Accounting for Variation in Policy Type and Length of Implementation when Exploring the Impacts of Outcomes-Based Funding on Bachelor's Degree Production

M. Kate Callahan; Dae Kim; X. Lindsey Liu, Cyril Cherian

Research for Action

Author Note

This research was supported by grants from the Lumina Foundation and the Bill and Melinda Gates Foundation. Opinions reflect those of the authors and do not necessarily reflect those of the granting foundations, their officers or employees.

Address queries to Kate Callahan, Research for Action, Philadelphia, PA, 19110, telephone: (267) 295-7766; email: kcallahan@researchforaction.org.

Abstract

As of 2018, 30 states have implemented or formed a task force for outcomes-based funding (OBF) policies. With the wide spread of the policies, we observe increasing variation in policy designs to address state strategic goals and completion agendas. No research has examined the impact of OBF policies on bachelor's degree production while accounting for variation in policy design and length of implementation. We created a 12-year balanced panel dataset from 476 public, four-year, degree granting institutions in 19 OBF states and 30 comparison states from 2005 to 2016. We examined the impact of OBF policies on both total number of bachelor's degree awarded and the logged value of number of degrees awarded. We used difference-in-differences models with a set of institution-level and state-level covariates to further investigate the effect of OBF on the two outcomes and how its effect differs by length of implementation and types of policy design. We found significant positive impact of OBF on outcomes in states in which OBF policies had been in place for a longer period of time and in states that implemented more robust versions of OBF policies. Our results stress the importance of recognizing several types of OBF policies when examining their impact.

Keywords: Outcomes based funding, higher education

Accounting for Variation in Policy Type and Length of Implementation when Exploring the Impacts of Outcomes-Based Funding on Bachelor's Degree Production

Introduction

Stagnating postsecondary attainment rates along with increases in average time-to-degree, tuition and student debt loads over the past thirty years have triggered intense public criticism of U.S. higher education and a demand for increased accountability from colleges and universities (Pew, 2011; Zumeta, 2011; Bound, Lovenheim, & Turner, 2012). In this environment, states have turned increasingly to performance funding to incentivize public colleges and universities to retain and graduate more students. Performance funding is a term used to describe a state or system-level higher education funding policy that links state funding to an institution's performance on a set of state-defined outcomes metrics (Burke & Minassians, 2003). Performance funding has continued to gain popularity as many states have reduced their level of investment in higher education and seek ways to maximize the impact of available funding. At the beginning of fiscal year 2018, 30 states had either adopted higher education performance funding policies or formed a task force to develop such policies (Snyder & Boelscher, 2018).

Performance funding "incorporates basic private-sector logic to financially reward organizations for desired outputs, with the expectation that organizations will alter practices to secure this money" (Li & Zumeta, 2016, p.2). Performance funding's focus on outputs marks a distinct shift from previous higher education oversight, which primarily focused on "inputs" (e.g. spending and enrollment), to a "new accountability" where public institutions are monitored for their performance on a range of outcomes such as certificate and degree production, on-time

graduation rates, course completion/ credit progression, and research productivity (Zumeta, 1998; McClendon, Hearn, & Deaton, 2006; Zumeta, 2011; Tandberg & Hillman, 2014).

Performance funding (PF) policies have evolved substantially over the past thirty-plus years. Both policymakers and institutional leaders contend that policy effectiveness must be considered within the context of both policy design (i.e. policy components such as percentage of funding, types of metrics, and mission differentiation) and how it is implemented (e.g. consistency of policy, years of implementation, and stop-loss/ hold-harmless provisions) (Callahan et al., 2017a; b; Meehan, Callahan & Stull-Hunter, 2017). To capture some of these differences the field has coined the term Outcome-Based Funding (OBF) to refer to more recent PF policies designed specifically to increase attainment by awarding funding based on measures of progression, efficiency, and completion(Dougherty & Reddy, 2013; Rutherford & Rabovsky, 2014; Hearn, 2015; Hillman, Tandberg, & Fryar, 2015; Snyder, 2015; Dougherty et al., 2016; Li & Zumeta, 2016; Gándara, Rippner, & Ness, 2017).

While PF/ OBF policies have proliferated and diversified over the past decade, most research on their effectiveness does not reflect the changing policy landscape and continues to examine performance funding as a homogenous policy. Because previous studies of OBF policy effectiveness do not account for variation in policy types, scholars have not been able to inform the field about whether and how differences in policy components and implementation affect intended outcomes. The current study seeks to bridge the gap between research and policy by utilizing a comprehensive typology of OBF policies being implemented across the United States to examine whether and how variation in performance funding policies and length of implementation affects the magnitude and timing of institutional response, and specifically bachelor's degree production.

Literature Review

Variation in OBF policies

OBF policies vary considerably in terms of both design and implementation. Snyder and colleagues constructed a typology which captures some of this variation by classifying all active OBF policies into four ordinal types (Snyder, 2015; Snyder & Fox, 2016; Snyder & Boelscher, 2018). The four-level typology, which has been refined twice since original publication in 2015 to reflect changes in OBF state policy, conveys the varying degrees to which a state's current OBF policy includes "critical elements" such as level of funding awarded, differentiation in metrics and weights by sector, prioritization of underrepresented students, and stability of the policy over multiple years. The 2015 and 2016 classifications assign types at the state level. The most recent 2018 classification assigns types at the sector-level to further document within-state variation in states that have different OBF policies for two- and four-year institutions, or OBF in only one sector. According to the 2016 statewide classification (Snyder & Fox, 2016), OBF policies in many states were either rudimentary models (i.e. Type I) or still developing promising practices (i.e. Type II). Robust OBF policies (i.e. Type IV) were relatively rare. Among the 25 states implementing OBF policies in any sector in 2016 only two, Ohio and Tennessee, were classified as implementing the most robust Type IV models, whereas there were four Type III states, eleven states Type II states and eight Type I states.

However, Snyder and Fox (2016) also found that OBF policies continue to evolve in many states, and observed that OBF policies in Colorado, Florida, Montana and Oregon had become more sophisticated and thus advanced in the typology from 2015 to 2016. This evolution towards more robust policies is evident in the most recent 2018 classification as well. Of the 19 states with OBF policies that apply to four-year sector institutions, seven states are classified

with robust (Type IV) policies compared to three states with rudimentary (Type I) policies and three states that are still developing promising practices (Type II). The evolution of OBF policies into relatively more robust models over time emphasizes the importance of both capturing important differences in policy components and tracking the number of years that the most recent iteration of the policy has been in place when estimating the effectiveness of OBF policies.

Measuring the effectiveness of OBF policies

To date there has been extensive research into the effectiveness, efficiency, and intended and unintended consequences of performance funding policies. None of this research has yielded strong evidence linking performance funding to significant improvements in postsecondary outcomes (Shin & Milton, 2004; Shin, 2010; Sanford & Hunter, 2011; Tandberg & Hillman, 2014; Tandberg, Hillman, & Barakat, 2014; Hillman, Tandberg & Fryar, 2015; Hillman, Fryar & Crespín-Trujillo, 2017). Studies focusing specifically on outcomes for four-year institutions have concluded that performance funding has little or no impact on first-time-in-college graduation rates (Shin & Milton, 2004); six-year graduation rates (Shin, 2010; Sanford & Hunter, 2011); and total bachelor's degree completions (Tandberg & Hillman, 2014). However, most OBF studies have theorized that performance funding functions as a homogenous policy across implementing states. Based on this assumption, scholars have employed a dichotomous classification, assigning states as either performance funding or non-performance funding when identifying treatment and comparison groups. Thus, it is possible that the generalized findings of non-effectiveness found when examining the impacts of *all* OBF policies on four-year degree production may be obscuring differential impacts among types of OBF policies that vary in policy design and length of implementation. Further, because most previous studies of OBF policy effectiveness do not account for variation in policy types, scholars have not been able to

inform the field about whether and how differences in policy components and implementation affect intended outcomes.

Notably, the handful of studies (Tandberg & Hillman, 2014; Rutherford & Rabovsky, 2014; Hillman, Fryar, & Crespín-Trujillo, 2017; Li & Kennedy, 2018) that have considered any type of variation among policies have found some evidence that policy design and implementation matter. For example, when examining the effects of performance funding in general on degree productivity, Tandberg and Hillman (2014) also considered how the length of policy implementation might influence policy effectiveness. Specifically, the authors used a quasi-experimental difference-in-differences design that included both state and year fixed effects to address potential bias from unobserved heterogeneity within and across states, and multiple comparison groups as robustness checks. Their study consisted of two models: one to examine the average treatment effect across all years of implementation, and a second model that included interaction terms for the treatment and each year the policy was implemented to test for a "duration interaction" - changes in the impact of the policy over years of implementation. Examining performance funding policies in place between 1990 and 2010, they found that the average effects of performance funding were small and insignificant in early years after policy adoption, yet grew to a significant positive effect on total bachelor's degree completions beginning in the seventh year of implementation. Importantly, the authors report substantial variation in the effectiveness of the policies across 19 states examined, where significant positive effects were found in four states (Kansas, Minnesota, New Jersey, and Washington), significant positive effects in only one of the two models (Arkansas, Missouri and Oklahoma), and no effects or negative effects in the remaining 12 states. This study highlighted the importance of the length of implementation to provide a more nuanced look into the impact of performance

funding. Additionally, Tandberg and Hillman call for future research capturing variation across state policies when more years of data become available and more states have implemented policies long enough for researchers to observe the emerging effects.

Two more recent studies have begun to address variation in policy design across OBF states. Acknowledging the temporal heterogeneity in policy implementation, Hillman, Fryar, & Crespín-Trujillo (2017) attempt to address differences in policy types by comparing the production of certificates, associate degree and bachelor's degrees in two states considered to have robust policies (i.e. Ohio and Tennessee) to other states with presumably weaker OBF policies, as well as states without any OBF. Although several inconsistencies in the identification of states with OBF and a lack of thorough documentation of how states were assigned to three different comparison groups threatens the validity of the study's findings, the authors' intent to examine the discrete effects of robust OBF policies continues to highlight the value of addressing policy variations when research examines the effect of OBF.

Li and Kennedy (2018) conducted the only study to date that attempts to capture variation in OBF policies by measuring the impacts of multiple policy types. Applying Snyder and Fox's 2016 typology of OBF policies, they included policy type as a covariate in their difference-in-differences model to explore how variation in policy types affect certificate production across 751 two-year colleges. To highlight the value of addressing nuance in policy design of OBF, the authors asserted homogeneity of OBF policies and their first model showed null effects on certificate production that were similar to those in previous literature that does not account for policy variation. However, examining the effects by different types of OBF policies, they found positive significant effects in states that implemented Type III and IV policies, which are the more rigorous models in Snyder and Fox's typology (2016). In contrast, Type I policies –

defined as rudimentary/ pilot policies-- showed negative to no effects across all outcomes. Their work calls for further research that considers variation among OBF policy designs and the temporal heterogeneity to examine the impact of OBF policies on additional institutional outcomes in both the two- and four-year sectors.

Informed by Li and Kennedy (2018) as well as earlier contributions of Hillman, Tandberg and colleagues (Tandberg & Hillman, 2014; Tandberg, Hillman, & Barakat, 2014), the current study examines the impact of OBF policies on bachelor's degree production using a multi-state analysis that combines two important components of OBF policy gleaned from prior research on OBF and performance funding: temporal heterogeneity in policy implementation and variation in the robustness of policy design. Following Li and Kennedy's (2018) analysis of the effects of OBF policy types on community college outcomes, this study is the first to apply the Snyder and Fox (2016) typology to examine the impacts of different types of OBF policies on bachelor's degree production over time.

Specifically, we conduct a difference-in-differences analysis and several robustness tests against multiple comparison groups and two outcome variables – total number of bachelor's degree awarded and its logged value to account for the large variations in the sizes of institutions. We also include the parallel trend test to further examine if our assumption can be validated. While our research is not free from limitation, our findings provide further evidence of the importance of including variation in both design and implementation when examining the impacts of large-scale state-level postsecondary policy. This study also offers empirical insights for policymakers and researchers to consider when developing or refining OBF policies. We address the following questions:

1) Has the implementation of OBF policies affected bachelor's degree production at public four-year institutions in participating states?
2) Does the impact of OBF on bachelor's degree production change over time?
3) Does the impact of OBF on bachelor's degree production vary between states that implemented different types of policies?

Conceptual Framework

Rabovsky (2012) outlines the causal logic of performance funding policies and argues that they are designed to restructure incentives to trigger higher education institutions to make changes that will improve performance as measured by student outcomes (p. 679). Rabovsky points out that most studies of performance funding jump directly to student outcomes without considering the mechanism of institutional response. We therefore turn to a conceptual framework developed by Naughton (2004) to explain how state accountability programs influence institutional change and how a state's perceived legitimacy among postsecondary institutions can influence institutional response, and in turn the impact of a given policy such as OBF on student outcomes.

Naughton's framework brings together rational-choice institutionalism as it is defined by economists and political scientists (e.g., North, 1990; Weingast, 2002) and the institutionalism of organizations as developed primarily by sociologists (e.g., DiMaggio & Powell , 1991; Scott, 2001). By integrating these two fields, he emphasizes the complex nature of institutional response to external incentives, pointing out that the decision-making process of institutional leaders is both rational and normative/cultural. Specifically, Naughton (2004) asserts that the resource exchange from state to institutions is dependent on campuses demonstrating their

legitimacy to the state. He also argues that institutional compliance with accountability policies depends on whether the state demonstrates its legitimacy to campuses, since when institutions perceive a state's governance and policies as legitimate, institutions are more likely to comply and respond to the state by aligning institutional policies and practices with the state's policies.

In addition, Naughton's framework emphasizes the bi-directional resource exchange between campuses and state governments and highlights the *symbolic* and *substantive* measures of legitimacy and resources within the exchange. The perceived legitimacy of states and postsecondary institutions and the resources they yield may be symbolic in nature (e.g., demonstrated through strong leaders and governance authority) or substantive (e.g. demonstrated by strong performance or processes). For example, the symbolic legitimacy of a state's resources (i.e. higher education funding) may be tied to the reputation of a governor, where a funding policy passed under a well-regarded governor would garner symbolic legitimacy. In contrast, a similar funding policy passed under a governor perceived to be antagonistic to public higher education might not be perceived as legitimate.

However, the substantive legitimacy of a state's resources would correspond more directly to the state's fiscal investment in higher education, where a funding formula passed in a state that substantially invests in public higher education would potentially have higher perceived legitimacy than a similar formula passed in a state that is making significant cuts to funding for higher education. Applied to higher education institutions, high symbolic legitimacy may be reflected by a strong president or board, while an institution's substantive legitimacy might be conveyed to the state through robust student enrollment and high graduation rates.

In respect to performance funding/ OBF policies, Naughton's conceptual framework suggests that there will likely be variation in how institutions respond to different types of OBF

policies. Depending on the perceived legitimacy of a state's policy and the resources being exchanged, campuses may choose to respond symbolically and/or substantively or not at all. Symbolic responses on the part of campuses can include "rhetorical alignments" between campus leaders, strategic plans, or institutional goals to align with the intent of a policy, as well as "tactics that buffer actions from further scrutiny" (Naughton, 2004, p.9) such as assembling a data team. More substantive reactions include actual changes in policies or programs to prioritize outcomes aligned with the policy's goals. We hypothesize that when states and institutions perceive one another as legitimate, accountability policies can influence institutional behavior, and in turn student outcomes.

Naughton's concepts of symbolic and substantial resources also suggest why some types of OBF policies would elicit more robust institutional responses which would in turn produce more substantial changes in outcomes (i.e. exhibited by increased bachelor's degrees) than other types. For example, OBF policies that affect high proportions of state funding and which provide stable funding from year to year can be predicted to garner greater substantive resources from institutions than OBF policies that award relatively low levels of funding and are contingent on new money being allocated. Moreover, OBF policies that are well-aligned with statewide attainment goals, include outcomes metrics that reflect variations in institutional missions, and prioritize outcomes for historically underrepresented students would be expected to have greater perceived legitimacy among institutions than policies that may not be aligned to statewide completion goals, do not differentiate metrics by sector/ mission, and do not prioritize underrepresented students. Thus, we hypothesize that an analysis of OBF impact that distinguishes among types of OBF policies may produce results that differ from studies that treat OBF as a single policy.

Furthermore, we hypothesize that there is a temporal dimension to institutional response. Specifically, while symbolic responses (e.g. aligning institutional strategic plans to the OBF policy and conducting internal analysis to predict institutional performance on OBF metrics) might occur quickly, substantive changes made by institutions to improve student outcomes (e.g., changes to student advising, implementing early warning systems, and increasing course availability) are likely to require additional planning and resources before full implementation. The literature on policy implementation describes the multi-stage nature of institutional response to state policies (de Lancer Julnes &Holzer, 2001; McLaughlin, 1987). Summarizing years of research on policy implementation, McLaughlin (1987) emphasizes that "slam bang policy effects" are uncommon and associated with more "ephemeral gains," while institutional responses that take more time are often more beneficial and stable (p.175). In line with this literature, we expect that substantive responses from institutions are likely to take additional time to influence longer-term student outcomes like increases in bachelor's degree production. Thus, we examine whether the impact of OBF increases over time (i.e. years of implementation) as symbolic responses are supplemented with more substantive ones.

Because there has been no empirically-driven analysis of the components of OBF policies that distinguishes robust and weak policies, we adopt the comprehensive state-by state typology developed by Snyder and colleagues at HCM Strategists. This typology, which originated in 2015 and was updated in 2016 and 2018, systematically sorts all current OBF policies by multiple design components into four ordinal types that are perceived to be related to policy strength and sophistication (Li & Kennedy, 2018). The typology framework (Snyder, 2015; Snyder & Fox, 2016; Snyder & Boelscher, 2018) aligns well with an application of Naughton's theory. Specifically, the robust characteristics of Type IV OBF policies (E.g. well-aligned with

state completion policies, recurring funding source, affects a relatively high level of funding, includes outcomes metrics that differentiate by institutional mission, prioritizes underrepresented students, and has been sustained for multiple years) would presumably convey more substantive legitimacy between states and institutions for increasing bachelor's degree production than more rudimentary Types I policies which do not have consistent or significant levels of funding attached and lack outcomes metrics for degrees/ credential completion. Such alignment enables us to investigate our hypothesis that there is a relationship between the conveyed substantive legitimacy (i.e. relative robustness) of OBF policies and the effectiveness of the policy in improving institutional outcomes.

To assign OBF policy types to each of the OBF treatment states in the analysis we use the Snyder and Fox (2016) typology rather than their 2018 version because 2016 is the most recent year of data included in our panel data set. Table 1 displays the key policy characteristics that are used to sort each policy into one of the four types, where Type I represents the most "rudimentary" policies and Type IV represents the most "robust" OBF policies as of FY16. Table 1 also indicates which state policies are assigned to each type in the last row.

It is important to note that each iteration of the Snyder et al. typology is intended to capture the characteristics of a state's most current policy at the time that the typology is published. Therefore, the Snyder and Fox (2016) typology reflects the characteristics of policies as of FY 2016. OBF policies evolve over time and so it is possible that the assigned type does not reflect a state's policy in all previous years. However, in the absence of a longitudinal data source that consistently tracks OBF policy types by year, this typology offers the best approximation of distinctions between policies available and has been used previously by Li and Kennedy (2018) to examine the effects of OBF policy types on outcomes in the two-year sector.

Table 1
Summary of Policy Characteristics Associated with the Four Outcomes-Based Funding (OBF) Policy Types in Snyder and Fox's 2016 Typology

Features of policy	Type I	Type II	Type III	Type IV
State completion/attainment goals	Maybe[a]	Maybe	Yes	Yes
Base funding (in addition to bonus or only base funding)	No	Yes	Yes	Yes
Percentage of funding tied to statewide analysis	Low[b]	Low	Moderate	High
Underrepresented students	Maybe	Maybe	Yes	Yes
Differentiation in metrics and weights by sector	No	No	Yes	Yes
Degree/Credential completion	Maybe	Yes	Yes	Yes
Formula driven	No	No	Maybe	Yes
Sustained for 2 or more consecutive fiscal years	No	No	No	Yes
States	IL, MA, MI, MO, ND	FL, IN, ME, MN, NM, OR, UT, PA	AR, CO, MT, NV	OH, TN

Note.

[a] All states in our analysis have degree completion information. Therefore, Type I refers to institutions with completion metrics, bonus funding and no base funding.
[b] Low (less than 5%), Moderate (5-24.9%), High (more than 25%)

Methods

Data and Sample

We constructed the panel data sample with the list of OBF states identified in Snyder and Fox's OBF typology (Snyder, 2015; Snyder & Fox, 2016; Snyder & Boelscher, 2018). Snyder and Fox (2016) reported that a total of 19 states were implementing OBF in four-year colleges and universities as of 2016. The comparison sample consists of 30 states that had yet to

implement OBF as of 2016, verified by state legislative documents and a database on funding models published by National Conference of State Legislatures (NCSL)[1]. South Carolina is the only state excluded from both the treatment and comparison groups. South Carolina is not included as an OBF state because it is not included in the 2016 Snyder & Fox typology. However, South Carolina did have a rudimentary form of OBF in place from 1996-2002, and Snyder and Boelscher (2018) note that South Carolina's policy took the form of early (i.e. pre-OBF) performance funding policies by rewarding inputs such as quality of faculty and quality of entering students. Because South Carolina does not fit clearly into either the treatment or comparison group, we do not include it in our anlaysis.

This study utilized data on public four-year colleges and universities from the U.S. Department of Education's Integrated Postsecondary Education Data System (IPEDS) for years 2005 through 2016. State-level economic growth and unemployment data for the same time were from the Bureau of Economic Analysis (BEA) and the Bureau of Labor Statistics (BLS). We constructed a 12-year, balanced panel dataset with a total of 476 four-year public institutions, 184 of which were from 19 OBF states (treatment group) and 292 institutions from 30 states with no OBF policies in place as the comparison group. Within the sample, institutions were included if they met the following criteria: identified in IPEDS as a public, four-year, degree-granting institution; remained open for all years in study period; were identified as a Research, Master's, or Baccalaureate institution according to the Carnegie classification; consistently rewarding more bachelor's degrees than associate degrees and reported the number of bachelor's degrees

[1] According to NCSL, Mississippi approved a performance-based funding model in fiscal year 2009 which was supported the implementation of the new allocation model through fiscal year 2014. However, no evidence can be found that funding had been allocated through the approved model, so we treated Mississippi as a comparison state.

awarded in all study years.[2] Additionally, institutions whose names included the words "Academy," "Community," "Graduate," "Maritime," or "Medical" were excluded from the sample. In some instances, regional campuses reported through their main campus in one year and as their own unit in another year. In such cases, we combined the observations from regional campuses with the observations for the main campus in the years that they reported separately in order to maintain consistency in data reporting.

Coding OBF Treatment Variable

We conducted a thorough review of OBF policies using Dougherty and Natow (2015) and National Conference of State Legislature (2016) to code the implementation year of OBF in each state. In this study, we define policy *implementation* as the first year in which funding was distributed based upon the OBF formula, as distinct from the year in which the policy was adopted. Rutherford and Rabovsky's (2014) analysis coded treatment as beginning in the year that the OBF policy was *adopted*—that is, the year that the policy was passed into law or regulation. However, there is substantial variation across states in the elapsed time between policy adoption and implementation, ranging from zero to two years across the 19 OBF states in this study (see Table 2). The driving premise of OBF policies is that institutions will be incentivized to improve their performance on state-specified outcomes by the differential disbursement of state dollars, because "[m]oney gives meaning as well as support to programs and priorities" (Burke, 2002, p. 267). Therefore, since the treatment is intended to be the funding itself, and dollars are not consistently dispensed at the point of policy adoption across

[2] In Pennsylvania, we include only institutions in the Pennsylvania State System of Higher Education (PASSHE) because other institutions that receive substantial public dollars in Pennsylvania are not included in its OBF policy.

OBF states, we use implementation year to ensure a more consistent measure of treatment across states.

To determine the first year of OBF implementation, we drew on state policy documents, budgets, and official websites of state higher education departments and interviews with state higher education officials[3] to identify the year in which OBF began to affect the amount of state dollars received by institutions. Besides contributing to arriving at an accurate implementation timeline, our interviews also confirmed the amount of funding tied to the formula, the metrics considered in the formula, and institutional response to the policy. In resolving inconsistencies, interviews with state higher education officials were given the strongest weight, followed by official budgets, official state documents, and official state websites. If state documents conflicted and we were unable to conduct interviews with state higher education officials, we coded the implementation year as the one that was used in the majority of sources.

Following these steps, 19 states with OBF policies in 2016 were identified. Table 2 presents the OBF treatment states in the sample and the OBF *adoption vs. implementation* timeline. This table indicates the importance of distinguishing between adoption and implementation, since implementation lagged at least 1 year and up to 3 years after adoption in 15 of the 19 OBF states.

Table 2.

States with OBF as of 2016: Adoption and Implementation Years (in State Fiscal Years)

State	Year adopted into Law or regulation	Year implemented (Funds distributed by formula)	Years Difference
Arkansas	2011	2014	3
Colorado	2014	2016	2

[3] We interviewed higher education officials from state departments of education in Arkansas, Indiana, Louisiana, Maine, Nevada, Ohio, Oregon, and Tennessee.

Florida	2014	2016	2
Illinois	2011	2013	2
Indiana	2009	2009	0
Maine	2013	2014	1
Massachusetts	2014	2015	1
Michigan	2013	2013	0
Minnesota	2013	2014	1
Missouri	2014	2015	1
Montana	2013	2015	2
Nevada	2011	2012	1
New Mexico	2012	2012	0
North Dakota	2013	2013	0
Ohio	2009	2010	1
Oregon	2015	2016	1
Pennsylvania	2012	2013	1
Tennessee	2010	2012	2
Utah	2015	2016	1

Outcome and Control Variables

We modeled two types of OBF outcome variables to test the robustness of OBF impact estimates to the choice of functional form: the total number of bachelor's degrees awarded, and the log of total degrees awarded. Using the log of total degrees awarded is analogous to estimating the percentage change in total degrees awarded due to OBF policies. The question of which of these two outcome measures is most appropriate depends on how OBF is hypothesized to affect degrees awarded among institutions. For example, if a state allocated more appropriations via OBF to institutions that produce more bachelor's degrees in total, the total number of degrees would be a more appropriate outcome measure. In contrast, the logged value of degree production might be a better measure if the funding an institution receives is

proportionate to its increase in the number degrees awarded from the previous year or a greater graduation rate. It is difficult to capture the exact calculation used in each OBF system, so we include both outcomes to address the possible variations of the allocation methods of state appropriation under OBF.

We add a broad range of time-varying, institutional and state-level controls. At the institutional level, all control variables are collected or generated from the data of the undergraduate division. We control for number of black students per 100 students, number of Hispanic students per 100 students, Pell dollars per FTE, tuition and fees, total part-time students enrolled, enrollment of the institution, percent of institutional operational cost in terms of revenue received from state. At the state level, we include state unemployment rate and state gross domestic product per capita to capture differences in state-level economy and labor market fluctuations. Table 3 presents summary statistics for the dependent and control variables between 2005 and 2016.

Table 3

Summary Statistics, Pooled 2005 through 2016

	Institutions in OBF states (n = 2,208)		Institutions in non-OBF states (n = 3,504)	
Variable	M	SD	M	SD
Number of bachelor's degrees awarded	2062.7	2016.8	2164.7	1993.2
Percent Black	10.5	16.7	16.5	23.6
Percent Hispanic	6.5	9.7	10.0	13.8
Total fall undergraduate enrollment (1,000)	112.7	91.6	112.5	89.8
Percent part-time students	22.5	14.6	19.2	13.5
Tuition and fees ($1,000)	8.0	2.5	7.2	2.4
Pell grants per FTE ($1,000)	1.5	0.7	1.6	0.8

Percent operational cost / state revenue	31.7	11.5	35.4	11.3
Unemployment rate	6.5	2.2	6.3	2.1
State GDP per capita ($1,000)	51.3	7.8	54.5	9.4

As shown in Table 3, on average, institutions in OBF states award fewer Pell Grant dollars per FTE and enroll fewer African American or Hispanic students. These institutions on average also have higher tuition and fees. OBF states also have slightly higher unemployment rates and lower GDP per capita than do comparison states.

Analytical Models

Staggered implementation of OBF policies at the state level provides an opportunity to apply a natural experimental design to examine the effects of OBF policies on bachelor's degree production. More specifically, we use a difference-in-differences (DID) framework that identifies the causal effects of OBF policies by contrasting the change in outcomes pre- and post-implementation, for the OBF (treatment) and Non-OBF (comparison) states (Ashenfelter, 1978; Bertland, Duflo, & Mullainathan., 2004). To estimate OBF policy effects, we implemented different versions of the following regression model:

$$y_{it} = X_{it}\beta + OBF_{it}\delta + c_i + \gamma_t + (c_i \times YEAR_t)\theta_i + \epsilon_{it} \quad (1)$$

where y_{it} represents the outcome variable for institution *i* in year *t* measured as the total number of bachelor's degrees or the logged number of bachelor's degrees. The variable OBF is the treatment variable that is equal to 1 if the institution is in a state that is implementing OBF in year *t*. The vector X_{it} represents institution and state level time-varying controls. Vectors c_i represents institution fixed effects, γ_t represents year fixed effects, and ϵ_{it} is the idiosyncratic errors within and across states. This model also includes a set of interactions between state fixed

effects and a linear time variable ($c_i \times YEAR_t$). This state-specific linear time trends allow degree production to take different trajectories (θ_i) across the states during the pretreatment period (Hillman, Fryer, & Crespin-Trujillo, 2017). Thus, if the parallel trend assumption is satisfied, the introduction of these linear trends should not meaningfully affect the estimates. In equation (1), the vector of coefficients β represents the effects of observed institution and state-level time varying controls. Finally, the coefficient δ is the average OBF treatment effect for all post-OBF years.

While Model 1 is a common specification of the DID model, δ provides a very restrictive version of an OBF effect estimate. That is, δ is estimated under the assumption that the impact of OBF is constant in all periods after exposure – in other words, that the impact of OBF is temporally homogenous. As Li and Kennedy (2018) argued, δ does not provide a valid estimate for the true policy effect because the policy has a lagged effect on the outcome and/or the policy effect changes over time. While *symbolic* responses (e.g. aligning institutional strategic plans to the OBF policy and conducting internal analysis to predict institutional performance on OBF metrics) might occur quickly, substantive changes made by institutions to improve student outcomes (e.g., changes to student advising, implementing early warning systems, and increasing course availability) are likely to require additional planning and resources before full implementation. As a result, they are likely to take additional time to influence student outcomes. If so, the impact of state OBF policies would increase over time (i.e. years of implementation) as *symbolic* responses are supplemented with more *substantive* ones. To estimate these temporal effects of OBF, in Model 2 we replace the single OBF policy variable with a vector of interaction terms between the treatment dummy and a set of dummies indicating years in operation:

$$y_{it} = X_{it}\beta + \sum_{k=0}^{n} OBF_{itk}\delta_k + c_i + \gamma_t + (c_i \times YEAR_t)\theta_i + \epsilon_{it} \qquad (2)$$

where the dummy variables $OBF_{k,it}$ are equal to one if the OBF happened to unit *i* for *k* years ago. In this model the number of treatment institutions identified by $OBF_{k,it}$ vary depending on how many years a state had been implementing OBF. For example, $OBF_{it0}$is equal to 1 for all institutions in each of the 19 OBF states (see Table 5) for the first year that OBF was implemented in that state. Likewise, $OBF_{1,it}$, $OBF_{2,it}$, …., and $OBF_{n,it}$ are set to equal 1 for all institutions in a given state that operated OBF for the 2nd, 3rd, …, and (n+1)th year, respectively. Accordingly, δ_0 - δ_n indicate the impact of OBF on the outcome measure for the first operation year through (n+1)th year.

The analytical models described in Model (1) and (2) above coded the OBF treatment dichotomously under the assumption that all OBF programs implemented in different states were equally robust, which is clearly not the case empirically. While this practice is common in the extant literature, this approach does not reflect the fact that OBF policies differ in substantive, qualitative ways across different states. To reflect variances in OBF policy design across states, we redefined "OBF treatment" by following Snyder and Fox (2016)'s typology that coded OBF policies into Type I, II, III, and IV, with Type IV being the most robust form of OBF. We employ this typology as an established coding documentation of variations in policy. As mentioned prior, in absence of any longitudinal measure of OBF policy variation across states, for purposes of this analysis we assign each OBF state to the policy type assigned by Snyder and Fox in 2016. While we acknowledge the evolving nature OBF policies, we classified states with

the assumption that although incremental change is not captured across years, in any OBF year prior to 2016 the type conveys the propensity of a given policy towards that type as of 2016. We acknowledge and further discuss the limitation of this assumption in our discussion of results. We therefore modified Equation 2 by replacing the year-specific OBF dummies with vectors of interaction terms between dummies indicating different years of OBF operation and dummies indicating different levels of program intensity (i.e., Type I, Type II, Type III, and Type IV).

$$y_{it} = X_{it}\beta + \sum_{k=0}^{n1} TYPEI_{k,it}\delta_{1k} + \sum_{k=0}^{n2} TYPEII_{k,it}\delta_{2k} + \sum_{k=0}^{n3} TYPEIII_{k,it}\delta_{3k}$$
$$+ \sum_{k=0}^{n4} TYPEIV_{k,it}\delta_{4k} + c_i + \gamma_t + (c_i \times YEAR_t)\theta_i + \epsilon_{it} \quad (3)$$

In Model (3), δ_{1k}, δ_{2k}, δ_{3k}, and δ_{4k} provide estimates of OBF effects from different years of operation for Type I, II, III, and IV, respectively. Based on our conceptual framework derived from Naughton (2004) and applying Snyder and Fox (2016)'s typology of OBF policy, we hypothesize that different types of OBF policies will lead to variation in institutional responses, where we would expect to see a more substantive and prompt response from institutions (as measured by increased bachelor's degree productivity) in states with a more robust type of OBF (e.g., Type IV) than in states with a more rudimentary form of OBF (e.g., Type I).

Since Models (1) – (3) were estimated using a panel dataset on public four-year universities and colleges that were grouped into clusters (i.e., states), unobserved errors (ϵ_{it}) could be uncorrelated across states but "clustered" or "correlated" within state (Cameron & Miller, 2013; Bertrand, Duflo, & Mullainathan, 2004). The clustering problem is caused by a common, unobserved shock at the cluster level (Hansen, 2007) and, if not properly accounted

for, can lead to underestimated standard errors. Following recent advice from Abadie et al. (2017) and (Cameron & Miller, 2013), we cluster standard errors at the state level where units are assigned to the treatment. More specifically, we estimated Liang-Zeger's cluster-robust standard errors.

Results

Our empirical results are organized as follows: First, we present the results of our first two models, which follow previous research by operationalizing OBF as monolithic policy when measuring impacts on bachelor's degree productivity overall and with respect to years of implementation. Given that there has been extensive documentation of variation in the robustness of OBF policies (Snyder and Fox, 2016; Snyder and Boelscher, 2018) and our conceptual framework informed by Naughton (2004) suggests that the nature of an institution's response will vary based on the perceived legitimacy of a given OBF policy, we then compare results from Models 1 and 2 to Model 3, which estimates impacts for four types of OBF policies. Finally, we present results from robustness checks of the common trends assumption.

Estimated Impacts of OBF as a Uniform Policy on Bachelor's Degree Production

Table 4 presents the parameter estimates of the model formalized in Model (1). The coefficient on the OBF Policy variable in each column reports the average treatment effects of implementing OBF policies across all operation years on the two bachelor's degree production measures: total number of bachelor's degrees awarded and logged total number of bachelor's degrees awarded. These estimates showed that, on average, there was no statistically significant change in either of the two outcomes in response to OBF polices.

Table 4.
Average Effects of OBF Policies on Bachelor's Degree Production

Variables	Total Number of Bachelor's Degrees	Logged Total Number of Bachelor's Degrees
OBF policy	-2.565	-0.00433
	(20.44)	(0.0112)
Black students per 100 FTE	-2.931	-0.00220
	(3.414)	(0.00139)
Hispanic students per 100 FTE	6.307	0.00206
	(4.444)	(0.00239)
Total fall undergraduate enrollment	13.56***	0.00235***
	(1.351)	(0.000295)
Part-time students per 100 FTE	-6.501*	-0.00162
	(2.617)	(0.00102)
Tuition and Fees ($1,000)	33.09***	-0.00702
	(7.121)	(0.00585)
Pell students per 100 FTE	-16.86	0.0182
	(31.23)	(0.0135)
% state revenue	3.168	-0.00119
	(1.586)	(0.000688)
Unemployment rate	6.197	0.00292
	(8.186)	(0.00457)
Income per capita ($1, 000)	-4.024	-0.00230
	(3.496)	(0.00216)
Constant	17,504*	28.72***
	(7,686)	(5.400)
Observations	5,712	5,712
R-squared	0.991	0.992

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table 5 presents the DID estimates of the effect of any type of OBF on the total number of bachelor's degrees and logged total number of bachelor's degrees over time as shown in Model 2. In each of years one through four of OBF implementation, we do not find any statistically significant impact of the policy on degree production. In each of implementation years five through eight, however, OBF showed significant positive effects on both outcome measures.

Table 5.
Effects of OBF Policies on Bachelor's Degree Production by Implementation Year

Variables	Total Number of Bachelor's Degrees	Logged Total Number of Bachelor's Degrees
1st Implementation Year	-1.364	0.00544
	(13.67)	(0.00968)
2nd Implementation Year	23.14	0.00664
	(20.71)	(0.0125)
3rd Implementation Year	47.26	0.0173
	(35.41)	(0.0208)
4th Implementation Year	59.61	0.0103
	(43.86)	(0.0259)
5th Implementation Year	144.6**	0.0743*
	(45.74)	(0.0310)
6th Implementation Year	138.6*	0.0911**
	(65.93)	(0.0313)
7th Implementation Year	186.9**	0.106**
	(65.41)	(0.0373)
8th Implementation Year	182.5*	0.110*
	(76.38)	(0.0429)
Observations	5,712	5,712
R-squared	0.992	0.992

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

In addition, there is a clear positive trend in the parameter estimates, suggesting that the effect of OBF on bachelor's degree production may become more positive over time. Results showed that the total number of bachelor's degrees awarded significantly increased by 144 during the fourth year of implementation and continuously increased during the each of the subsequent years. We also found similar results in the logged total number of bachelor's degrees. During years five through eight of implementation, the bachelor's degree production for the treatment institutions increased 7.4% to 11%, which was significantly higher than increases observed in non-OBF states.

It should be noted that states in the treatment group implemented OBF in different years, meaning that the coefficient estimate for each subsequent post-treatment year is driven by fewer

treatment states than the year before it. That is, the effect in the first implementation year is estimated for all 19 OBF states, the effect of the second year is estimated for 14 states that implemented OBF policies for two or more years, and so on. Specifically, the estimated effect of OBF in the fifth year was produced by five states (i.e., Indiana, Nevada, New Mexico, Ohio, and Tennessee); the estimated effect in sixth and seventh years were driven by only two OBF states (Indiana and Ohio); and the estimated effect in the eighth year was driven solely by Indiana, as it was the first state to implement OBF.

Estimated Impacts of Distinct Types of OBF on Bachelor's Degree Production

As discussed earlier, OBF effects shown in Tables 4 and 5 were estimated under the baseline hypothesis that all OBF policies implemented in different states are uniform and equally robust, which is clearly not the case empirically. To test our alternative hypothesis that variation in the composition and robustness of OBF policies impacts bachelor's degree productivity, we examined effects across states that implemented four types of OBF polices using the analytical model described in Model (3). Table 6 shows the states that are classified under each of the four OBF policy types and the total years of policy implementation in each as of 2016. Again, it is important to note that implementation years vary across the OBF states within each Type (see Table 6), and so the estimated effects for a given type in a given implementation year are driven by the states within that type that had implemented for that given length of time.

Table 6.

Typology of OBF as of Fiscal Year 2016 (Snyder & Fox, 2016) and Years of Implementation

Typology	State (Years of Implementation)
Type I	Illinois (4), Massachusetts (3), Michigan (4), Missouri (2), North Dakota (4)
Type II	Florida (1), Indiana (8), Maine (3), Minnesota (3), New Mexico (5), Oregon (1), Utah (1) Pennsylvania[4] (4)
Type III	Arizona (3), Colorado (1), Montana (2), Nevada (5)
Type IV	Ohio (7), Tennessee (4)

Type I effects. As shown in Table 6, the five states with Type I policies (i.e., the most rudimentary type of policies) as of 2016 had implemented OBF for two to four years. Table 7 reports estimated effects of OBF policies by program type and implementation year. Results shown in Table 7 suggest that Type I OBF policies do not significantly impact bachelor's degree production in these states during the first four years of implementation. All estimated OBF effects on the total number of bachelor's degrees are positive but statistically insignificant in years one through four. Estimated effects on the log of total bachelor's degrees are also statistically insignificant with mixed signs.

Type II effects. Across the eight states with Type II policies as of 2016, the length of policy implementation varied from one to eight years. Similar to Type I policies, no significant effects were observed for Type II policies during the first two years of policy implementation. However, bachelor's degree production significantly increased in Type II states beginning in implementation year three. The Year 4 increase in total bachelor's degrees did not reach significance at the $p<.05$ level but indicates a consistently upward trend in total bachelor's

[4] PASSHE institutions only.

Table 7.
Effects of OBF Policies on Bachelor's Degree Production by Type and by Implementation Year

Year of OBF	**OBF Type I (n=696)**		**OBF Type II (n=888)**		**OBF Type III (n=348)**		**OBF Type IV (n=276)**	
	Total Number of Bachelor's Degrees	*Logged Total Number of Bachelor's Degrees*	*Total Number of Bachelor's Degrees*	*Logged Total Number of Bachelor's Degrees*	*Total Number of Bachelor's Degrees*	*Logged Total Number of Bachelor's Degrees*	*Total Number of Bachelor's Degrees*	*Logged Total Number of Bachelor's Degrees*
1st	14.89	0.00216	-24.59	-0.00540	28.74	0.0428	12.45	0.0141
	(19.21)	(0.0129)	(20.97)	(0.00895)	(52.67)	(0.0347)	(61.55)	(0.00737)
2nd	17.41	-0.00431	23.76	0.00708	57.63	0.0110	143.4**	0.0454***
	(41.75)	(0.0172)	(17.81)	(0.0149)	(52.91)	(0.0413)	(41.42)	(0.0101)
3rd	22.10	0.000273	65.74*	0.0179	57.19	-0.0113	253.7***	0.0796*
	(78.97)	(0.0382)	(29.15)	(0.0201)	(117.1)	(0.0548)	(42.95)	(0.0336)
4th	60.52	-0.0184	84.90	0.0237	158.2	-0.00531	200.5***	0.0609*
	(74.96)	(0.0450)	(45.87)	(0.0414)	(81.49)	(0.0450)	(38.36)	(0.0257)
5th			149.7**	0.108*	140.1	-0.0221	321.4***	0.116**
			(51.42)	(0.0431)	(93.06)	(0.0526)	(50.00)	(0.0351)
6th			156.7**	0.119**			379.7***	0.142***
			(47.82)	(0.0424)			(61.05)	(0.0372)
7th			213.0***	0.142**			416.0***	0.156***
			(58.12)	(0.0489)			(66.60)	(0.0410)
8th			217.5**	0.141*				
			(64.96)	(0.0558)				

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

degrees. In implementation Year 5 a statistically significant effect reemerges and extends through Year 8. As shown in Table 7, the total number of bachelor's degrees rose significantly from about 65.74 in the third year of implementation to 217.5 in year eight of implementation. The log of total bachelor's degrees awarded also increased significantly in Type II states, but not until the fifth year of operation and thereafter. More specifically, logged values for bachelor's degrees rose on average 10.8% in Year 5 to 14.1% in Year 8.

In general, positive effects became stronger over time for both outcome measures. When considering these outcomes, it is important to note that only five of the eight Type II states had implemented policies for at least three years, three states had implemented for at least four years, and two states for at least five years. Outcomes found for years six through eight reflect institutions in a single Type II state (i.e. Indiana). The difference in impacts observed between Type I and Type II OBF policies shows that we can reject the baseline hypothesis that all OBF policies are monolithic and equally robust and supports our alternative hypothesis that impacts on bachelor's degree productivity will vary by policy type, and more robust OBF policy types will yield stronger effects than more rudimentary policy types.

Type III effects. Half as many states were implementing Type III policies versus Type II policies as of 2016. Of the four states implementing Type III OBF policies, years of OBF implementation varied from one to five years (see Table 6). Overall, we found no significant effects for Type III policies on either of the bachelor's degree outcomes across all five implementation years. While the lack of significant effects supports the hypothesis that different types of OBF will produce different effects, it does not support the hypothesis that more robust types of OBF will produce more intense effects.

Type IV effects. Our analysis of the most robust OBF Type IV policy type included institutions in only two states, Ohio and Tennessee, where OBF had been implemented for seven and four years respectively. In line with our alternative hypothesis, OBF policies that were classified as the most robust also yielded the most significant and prompt effects on bachelor's degree productivity when compared to the other three types. In these two states both total number and percent of bachelor's degrees awarded significantly rose in the second year of OBF operation and the estimated effect size generally increased over time. The total number bachelor's degrees awarded increased by 143.4 in the second year of operation and consistently increased to 416 as of the seventh year. In terms of percent changes, the increase ranged from 4.5% to 15.6% during the same period.

In sum, when compared to results from Models 1 and 2, results from Model 3 suggest that different types of OBF polices yield different impacts on bachelor's degree production and suggest that we can reject the baseline hypothesis that OBF policies are monolithic and have uniform impacts. Results from Model 3 also show that years of implementation continues to be a key factor within OBF types where significant impacts were seen in Year 2 of implementation across institutions in states implementing the most robust type of OBF policies (see Table 7), but did not emerge until Year 5 when all OBF types were pooled together (see Table 5). While previous studies have found limited evidence tying OBF policies to increased outcomes, these results suggest that some types of OBF policies (i.e. Type II and Type IV) do significantly increase degree production over time, where other types (Type I and Type III) have not produced measurable impacts.

Robustness Checks

Prior to running Models 1-3, we first examined the common trend assumption (CTA), a key identifying assumption of the difference-in-difference (DID) estimation of OBF effects. As Heckman et al. (1997) note, the validity of a causal inference of estimated OBF policy effects relies on whether the treated and untreated units (institutions) would have had common trends in the absence of treatment. If there were transitory shocks prior to implementation of OBF policies that affected the treated and untreated units differently, such events during pre-OBF years may bias DID estimates of OBF effects. We conducted an event-history analysis to examine the presence or absence of consistent pre-OBF trend in Bachelor's degree production (i.e., total number of bachelor's degrees) between treated and untreated institutions by following an empirical strategy in the displaced worker literature, as in Jacobsen et al. (1993) and Stevens (1997); the school closings literature, as in Brummet (2014); and the performance funding literature, as in Li and Kennedy (2018). That is, we estimated the following even-history model:

$$y_{it} = X_{it}\beta + \sum_{h=-m}^{-1} PreOBF_{h,it}\omega_h + \sum_{k=0}^{n} OBF_{k,it}\delta_k + c_i + \gamma_t + \epsilon_{it} \qquad (4)$$

where y_{it} is the total number of bachelor's degrees awarded. This model includes dummy variables identifying treatment institutions for each of both pre- and post-OBF years. Coefficients on the pre- and post OBF dummies (ω_h and δ_k) measure differences in the means of a given outcome between the treatment and comparison institutions in each of pre- or post OBF operation years, controlling for observed institution and state time-varying covariates (X_{it}) and institution (c_i) and time fixed effects (γ_t).

Finding evidence that the pre-period differences in bachelor's degree production between the treatment and comparison institutions did not vary significantly in pre-OBF years or not rejecting H_o: $\omega_{-1} = \omega_{-2} = \omega_{-3} = \cdots = \omega_{-m}$ will be supportive of the common pre-trend

assumption up to "m^{th}" pre-OBF year. The full parameter estimates of this event-history model are reported in Appendix A, and step-by-step joint hypothesis test results up to 11th pre-OBF year are reported in Table 8 below.

Table 8.
Joint Hypothesis Test Results on the Common Pre-OBF Trend on Bachelor's Degree Production

Joint Hypothesis of Common Trend	F-statistic	*p*-value
Pre-Year 1 & Pre-Year 2 (H_o: ω_{-1}= ω_{-2})	0.34	0.5624
Pre-Year 1 through Pre-Year 3 (H_o: ω_{-1}= ω_{-2}= ω_{-3})	0.31	0.7351
Pre-Year 1 through Pre-Year 4 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-4}$)	0.33	0.8047
Pre-Year 1 through Pre-Year 5 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-5}$)	0.38	0.8207
Pre-Year 1 through Pre-Year 6 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-6}$)	0.50	0.7736
Pre-Year 1 through Pre-Year 7 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-7}$)	0.67	0.6729
Pre-Year 1 through Pre-Year 8 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-8}$)	0.92	0.4961
Pre-Year 1 through Pre-Year 9 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-9}$)	0.87	0.5515
Pre-Year 1 through Pre-Year 10 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-10}$)	1.05	0.4290
Pre-Year 1 through Pre-Year 11 (H_o: $\omega_{-1} = \omega_{-2} = \cdots = \omega_{-11}$)	4.44	0.0002

Joint hypothesis test results reported in Table 8 suggest that pre-OBF trends in bachelor's degree production between the institutions in OBF states and their comparison institutions were stable from pre-OBF years 1 through 10, strongly supporting the common pre-trend assumption of our difference-in-difference model during these pre-years. However, when a dummy variable for the 11th pre-OBF year was added the joint hypothesis for common pre-OBF trends was rejected. Note that the coefficient on this dummy (ω_{-11}) captures the average difference between the number of bachelor's degrees awarded in Year 2005 (the first year of the 12-year panel data used for this study) by institutions from the four OBF states that first implemented OBF policies in Year 2016 (i.e., Florida, Colorado, Utah, and Oregon) and the number of

bachelor's degrees awarded in Year 2005 by all comparison institutions in non-OBF states[5]. Since the introduction of the four most recent OBF states to the sample led to rejecting the joint hypothesis that is supportive of CTA, we re-estimated the main impact models after excluding these four states to explore how the departure from CTA would have potentially affected OBF impact estimates.

The results from the common trend assumption analysis indicates that a common trend can be established for the first through the tenth pre-implementation years but not across all eleven pre-implementation years among all states. To account for the possibility that the latest implementers (Colorado, Florida, Oregon, and Utah) are different in state and institutional characteristics which might lead to differences in the effectiveness of OBF policies, we examined the same models (1) through (3) without these four states. The findings show no changes in the average treatment effects when we examined all OBF policies together (Model 1); a significant and positive effect showed one year earlier when we examined the effects by different lengths of implementation; and no changes in the effects of Type I and Type IV policies, but significant and positive effects of Type II and Type III policies emerged one year earlier. Overall, we do not find significant changes in the findings when we excluded the most recent implementers[6].

Conclusion and Policy Implications

In this study we first examine the impacts of OBF following the assumptions of previous studies by measuring the impacts of any type of OBF on bachelor's degrees overall and over time. We then test an alternative hypothesis that OBF policies vary, and that more robust types of

[5] Estimated coefficients on dummies for pre-OBF year 1 through pre-OBF year 10 ranged from 89.0 to 114.3 and statistically no different from each other. But, the coefficient on the pre-OBF year 11 dummy reduced to -7.8 which resulted in the rejection of the last joint hypothesis in Table 8.

[6] The results our robustness check of the three models are available upon request.

OBF policies would be expected to elicit more substantive institutional responses which would lead to more substantial changes in outcomes over a shorter implementation period. Specifically, we examine the impacts of four types of OBF policies on two measures of bachelor's degree productivity over time. Overall, our findings suggest that OBF policy type influences policy effectiveness as well as the timing of impacts, and therefore, have important implications for state policymakers and researchers.

For 19 states that implemented any version of OBF, our findings showed that on average, OBF has no significant impact on the total number of bachelor's degree awarded (see Table 4). Though not significant, the coefficients for the two outcomes showed a mixed trend: positive for total bachelor's degree production but negative for the logged value. These mixed results and insignificant findings are aligned with previous literature where researchers assume a homogenous effect of OBF across states regardless of the staggered implementation timeline and variations in policies (Tandberg, Hillman, & Barakat, 2014; Tandberg & Hillman, 2014; Hillman, Tandberg, & Fryar, 2015).

In our second Model we account for the temporal heterogeneity in policy implementation across states to determine whether the effects of *any* type of OBF vary based on years of implementation. Our findings show a lag between policy implementation and policy effects on bachelor's degree productivity, where positive impacts only begin to emerge after five years of policy implementation (see Table 6). These results further substantiate findings from earlier studies that also accounted for temporal heterogeneity (Tandberg, Hillman, & Barakat, 2014; Tandberg & Hillman, 2014), and support the conclusion that policy effectiveness is associated with length of policy implementation.

The unique contribution of this study, however, was to account for OBF policy differences that might affect institutional response in addition to temporal heterogeneity. Specifically, our results show that OBF impacts varied substantially by policy type, with significant positive impacts on bachelor's degree productivity for institutions in states with Type II and Type IV policies, but no measurable impacts on bachelor's degree productivity for institutions in states with Type I and Type III policies. Overall, we found that the most robust type of OBF policies (i.e. Type IV) produced the largest impact estimates of any type. In addition, our examination of the effect of the longevity of policy implementation show significant impacts for Type IV on both measures of bachelor's degree productivity as early as the second year of implementation, with steadily increasing effects through all subsequent years of implementation. When we examine the impacts for Type II (i.e. a less robust type) significant positive impacts for both measures of bachelor's degree productivity emerge as well, but not until the fifth year of implementation (see Table 7).

The impact of Type IV policies may emerge earlier because they are hypothesized to garner higher perceived legitimacy among institutions (Dougherty & Reddy, 2013; Snyder &Fox, 2016; Hillman, Fryar, & Crespín-Trujillo, 2017). These findings support our hypothesis that the level of perceived legitimacy conveyed by substantive resources can influence both the magnitude of the effects produced by OBF policies and how long it takes us to observe such effects. For state policymakers considering the adoption of OBF, these results emphasize the importance of designing a robust policy for both the magnitude of policy impacts and the expediency of institutional response. Moreover, policymakers in states with more rudimentary (Type I) policies might consider refining existing policies to include components of more robust (Type IV) policies.

While the results for Type IV policies clearly align with our hypothesis, the significant impacts found for Type II policies and non-significant effects for Type III policies do not. One explanation could be related to the specific components that distinguish Type II and Type III policies. For example, one key component that Snyder and Fox (2016) use to distinguish Type II policies from Type III policies is the relative percentage of state postsecondary funding tied to OBF, where Type II policies have percentages of funding less than five percent and Type III policies award between 5% and 24.9% of state postsecondary funding. Based on this component, Indiana is classified as having a Type II policy in 2016. Notably, Indiana's OBF policy had been implemented for more consecutive years (i.e. 8 years) than any other policy in our sample. While Snyder and Fox (2016) highlight formula stability as a component associated with more robust policies in their typology, this component was not weighted as heavily as percentage of funding when assigning 2016 types. Therefore, future research should consider empirically testing the relative effectiveness of different policy components such as percentage of funding and policy stability to provide more precise guidance for defining the components of robust policies and distinguishing policy types.

Another key factor to consider is how specific state contexts influence the perceived legitimacy of different OBF policy components. For example, previous research on OBF policy implementation in Indiana found that institutional leaders perceive the percentage of state funding affected by Indiana's policy to be substantial enough to warrant response (Callahan et al. 2017a). This research also found that institutional leaders in Indiana did not associate additional policy legitimacy with a higher percentage of state funds attached to OBF, but viewed the existing percentage as appropriate given their state context. In a recent analysis of state-specific responses to the National College Completion Agenda, Rubin and Hearn (2018) emphasize the

importance of states crafting completion policies that fit their state context. Thus, when crafting robust OBF policies it is important for policymakers to consider the main components of robust OBF policies, their urgency for policy impacts, as well as the perceived legitimacy each of these factors will have with institutions given the postsecondary landscape in their states. More nuanced research that examines the relative impacts of the key components of robust OBF polices in different state contexts could benefit the policy design process.

Limitations

While this study provides the first analysis of how diverse types of OBF policies and implementation timelines impact bachelor's degree production over time, it is not without limitations. First, like Li and Kennedy (2018) we utilize the Snyder and Fox (2016) typology to assign policy types to states. Since Snyder and colleagues began assigning OBF policy types in 2015, no types are available for earlier years of policy implementation. Because of the lack of retrospective data we applied the 2016 type to all previous years of implementation built upon the assumption that as of 2016, all OBF states are classified in corresponding types and from the first year of implementation all policies are in the process of becoming their 2016 type. Undoubtedly policy changes have been happening throughout all years and across states and this likelihood calls for extensive research and continued conversation between practitioners and researchers to establish a more accurate reflection of the nuance of each state's policy. In order to more precisely track the impacts of OBF policy types and identify the specific policy components that are most critical to shifting outcomes, the fields needs a comprehensive data set that retroactively tracks each OBF policy and its key components across all years of implementation.

In addition, while IPEDS data support rigorous cross-state comparisons, available institution-level data do not allow us to further disentangle the variations in policy impacts for focus populations identified in many policies (e.g., state residents, Pell grant recipients, adults, underrepresented minorities, etc.). Research that examines the effects of OBF on focus populations should consider student-level data such as Statewide Longitudinal Data System for a more nuanced look on whether these positive impacts are closing equity gaps.

Appendix A

Event History Model Results

VARIABLES	Total Number of Bachelor's Degree
pre_obf1_hcm	98.34
	(338.0)
pre_obf2_hcm	105.8
	(335.3)
pre_obf3_hcm	113.0
	(334.3)
pre_obf4_hcm	106.8
	(333.6)
pre_obf5_hcm	95.21
	(332.1)
pre_obf6_hcm	105.3
	(333.2)
pre_obf7_hcm	89.00
	(337.3)
pre_obf8_hcm	98.44
	(338.6)
pre_obf9_hcm	114.3
	(343.8)
pre_obf10_hcm	100.6
	(361.5)
pre_obf11_hcm	-7.820
	(386.5)
post_obf0_hcm	90.72
	(341.3)
post_obf1_hcm	108.2
	(337.3)
post_obf2_hcm	147.2
	(331.1)
post_obf3_hcm	144.9
	(333.3)
post_obf4_hcm	200.6
	(333.5)
post_obf5_hcm	212.8
	(328.7)
post_obf6_hcm	234.0
	(333.5)
post_obf7_hcm	157.4
	(348.2)
Observations	5,712
R-squared	0.990

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

References

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *The Review of Economics and Statistics, 60*(1), 47-57. doi:10.2307/1924332

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2017). *When Should You Adjust Standard Errors for Clustering?* doi:10.3386/w24003

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. *The Quarterly Journal of Economics, 119*(1), 249-275.

Bound, J., Lovenheim, M. F., & Turner, S. (2012). Increasing time to baccalaureate degree in the United States. *Education Finance and Policy*, *7*(4), 375-424.

Burke, J. C. (Ed.). (2002). *Funding public colleges and universities for performance: Popularity, problems, and prospects.* Albany, NY: Rockefeller Institute Press.

Burke, J. C., & Minassians, H. P. (2003). *Performance reporting: "Real" accountability or accountability "lite": Seventh Annual Survey 2003.*New York City: Nelson A. Rockefeller Institute of Government, State University of New York.

Callahan, M. K, Meehan, K., Shaw, K. M., Kim D. Y., Hunter V. R., Lin J., Slaughter A., and Leonard W. (2017a). *Implementation and impact of outcomes-based funding in Indiana. (Research Brief).* Philadelphia, PA: Research for Action. Available: https://8rri53pm0cs22jk3vvqna1ub-wpengine.netdna-ssl.com/wp-content/uploads/2017/03/RFA-OBF-in-Indiana-Full-Brief-February-2017-updated.pdf. Accessed 15 January 2017.

Callahan, M. K, Meehan, K., Shaw, K. M., Kim D. Y., Hunter V. R., Lin J., Slaughter A., and Leonard W. (2017b). *Implementation and impact of outcomes-based funding in Tennessee. (Research Brief).* Philadelphia, PA: Research for Action. Available: https://aefpweb.org/sites/default/files/webform/42/A%20Mixed%20Methods%20Examination%20of%20OBF%20in%20TN%20-%20Callahan%20et%20al..pdf. Accessed 15 March 2017.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372.

de Lancer Julnes, P., and M. Holzer. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*. 61 (6).

DiMaggio, P. J., & Powell W. W. (1991). *The new institutionalism in organizational analysis. Chicago,* IL: University of Chicago Press.

Dougherty, K. J., & Natow, R. S. (2015). *The politics of performance funding for higher education: Origins, discontinuations, and transformations.* Baltimore, MD: Johns Hopkins University Press.

Dougherty, K. J., Jones, S. M., Lahr, H., Natow, R. S., Pheatt, L., & Reddy, V. (2014). Performance funding for higher education: Forms, origins, impacts, and futures. *The ANNALS of the American Academy of Political and Social Science, 655*(1), 163-184.

Daugherty, K. J., & Reddy, V. (2013). *Performance funding for higher education: What are the mechanisms. What are the impacts*. [Working Paper] Community College Research Center. Available: http://onlinelibrary.wiley.com/doi/10.1002/aehe.v39.2/issuetoc. Accessed 01 December 2016.

Gandara, D., Rippner, J. A., & Ness, E. C. (2017). Exploring the 'how'in policy diffusion: National intermediary organizations' roles in facilitating the spread of performance-based funding policies in the states. *The Journal of Higher Education, 88*(5), 701-725.

Hansen, B. E. (2007). Least squares model averaging. *Econometrica*, *75*(4), 1175-1189.

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies,64*(4), 605-654. Retrieved from http://www.jstor.org/stable/2971733

Hearn, J. C. (2015). *Outcomes-based funding in historical and comparative context.* (Lumina Issue Papers). Athens, GA: Lumina Foundation. Available: http://files.eric.ed.gov/fulltext/ED561322.pdf. Accessed 02 August 2016.

Hillman, N. W., Tandberg, D. A., & Fryar, A. H. (2015). Evaluating the impacts of "new" performance funding in higher education. *Educational Evaluation and Policy Analysis*, *37*(4), 501-519.

Hillman, N. W., Hicklin Fryar, A., & Crespín-Trujillo, V. (2018). Evaluating the impact of performance funding in Ohio and Tennessee. *American Educational Research Journal*, *55*(1), 144-170.

Jacobsen, F. M., Mueller, E. A., Rosenthal, N. E., Rogers, S., Hill, J. L., & Murphy, D. L. (1994). Behavioral responses to intravenous meta-chlorophenyl-piperazine in patients with seasonal affective disorder and control subjects before and after phototherapy. *Psychiatry Research*, *52*(2), 181-197.

Li, A.,Y, & Zumeta, W. M. (2016). *Performance funding on the ground: Campus responses and perspectives in two states.* Research Dialogue, 128, Available: https://www.tiaainstitute.org/sites/default/files/presentations/2017-02/rd_performance_funding_on_the_ground.pdf. Accessed 22 January 2017.

Li, A. Y., & Kennedy, A. I. (2018). Performance funding policy effects on community college outcomes: Are short-term certificates on the rise?. *Community College Review*, *46*(1), 3-39.

McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Education Evaluation and Policy Analysis*, *9*(2), 171-178.

McLendon, M. K., Hearn, J. C., & Deaton, R. (2006). Called to account: Analyzing the origins and spread of state performance-accountability policies for higher education. *Educational Evaluation and Policy Analysis*, *28*(1), 1-24.

Meehan, K., Callahan, M. K., Stull-Hunter, G. (2017). Adoption versus implementation of outcomes-based funding in three states. Paper presented at the annual meeting of the Association for the Study of Higher Education, Houston, TX.

National Conference of State Legislatures. (2016). *Performance-based funding for higher education.* Available: http://www.ncsl.org/research/education/performance-funding.aspx. Accessed 20 December 2016.

Naughton, B. A. (2004). *The efficacy of state higher education accountability programs*. Unpublished Paper, Stanford University.

North, D. C. (1990). Institutions. *Journal of Economic Perspectives*, *5*(1), 97-112.

Pew Research Center. (2011). *Is college worth it? College presidents, public assess value, quality and mission of higher education.* Washington D.C. Available: http://www.pewresearch.org/wp-content/uploads/sites/3/2011/05/higher-ed-report.pdf

Rabovsky, T. M. (2012). Accountability in higher education: Exploring impacts on state budgets and institutional spending patterns. *Journal of Public Administration Research and Theory*, *22*(4), 675-700.

Rutherford, A., & Rabovsky, T. (2014). Evaluating impacts of performance funding policies on student outcomes in higher education. *The ANNALS of the American Academy of Political and Social Science*, *655*(1), 185-208.

Sanford, T., & Hunter, J. M. (2011). Impact of performance funding on retention and graduation rates. *Education Policy Analysis Archives*, *19*, 33.

Shin, J. C. (2010). Impacts of performance-based accountability on institutional performance in the US. *Higher Education*, *60*(1), 47-68.

Shin, J. C., & Milton, S. (2004). The Effects of Performance Budgeting and Funding Programs on Graduation Rate in Public Four-Year Colleges and Universities. *Education Policy Analysis Archives*, *122*(2): 1-27.

Snyder, M. (2015). Driving better outcomes*: Typology and principles to inform outcomes-based funding models.* Washington, DC: HCM Strategist. Available: http://hcmstrategists.com/drivingoutcomes/wp-content/themes/hcm/pdf/Driving%20Outcomes.pdf. Accessed 03 April 2016.

Snyder, M, & Fox, B. (2016). *Driving better outcomes: Fiscal year 2016 state status and typology update.* Washington, DC: HCM Strategists. Available: http://hcmstrategists.com/wp-content/uploads/2016/03/2016-Report.pdf. Accessed 09 September 2016.

Snyder, M., & Boelscher, S. 2018. *Driving Better Outcomes: Fiscal Year 2018 State Status & Typology Update.* Washington, DC: HCM Strategists. Available: http://hcmstrategists.com/resources/driving-better-outcomes-fiscal-year-2018-state-status-typology-update/. Accessed 14 March 2019

Stevens, A. 1997. Critical notice. *Journal of Analytical Psychology*, *42*: 671-690. doi:10.1111/j.1465-5922.1997.00671.x

Tandberg, D. A., & Hillman, N. W. (2014). State higher education performance funding: Data, outcomes, and policy implications. *Journal of Education Finance*, 222-243.

Tandberg, D. A., Hillman, N., & Barakat, M. (2014). State Higher Education Performance Funding for Community Colleges: Diverse Effects and Policy Implications. *Teachers College Record*, *116*(12): 1-31.

Weingast, B. R. (2002). *Rational choice institutionalism. In Political science: The state of the discipline, edited by Ira Katznelson and Helen Milner*, pp. 660-692. New York City, NY: W. W. Norton.

Zumeta, W. M. (1998). *Public university accountability to the state in the late twentieth century: Time for a rethinking?* Review of Policy Research, 15(4): 5-22. doi: 10.1111/j.1541-1338.1998.tb01088.x.

Zumeta, W. M. (2011). What does it mean to be accountable?: Dimensions and implications of higher education's public accountability. *The Review of Higher Education*, *35*(1), 131-148.