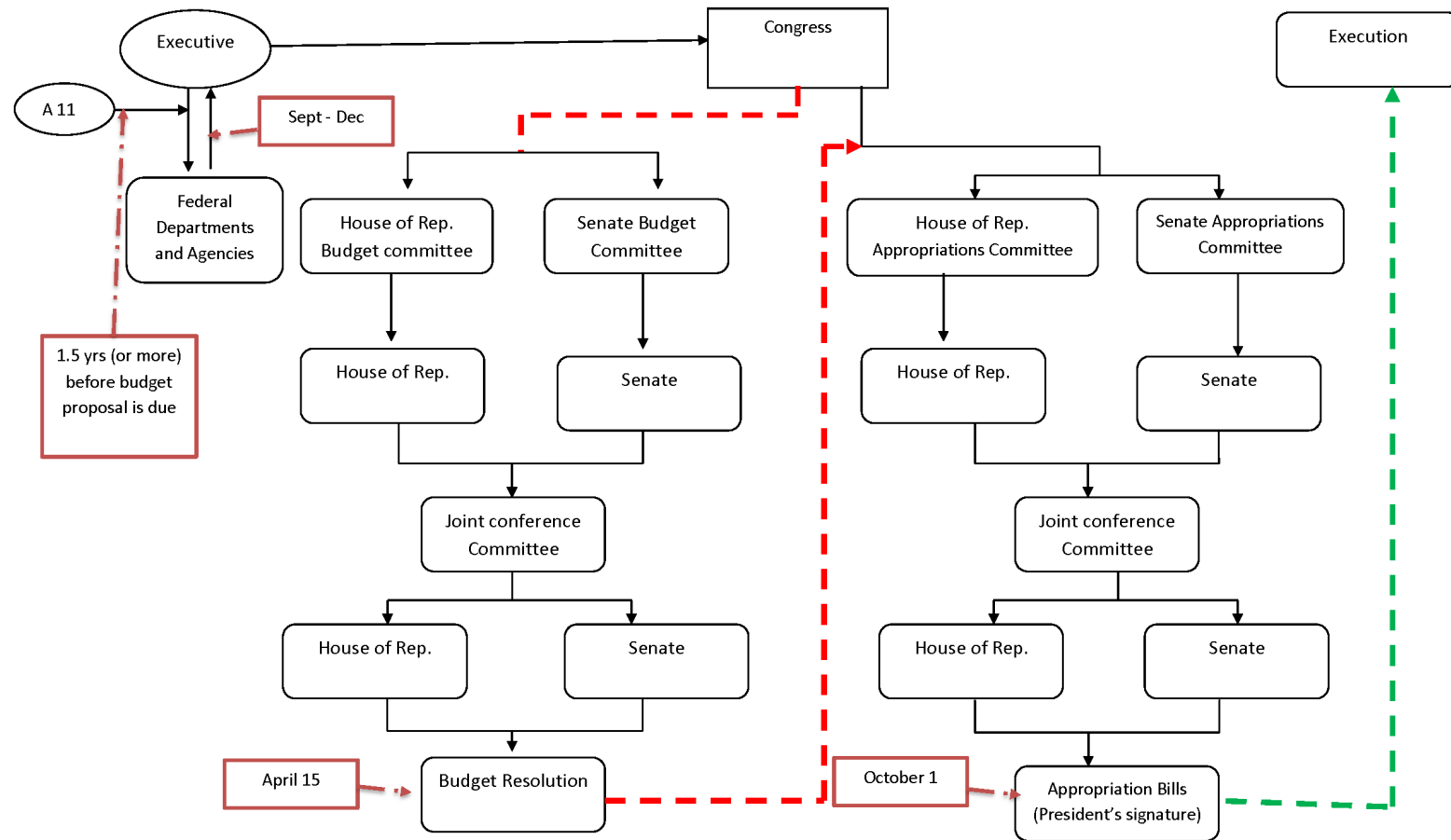
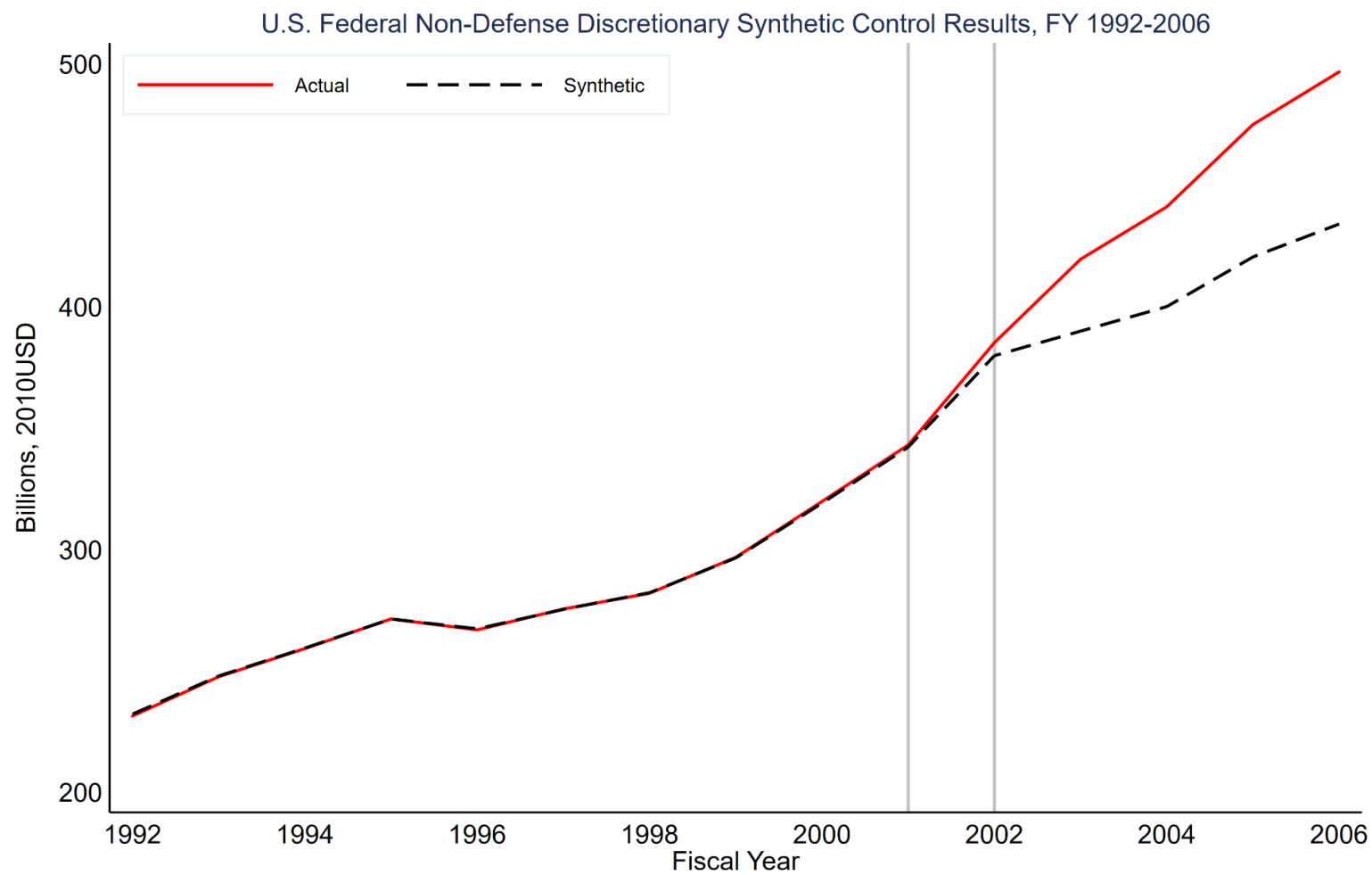


Figure 5



Left Vertical Line: End of synthetic model pre-treatment period. Right Vertical line: Begin treatment period.

Figure 6: Placebo Results from Synthetic Control Analysis

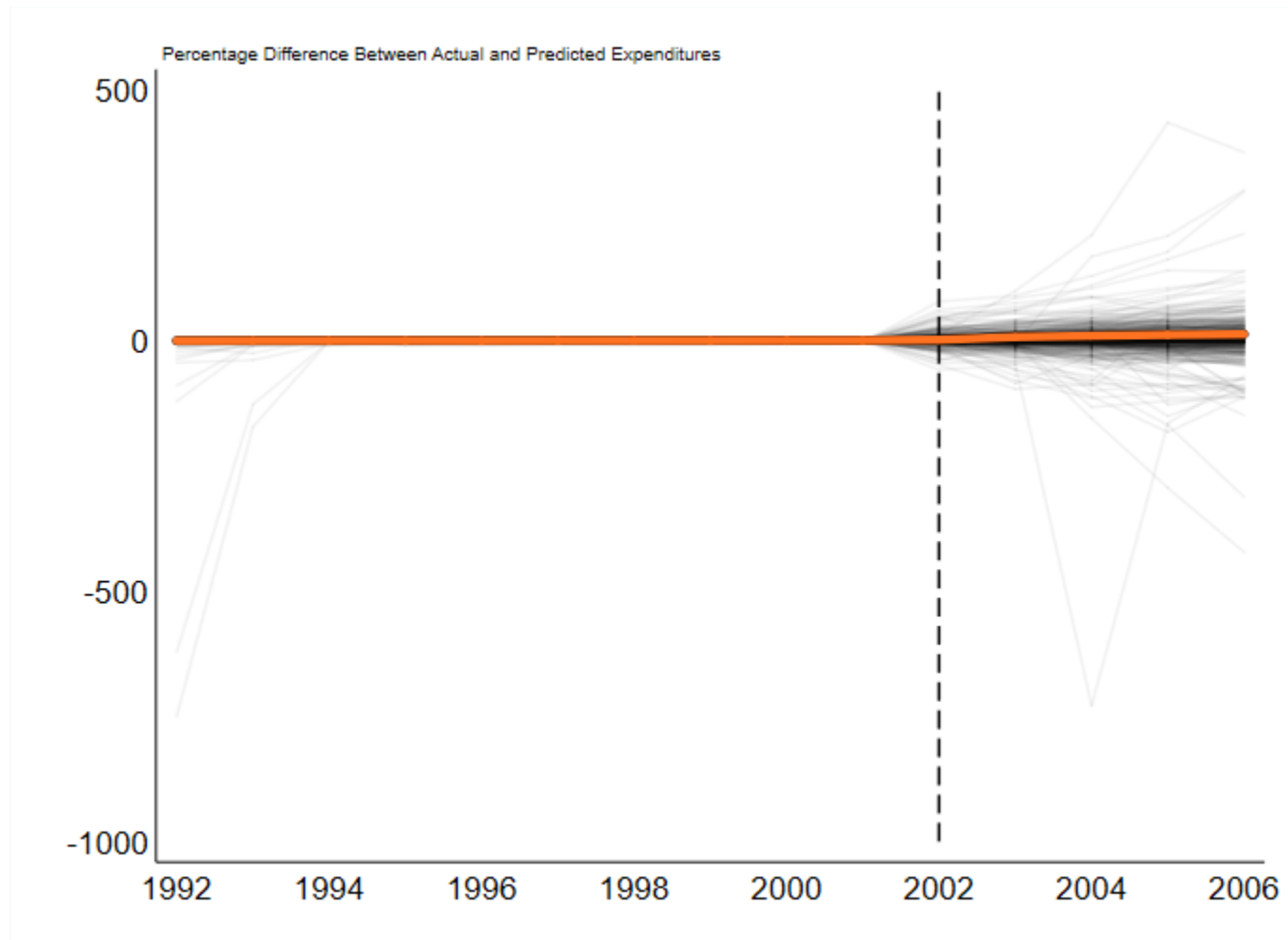


Figure 7: Distribution of Placebo Results by Treatment Year

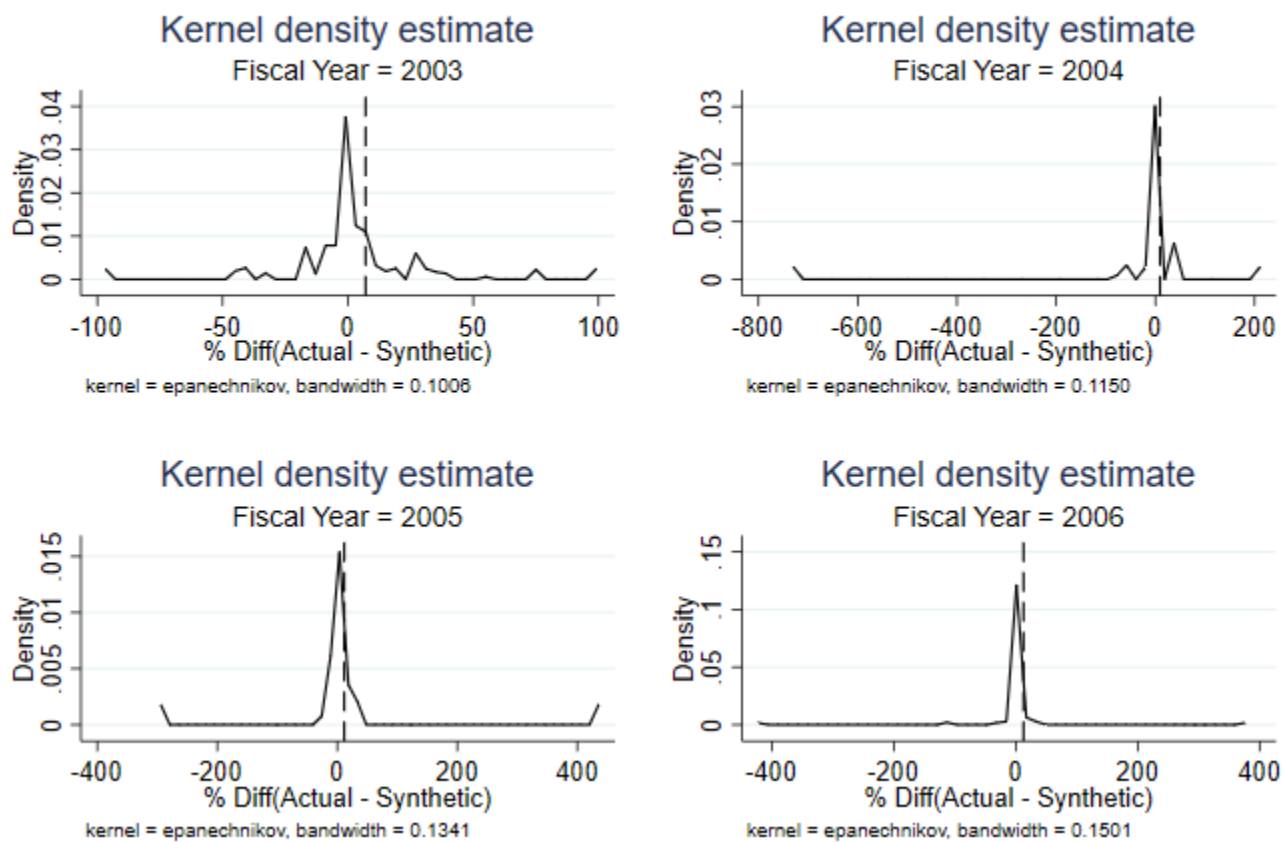


Figure 8: Synthetic Control Results from Monte Carlo Donor Pool

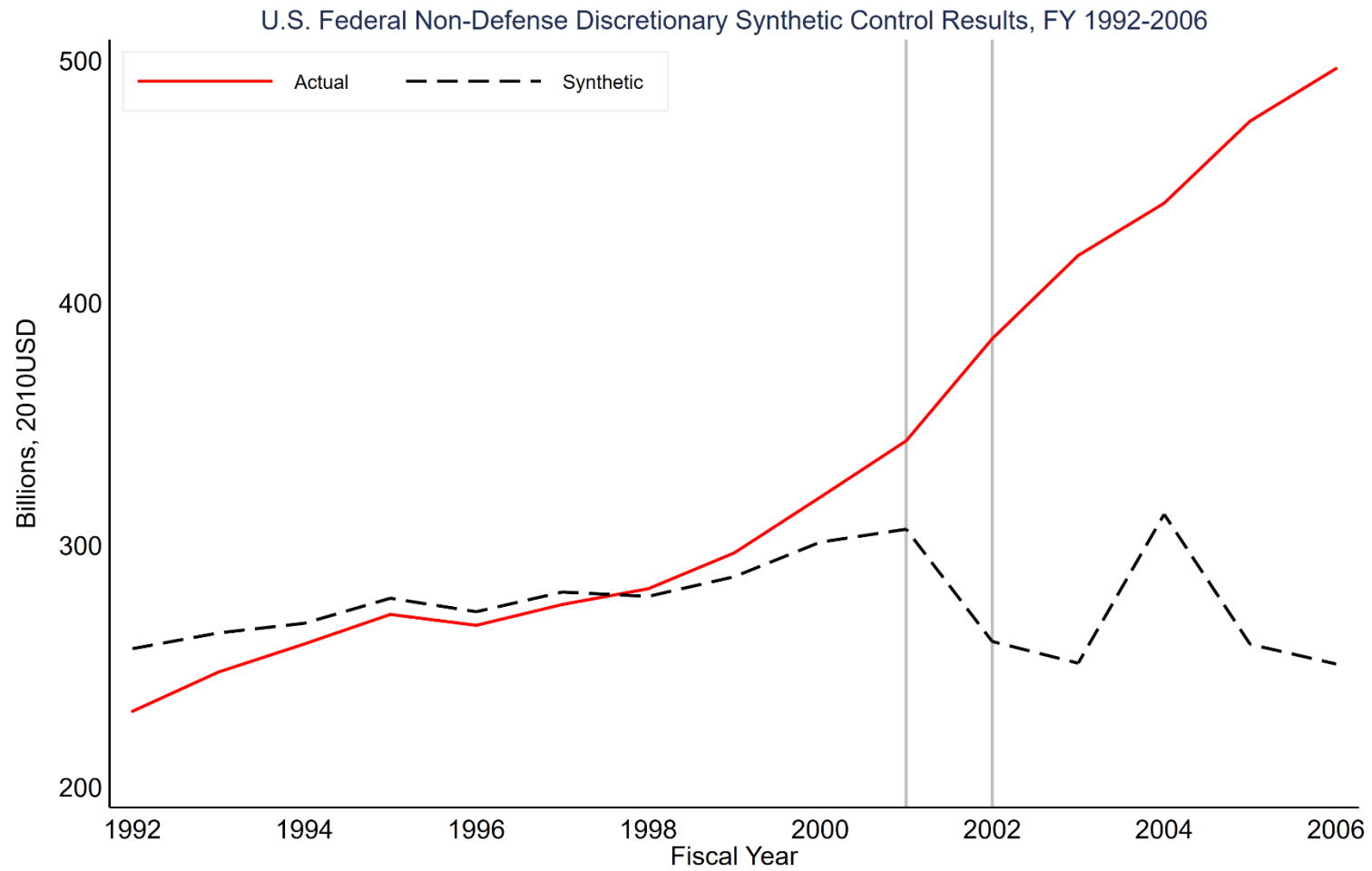


Table 1: BEA90 Spending Limits Versus Actual Spending, 1991-2002

| Spending Limit | | | Amount Actual Above or Below (-) Limit | |
|-----------------------|-----------------------------|----------------|-------------------------------------------------------|----------------|
| Year | Budget Authority | Outlays | Budget Authority | Outlays |
| 1991 | \$492 | \$514 | \$10 | -\$14 |
| 1992 | \$503 | \$525 | \$14 | -\$6 |
| 1993 | \$511 | \$534 | \$11 | \$5 |
| 1994 | \$511 | \$535 | \$2 | \$7 |
| 1995 | \$518 | \$541 | -\$16 | \$4 |
| 1996 | \$519 | \$547 | -\$18 | -\$15 |
| 1997 | \$528 | \$547 | -\$17 | \$0 |
| 1998 | \$531 | \$548 | -\$1 | \$4 |
| 1999 | \$533 | \$559 | \$49 | \$13 |
| 2000 | \$537 | \$564 | \$47 | \$51 |
| 2001 | \$542 | \$564 | \$122 | \$85 |
| 2002 | \$553 | \$562 | \$182 | \$172 |

All dollars in billions U.S. (nominal).

Source: Congressional Budget Office (2003), Table A-2.

Table 2: Predictors Selected by Lasso Regression Accounting for >1% of Synthetic USA

| Share of Prediction | Country | Indicator |
|--------------------------------|--------------------|---------------------------------------------------|
| 17.4% | Jamaica | Total Central Government Debt |
| 9.5% | Tonga | Total Gross Domestic Product |
| 8.4% | Dominican Republic | Manufacturing Value Added to GDP |
| 8.0% | Costa Rica | Extra-Budgetary Central Government Revenue |
| 3.9% | Malaysia | General Government Final Consumption Expenditures |
| 3.3% | Belize | Gross Domestic Product |
| 2.3% | Pakistan | General Government Final Consumption Expenditures |
| 1.1% | Macedonia | Manufacturing Value Added to GDP |

Prediction shares reported for 2006, intercept accounts for 49.9% of prediction.

Table 3: Synthetic Control Results for Federal Non-Defense Discretionary Expenditures

| | | Treatment Years (Post-BEA90) | | | | |
|----------------------------|---------------|-------------------------------------|---------------|---------------|---------------|-------------|
| | FY2002 | FY2003 | FY2004 | FY2005 | FY2006 | Post |
| Actual Spending | 385.0 | 419.4 | 441.0 | 474.9 | 496.6 | 458.0 |
| Estimated Treatment Effect | 5.3 | 29.6 | 41.1 | 54.5 | 62.7 | 47.0 |
| %(Actual-Synthetic) | 1.4 | 7.1 | 9.3 | 11.5 | 12.6 | 10.1 |
| p-value | 0.42 | 0.27 | 0.27 | 0.28 | 0.29 | 0.30 |