ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 20-270

Incentivizing Equity? The Effects of Performance-Based Funding on Race-Based Gaps in College Completion

Monnica Chan
University of Massachusetts Boston

Zachary Mabel
Georgetown University

Preeya Pandya Mbekeani
Brown University

Performance-based funding models for higher education, which tie state support for institutions to performance on student outcomes, have proliferated in recent decades. Some states now tie most of their higher education appropriations to completion outcomes and include bonus payments for historically underrepresented groups to address equity gaps in postsecondary attainment. Using a Synthetic Control Method research design, we examine the heterogenous impact of these funding regimes in Tennessee and Ohio on completion outcomes for racially minoritized students and students from historically overrepresented racial groups. Across both states, we generally estimate null or negative effects on credentials conferred to racially minoritized students and null or positive effects on credentials conferred to students from historically overrepresented racial groups. As a result, we find that performance-based funding policies widened the racial gap in certificate completion in Tennessee and in baccalaureate degree completion in Ohio. Across both states, the estimated impacts on associate degree outcomes are also directionally consistent with performance-based funding exacerbating racial inequities in associate degree attainment.

VERSION: July 2022

Suggested citation: Chan, Monnica, Zachary Mabel, and Preeya Pandya Mbekeani. (2022). Incentivizing Equity? The Effects of Performance-Based Funding on Race-Based Gaps in College Completion. (EdWorkingPaper: 20-270). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/g53d-5105

Incentivizing Equity? The Effects of Performance-Based Funding on Race-Based Gaps in College Completion

Monnica Chan
University of Massachusetts Boston

Zachary Mabel
Georgetown University

Preeya Pandya Mbekeani
Brown University

March 22, 2022

We are grateful to Felipe Barrera-Osorio, Denisa Gándara, Joshua Goodman, Nicholas Hillman, Larry Katz, Robert Kelchen, Bridget Long, Nozomi Nakajima, two anonymous reviewers, and participants at the 2021 SHEEO Convening on Public Higher Education Funding, 2019 Association for Education Finance and Policy, Society for Research on Educational Effectiveness, and Association for Public Policy and Management annual research conferences for helpful comments and suggestions on earlier versions of this paper. We also thank staff at HCM Strategists, the Ohio Department of Higher Education, State Higher Education Executive Officers Association, Tennessee Board of Regents, and Tennessee Higher Education Commission for sharing their insights into state funding formula design and implementation. The views in this paper do not reflect those of the University of Massachusetts-Boston, Georgetown University, or Brown University. All errors, omissions, and conclusions are our own.

Abstract

Performance-based funding models for higher education, which tie state support for institutions to performance on student outcomes, have proliferated in recent decades. Some states now tie most of their higher education appropriations to completion outcomes and include bonus payments for historically underrepresented groups to address equity gaps in postsecondary attainment. Using a Synthetic Control Method research design, we examine the heterogenous impact of these funding regimes in Tennessee and Ohio on completion outcomes for racially minoritized students and students from historically overrepresented racial groups. Across both states, we generally estimate null or negative effects on credentials conferred to racially minoritized students and null or positive effects on credentials conferred to students from historically overrepresented racial groups. As a result, we find that performance-based funding policies widened the racial gap in certificate completion in Tennessee and in baccalaureate degree completion in Ohio. Across both states, the estimated impacts on associate degree outcomes are also directionally consistent with performance-based funding exacerbating racial inequities in associate degree attainment.

Keywords: accountability, equity, higher education, policy analysis, synthetic control

Published Citation & Version:
Chan, M., Mabel, Z., & Mbekeani, P.P. (2022). Incentivizing Equity? The Effects of Performance-Based Funding on Race-Based Gaps in College Completion. *The Journal of Higher Education*. https://doi.org/10.1080/00221546.2022.2082762

Incentivizing Equity? The Effects of Performance-Based Funding on Race-Based Gaps in College Completion

Introduction

The economic return to completing college is large and increasing (Autor, 2014; Avery & Turner, 2012; Carnevale et al., 2016), yet economically disadvantaged and racially minoritized students are significantly less likely to graduate than their high-income and White peers and these disparities have widened over time (Bailey & Dynarski, 2011; Chetty et al., 2017). With social mobility declining in the United States and the payoff to college credentials increasing, raising completion rates among historically underrepresented populations is imperative to creating equitable opportunities for economic prosperity.

Over the last twenty years, state policymakers and higher education advocacy organizations have turned to performance-based funding (PBF) as one strategy to increase degree attainment. PBF policies tie state support for public higher education institutions to performance on student outcomes, such as year-to-year persistence and degree completion. PBF policies first gained popularity in the 1990s and were often structured as bonus payouts on top of base funding appropriated to colleges. The weak effects and unintended consequences of past policies informed the development of a new generation of policies in the 2010s (Dougherty & Natow, 2015). These policies, commonly dubbed “PBF 2.0”, typically determine base funding levels (rather than bonus payments), tie funding to attainment outcomes, and award premiums for positive outcomes among historically underserved students, including racially/ethnically minoritized, older adult, and low-income students (Dougherty & Natow, 2015). As of 2020, 41 states had ever adopted PBF funding models (Ortagus et al., 2020), of which more than half prioritized performance of underserved student groups through "equity provisions".[1]

Two early and vigorous adopters of PBF 2.0 policies with equity provisions were Tennessee and Ohio. Early iterations of PBF in both states followed the bonus payment structure popular in decades past. However, in the 2010s, both states began tying the majority of state appropriations to institutional degree performance. They also included additional weights in their new funding formulas for graduating students from traditionally underserved groups.

We examine the extent to which degree completion trends shifted for different student subgroups after the adoption of the PBF 2.0 policies in Tennessee and Ohio. Prior research suggests that the adoption of PBF 2.0 in these states did not increase bachelor's degree completion overall and may have shifted students at two-year colleges to earn certificates instead of associates degrees (Hillman, et al., 2018; Ward & Ost, 2019). We are aware of no studies, to date, that analyze whether the new PBF models in Tennessee and Ohio differentially impacted credential completion among majority and minoritized student groups or reduced race-based completion gaps. Understanding this is important for policymakers interested in leveraging these policies towards more equitable educational outcomes in the context of a racialized higher education system.

This study contributes to the literature on the effects of PBF policies by examining two questions:

i) Did the adoption of PBF 2.0 policies in Tennessee and Ohio impact the number of certificates and degrees awarded to underrepresented racially minoritized (URM) students?[2]

---

[1] Authors' calculation based on prior reports, articles, and own research.
[2] We define URM as Black, Latino, or American Indian/Alaskan Native students and non-URM as White or Asian students. Although students from some Asian subgroups are also underrepresented in higher education, we are

ii) Did adoption reduce completion gaps between URM and non-URM students?

To answer these questions, we employ the Synthetic Control Method (SCM), a popular econometric strategy for comparative case studies (Abadie & Gardeazabal, 2003; Abadie, Diamond & Hainmuller, 2010; 2015; Peri & Yasenov, 2018). Using this approach and Integrated Postsecondary Education Data System (IPEDS) data on public colleges and universities operating in the United States, we compare changes in the number of credentials conferred to non-URM versus URM students after the adoption of PBF 2.0 in Tennessee and Ohio to "synthetic" counterfactual states. We construct comparison groups separately for Tennessee and Ohio using the set of states that did not utilize PBF policies from 2004-2015 based on our review of the literature and state policy documents.[3,4]

Across both states, we find evidence that PBF 2.0 differentially impacted credential production by race. In general, our findings point to null or negative effects on credentials conferred to URM students and null or positive effects on credentials conferred to non-URM students relative to the trends in the synthetic states. PBF 2.0 widened the racial gap in certificate completion in Tennessee such that, by 2015, the share of less-than-two-year certificates awarded to URM students declined by 9.1 percentage points relative to the counterfactual. PBF 2.0 also widened the racial gap in baccalaureate degree completion in Ohio; it increased the number of bachelor's degrees awarded to non-URM students by 10-15 percent per year between 2011- 2015 but had no impact on degrees awarded to URM students. Across both states, the estimated impacts

---

unable to disaggregate the data by Asian subgroup across all years of our analysis. We exclude students of "other" racial categories, comprised of students of two or more races, non-resident aliens, and unknown race, because we are unable to determine URM status. Our results do not meaningfully change when we restrict the non-URM category to White students or when the share of degrees awarded to URM students is defined relative to all other students, rather than White and Asian students.

[3] We index school years by the fall term throughout the paper (e.g., 2009 refers to the 2009-10 school year).

[4] We used multiple sources to identify states that did not utilize PBF. We describe our approach in Appendix A.

on associate degree outcomes are not statistically significant but are directionally consistent with widening racial inequities in associate degree attainment following PBF 2.0 adoption.

While many states are actively modifying their PBF policies, our results suggest that allocating the majority of state appropriations to colleges based on performance outcomes may be insufficient to close race-based attainment disparities and shift systems that have underserved certain student populations, even when the formulas include bonus payments, as they do in Tennessee and Ohio. Our findings suggest that public institutions in those states may have focused on increasing attainment overall to avoid or minimize funding losses instead of prioritizing completion among historically underserved groups. These formulas perpetuated, and, in some cases exacerbated, race-based attainment disparities.

Background

Performance Management and Equity

PBF models in higher education are grounded in principal-agent (PA) theory. According to this framework, state governments ("principals") seek to hold higher education institutions ("agents") accountable for performance across a range of student outcomes. Performance monitoring and financial incentives/losses form the key policy levers and are intended to mitigate gaming and resistance by impacted institutions. In practice, however, other obstacles, including lack of resources and lack of alignment between institutional missions and PBF metrics, may also arise (Dougherty & Natow, 2020).

Early iterations of PBF policies did not consider equity, instead focusing on efficiency and outcomes (Dougherty & Natow, 2015; Jones et al., 2017). Critics and scholars advocated that these policies could be more effective and better attend to equity by including premia for underserved groups and differentiated goals and benchmarks for institutions (Jones et al., 2017;

Zerquera & Ziskin, 2020). However, even with these features, the approach to equity is still cast in the framework of outputs and incentives. More generally, the race evasiveness of PA theory may limit its ability to consider and mitigate the racialized impacts of PBF policies. We consider PBF in the context of racialized public higher education systems (Dowd & Bensimon, 2015; Rodriguez et al., 2021), to generate hypotheses about the impacts of PBF on race-based attainment gaps.

PBF policies may exacerbate race-based completion gaps through effects within institutions. Colleges may respond to these policies by enrolling or shifting resources towards historically advantaged students who, on average, have a higher likelihood of completion (Birdsall, 2018; Dougherty, et al., 2016; Gándara & Rutherford, 2020; Kelchen & Stedrak, 2016; Umbricht, et al., 2017). In Tennessee, administrators have reported that equity premia were too small to increase attainment among historically underrepresented student groups (Ness et al., 2015). By pursuing the easiest path to "improve" in the near-term, institutions may exacerbate race- and class-based gaps. Equity premia may mitigate such "cream skimming" efforts (Kelchen, 2018) but have had null to small enrollment effects among underrepresented student groups in prior studies (Gándara & Rutherford, 2018; Kelchen, 2018).

PBF policies may also exacerbate completion gaps due to longstanding inequalities between institutions that serve historically disadvantaged versus advantaged students. For example, PBF policies may exacerbate existing resource disparities across institutions (Favero & Rutherford, 2020; Hagood, 2019; Jones et al., 2017). In Tennessee and Ohio, the adoption of new PBF policies resulted in minority-serving institutions (MSIs) losing $800 and $1,200 in annual state appropriations per FTE, respectively (Hillman & Coral, 2018). Broad-access institutions, which serve disproportionate shares of minoritized and low-income students, also have a long history of being funded at much lower levels than selective institutions (Davies & Zarifa, 2012). These institutions, along with MSIs, have not been afforded the foundational resources and organizational capacity to achieve the intended aims of PBF policies . Lastly, mission differences between institutions that are correlated with the racial composition of the student body may lead to differential burdens and frictions arising from these seemingly "neutral" policies (Ziskin et al., 2018; Zerquera & Ziskin, 2020). In summary, even in states that have adopted PBF policies with generous equity provisions, race-based attainment gaps may remain unchanged or worsen because PBF policies do not directly address the stratification and racialization of public higher education in the United States.

Performance Funding in Tennessee and Ohio

Tennessee and Ohio offer the best opportunity to examine the potential for PBF with equity provisions to reduce disparities in college outcomes. They are the only states to have allocated most state funding via PBF by 2010, which allows a long enough time horizon to examine impacts on college completion, particularly in the four-year sector.[5] Both states have also embraced equity provisions that reward public institutions with funding premiums for increasing persistence and degree completion for adult, racially minoritized, and/or low-income students. The PBF 2.0 models in Tennessee and Ohio have been studied extensively (e.g. Dougherty, et al., 2016; Hillman, et al., 2018; Kelchen, 2018; Li & Ortagus, 2019; Ward & Ost, 2019), but we are unaware of studies that examine the evolution of race-based attainment gaps after the adoption of the new PBF regimes.

[5] North Dakota, Nevada and Missouri also tied most state funding to institutional performance, but not until 2013.

In 2009, Ohio outlined new formulas for its public university and community college sectors focused on student course and degree completion.[6] For four-year institutions, the new system rewards degree production and equity across underrepresented groups in degree outcomes (Morris, 2009; Zumeta & Li, 2016). The share of funds awarded to four-year institutions based on student progression and completion is large (80-100 percent), and the new formula places greater weight on degree completion than previously. However, to smooth the transition to PBF 2.0, Ohio capped annual funding losses for four-year institutions at 3-5 percent of prior year levels between 2009 and 2013.

The PBF 2.0 formula in Ohio also awards premiums to four-year institutions based on the progression and completion of four historically underrepresented student populations: adult, low-income, academically underprepared, and racially minoritized students.[7] The specific premium amounts have varied across groups and over time, but across most treated years in this study (2009-2015), four-year institutions could receive an additional 30-40 percent in funding for historically underrepresented student performance. Bonus payments are also stackable for students who fit into more than one premium category.[8]

At two-year institutions in Ohio, the PBF 2.0 formula awards funding based on academic progress, completion, and transfers. Between 2009 and 2015, the share of funds awarded for student progress and completion increased incrementally as the state transitioned away from its historical, enrollment-based funding model. During the transition, Ohio allocated funding to

[6] We define the year in which the Tennessee and Ohio legislatures passed new PBF policies as the first treated year, rather than the first year of implementation. This allows us to capture anticipatory effects that may have arisen as colleges prepared for funding changes. Our results for two-year institutions in Ohio are robust to defining 2013 as the first treated year, which corresponds to when the legislature articulated the equity provisions for community colleges.
[7] Ohio defines racially minoritized students as Native American, African American, or Hispanic (Turocy & Perch, 2013).
[8] For example, bonus payments for students who were low-income and URM were approximately 60-80 percent over the period we study.

institutions based on the number of students that completed courses, earned 15 and 30 semester credit hours, earned an associate degree, and completed developmental math courses (Turocy, 2013; Ohio Board of Regents, 2013). The state increased its share of community college funding based on these criteria to 50 percent in 2013 (Boelscher & Snyder, 2019).

Ohio introduced additional features into its PBF 2.0 model for two-year institutions in 2014. The state began awarding premiums for the academic progression and degree completion of adult, low-income, and racially minoritized students. Funding for completion of certificate programs over 30 credit hours (equivalent to programs that exceed one year) was also introduced in 2014. As with the four-year sector, a stop-loss provision initially capped annual funding losses at 3-5% of prior year funding levels and was phased out in 2014.

Tennessee passed the Complete College Tennessee Act (CCTA) in 2010, which overhauled its funding formula and ended enrollment-based funding. Like in Ohio, the new PBF 2.0 system awards most funding (80-100 percent) to institutions based on student progression, degree production, and efficiency. Both two- and four-year institutions are included in the system, but they are subject to different metrics and weights. Whereas Ohio implemented stop-loss provisions, Tennessee initially made additional funds available to institutions to protect against severe financial losses resulting from the shift away from enrollment-based appropriation levels.

Tennessee also offers equity-based funding premiums to institutions under the PBF 2.0 regime. Specifically, since 2010, adult students and students from low-income families have garnered a 40 percent premium. The state did not include premiums for URM students specifically over the study period, but adult and low-income college students in Tennessee are more likely to be students of color. The policy in Tennessee therefore allows us to explore whether the equity provisions of PBF 2.0 have spillover effects for other historically disadvantaged groups that are not explicitly prioritized.

Research Design

Data

We use data primarily from the Integrated Postsecondary Education Data System (IPEDS). IPEDS contains annual institution-level administrative data for nearly all public and private non-profit two- and four-year colleges and universities in the United States. The data include the overall number of certificates and degrees conferred at each institution and separately by race/ethnicity during each school year.

We constructed three outcomes separately for URM and non-URM students with this data: the logged number of less-than-two-year certificates conferred at public two-year colleges, the logged number of associates degrees conferred at public two-year colleges, and the logged number of bachelor's degrees conferred at four-year public colleges and universities.[9] We constructed a fourth outcome, the proportion of certificates/degrees awarded to URM students, to examine if PBF policies in Tennessee and Ohio altered race-based completion gaps.

There are several limitations to IPEDS race/ethnicity data. The coarse measures of race and ethnicity in IPEDS may gloss over important experiential differences within student subgroups, and students' racial and ethnic identities may change over time (Viano & Baker, 2020). Despite these limitations, IPEDS is the only dataset that links student race/ethnicity to institutional degree conferrals across the census of public postsecondary institutions. IPEDs can therefore provide an understanding of whether and how PBF 2.0 policies differentially impact student subpopulations.

[9] Less-than-two-year certificates include both short- (less than one-year) and moderate-term certificates (one-to-two-years,). We include both certificate types in our estimation because some states define "short-term" and "long-term" certificates differently. Short-term certificates account for 63% of all less than two-year certificates in our analytic sample.

We used time-varying institutional characteristics as covariates to construct the synthetic counterfactual states. The covariates selected account for several factors, including enrollment size, student body composition, affordability, and generosity of public funding shown to influence credential production in prior research (Russell, 2019; Dale & Krueger, 2014; Goldrick-Rab, 2016; Deming & Walters, 2017). From IPEDS, we constructed enrollment-weighted, inflation-adjusted measures of the average net price, instructional expenditures per full-time-equivalent student, and the share of education and general expenditures paid for by state appropriations at the state-by-year level.[10] We also aggregated the number of undergraduate URM, non-URM, and students of "other" races enrolled at public institutions up to the state-by-year level.

Lastly, because our analytic window spans the Great Recession and its severity varied across the country, we used county-level unemployment data from the U.S. Bureau of Labor Statistics to construct the annual unemployment rate within the commuting zone of each public institution as a proxy for local economic hardship. Like the other covariates, we then aggregated the unemployment measure up to the state-by-year level.

Samples

We constructed a state-by-year dataset spanning the 2004-2015 school years. Prior to aggregating to the state level, we restricted the IPEDS data to public, degree-granting two- and four-year colleges and universities (N=1,695). We then excluded institutions that, over the study period, changed their degree-granting status or were missing covariates.[11,12] This resulted in a

[10] We inflation-adjusted using the Consumer Price Index (CPI).

[11] 143 institutions were missing covariates in one or more years and 392 institutions reported an inconsistent degree-granting status over the study period. There is some overlap between groups. We excluded the former to avoid constructing counterfactuals using imputed data. We excluded the latter to avoid conflating the effects of PBF 2.0 with degree-granting status changes, which may also have influenced credential conferrals in the post-treatment years.

[12] To assess the robustness of our estimates to both exclusion criteria, we estimated effects on a state-by-year dataset constructed from a broader set of institutions by interpolating missing data where possible and including institutions

balanced panel of 1,250 institutions across all states. Our analytic sample accounts for 80 percent of certificates and bachelor's degrees and 82 percent of associate degrees awarded to undergraduates attending public institutions between 2004 and 2015. We aggregated the data to the state-by-year level by summing enrollment and degree counts and by calculating enrollment-weighted averages of institutional revenues, expenditures, and local unemployment rates across all public institutions in each state and year.

After constructing the state-level dataset, we restricted the set of states to Ohio, Tennessee and the 16 states identified as having no PBF policies from 2004-2015.[13,14] This set of non-PBF states provides the cleanest counterfactual for estimating the effects of the new PBF 2.0 policies in Tennessee and Ohio. In Table 1, we present descriptive characteristics of public institutions operating in Tennessee, Ohio, and the donor pool states. There are large average differences on numerous characteristics between institutions in the donor and treated states, making it not possible to estimate unbiased policy impacts by simply comparing mean outcomes. Below, we discuss our strategy for estimating impacts in light of these differences.

[Table 1]

Analytic Strategy

Most previous empirical studies in the PBF literature have estimated impacts using a difference-in-differences (DID) design (e.g., Birdsall, 2018; Hillman & Corral, 2018; Kelchen &

---

that changed their degree-granting status. The broader sample includes nearly all public institutions that did not close or convert to a private institution between 2004-2015 (N=1,543) and accounts for 97.5 percent of all credentials conferred by public institutions over the period we examine. We estimate nearly identical effects, both statistically and substantively, across the two samples. These results are available from the authors upon request.

[13] See Table 2 for the list of states by sector included in the donor pool and Appendix A for information on how states were identified.

[14] We also conducted supplemental analyses in which we expanded the set of comparison group states to include those that implemented "rudimentary" PBF policies over the study period. This yielded a total of 33 states in the donor pool. Results from this larger sample are more difficult to interpret because some states in the comparison group used PBF to allocate institutional funding. Nevertheless, the results are generally consistent with those estimated off the subset of non-PBF states, although the effects are estimated less precisely in the larger sample.

Stedrak, 2016; Umbricht, et al., 2017). The DID design assumes that in the absence of the policy change, pre-treatment outcome differences between treated and comparison units would have remained unchanged in the post-treatment period. As illustrated in Appendix Figure A1, this "parallel trends" assumption is violated across many of our outcomes of interest in both Ohio and Tennessee.[15] Many of the point estimates in the pre-treatment period are large in magnitude and statistically significant, indicating that the outcome differences between public institutions in Tennessee/Ohio and comparison states were changing over time, even before the adoption of PBF 2.0. We conclude that DID is inappropriate for understanding the effects of PBF across student subpopulations in our context.

We use the Synthetic Control Method (SCM) to address the fact that the parallel trends assumption is violated in a DID framework. SCM takes a data-driven approach to finding a comparison group in small-sample comparative studies in which the parallel trends (and levels) assumption is feasible (Abadie & Gardeazabal, 2003; Abadie et al., 2010; 2014). We therefore continue to assume that outcomes in Tennessee/Ohio and their respective comparison group would have evolved similarly over time in the absence of the PBF 2.0 policies. SCM ensures this assumption is reasonable by constructing a "synthetic" state that resembles Tennessee/Ohio in both the mean values and trends of outcomes and covariates over the pre-treatment period. This is accomplished by identifying a vector of weights that is applied to the set of comparison states to

[15] We examined the appropriateness of a DID-based estimation strategy across multiple comparison group definitions (e.g. regional and national comparisons). We also examined the feasibility of conducting a state-level, rather than institution-level, DID analysis. The parallel trends assumption was violated in all cases.

minimize the difference between Tennessee/Ohio and the comparison group in the pre-treatment period.[16]

The improvement in internal validity from using SCM comes at a cost. By aggregating the data to the state level, we lose the ability to examine treatment effect heterogeneity across institution types.[17] We are therefore limited in our ability to explore the mechanisms of the effects we estimate. Nevertheless, we consider this trade-off reasonable in order to estimate plausibly unbiased causal impacts.

We tested 18 approaches to constructing the counterfactuals separately for Tennessee and Ohio. We modeled the outcomes and covariates over the pre-treatment years (2004-2008 for Ohio and 2004-2009 for Tennessee) in three ways – averaging over all pre-treatment years, using the values in the last two pre-treatment years, and using the values in the last three pre-treatment years – and tested each resulting outcome-by-covariate combination.[18] In addition, we explored using the full set of institutional characteristics described in the Data section in the vector of covariates as well as a more parsimonious vector that only included the URM, non-URM, and “other” race enrollment measures in the list of covariates.

Because we estimate impacts over multiple outcomes (i.e., less-than-two-year certificates, associate degrees, and bachelor’s degrees) and groups (i.e., URM, non-URM, and the share of credentials conferred to URM students) in each state, different approaches minimized the

[16] The “optimal” vector of weights minimizes the root mean-squared prediction error (RMPSE) in the pre-treatment period.

[17] Recent extensions to SCM allow for inference with multiple treated units. However, like the DID models, the institution-level SCM specifications created poorly fitted synthetic control groups in the pretreatment period. We therefore present results from standard SCM models in which the identifying assumptions are plausible, but for which data aggregation to the state level is required.

[18] We did not construct synthetic control groups using outcome and covariate values in every pre-treatment year to avoid overfitting to the data. Likewise, we did not construct synthetic states using only pre-treatment outcome values because this approach did not differentiate well among the donor pool to select the subset of states that best represented Tennessee/Ohio in the pre-treatment years.

difference between the treatment and comparison groups across each outcome-by-group combination. In our main results, we present effect estimates using the "optimal" criterion that minimized the root mean-squared prediction error (RMSPE) for each outcome-by-group combination.[19] As a robustness check, we examine the sensitivity of using sub-optimal criterion to construct the counterfactuals and estimate effects.

In Tables 2 and 3, we report the optimal weights assigned to donor states to construct the counterfactuals for each outcome-by-group combination in Tennessee and Ohio, respectively. The weights vary considerably across donor pool states for each outcome-by-group combination. This is reasonable given the differences in URM enrollment and degree completion across Tennessee, Ohio, and the donor states. Importantly, the weights converge on a subset of the donor pool states for each outcome-by-group combination. This indicates good differentiation between the possible donor pool states; those states with positive weights best represent Tennessee and Ohio over the pre-treatment period (McClelland & Gault, 2017). For example, to estimate impacts on the log number of certificates conferred to URM students in Tennessee, Alabama, Connecticut, New Jersey, and West Virginia receive all the weight to construct the counterfactual, whereas for non-URM students, Alabama, California, New Hampshire, New Jersey, Rhode Island, and West Virginia are weighted to construct the counterfactual. We observe similar variability with respect to the donor state weight assignments in Ohio, although broadly speaking, states in the South receive relatively less weight to construct the counterfactuals for Ohio as compared to Tennessee.

[Tables 2 & 3]

[19] Although the counterfactual is constructed from a unique weighting of the donor pool states for each outcome-by-group estimate, the weights consistently derive our best approximation of what the outcome-by-group trend in the treated state would have been in the absence of PBF 2.0. All our results therefore share a common interpretation and can be compared and contrasted.

After constructing the synthetic control groups, we estimate the effect of performance funding in Tennessee/Ohio on completion outcomes by calculating the difference between outcomes in the treated state and the synthetic state in each post-adoption year. To conduct hypothesis tests, we use a permutation-based test to report an empirical *p*-value (Abadie et al., 2010). We implement the test by restricting the sample to comparison states, choosing one state as the placebo treated state, and then re-running the SCM model to estimate placebo treatment effects. We repeat this process over all states in the donor pool and compare the proportion of placebo effect estimates that are at least as large as the estimated effects in Tennessee/Ohio. Because placebo matches vary in quality across donor states, we report adjusted *p*-values that account for this.[20]

Results

Graphical Evidence

The graphical evidence suggests that the synthetic counterfactuals for Tennessee and Ohio provide a reasonable approximation of the outcome paths that would have materialized in the absence of PBF 2.0. In Figure 1, we plot the number of associate degrees conferred to URM and non-URM students in Tennessee and Ohio and in their respective counterfactual states before and after PBF-adoption. In Tennessee, the synthetic control for non-URM students closely tracks the actual outcome path of non-URM students in the first six pre-treatment years. The synthetic control for URM students aligns less well in the first three pre-treatment years, but almost perfectly tracks the trend in associate degrees conferred to URM students in the three years immediately before PBF 2.0 took effect. The synthetic control groups for Ohio mirror the completion levels and trends

[20] Instead of reporting the proportion of placebo effect estimates at least as large as the effect estimate in the actual treated states, the adjusted p-values report the proportion of ($\frac{placebo\ effect}{pre-treatment\ RMSPE}$) ratios at least as large as the ratio in the actual treated states.

for URM and non-URM students almost perfectly over the entire pre-treatment period. Taken together, we conclude that SCM generates reasonable counterfactuals for estimation of policy effects.

Comparing the completion trends in Tennessee and Ohio after the introduction of PBF 2.0 to trends in the synthetic states suggests that these policies exacerbated race-based disparities in associate degree attainment. In Tennessee, the actual and counterfactual trends in Figure 1 for URM students are nearly identical over the post-treatment period, indicating that PBF 2.0 had no impact on the number of associate degrees conferred to URM students. In contrast, the number of associate degrees conferred to non-URM students in Tennessee lies above the synthetic counterfactual in all six post-treatment years, suggesting that PBF 2.0 increased the number of associate degrees conferred to non-URM students. In Ohio, we observe a different pattern of results, which suggest that the introduction of PBF 2.0 decreased the number of associate degrees conferred to URM students and had no impact on the number of associate degrees conferred to non-URM students. Taken together, the patterns provide suggestive evidence of differential policy impacts by race, implying that the adoption of PBF 2.0 policies in Tennessee and Ohio widened racial disparities in associate degree attainment.

SCM Estimates of Effects on Certificate, Associate, and Bachelor's Degree Conferrals

The SCM results corroborate the graphical evidence that PBF 2.0 in Tennessee and Ohio had differential impacts on attainment by race/ethnicity. In Table 4, columns 1-3, we report impacts in Tennessee on less-than-two-year certificate production in each post-adoption year. The estimates for URM students are positive and significant in the second and third post-treatment years; certificates conferred to URM students increased by factors of 2 and 1.4, respectively, in those years.[21] The point estimates are closer to zero and statistically insignificant thereafter. In contrast, PBF 2.0 in Tennessee led to larger and sustained growth in certificate conferrals to non-URM students. PBF 2.0 increased the number of certificates conferred to non-URM students by factors of 1.5-4 across all post-treatment years. As a result, the proportion of certificates conferred to URM students decreased after PBF 2.0 took effect. The share of certificates awarded to URM students declined by 4.6 percentage points in the first year of adoption; six years after adoption, the share was 9.1 percentage points lower than the synthetic control state.

PBF 2.0 did not impact the number of certificates conferred to either URM or non-URM students in Ohio. The estimates for both groups are statistically insignificant (Table 4, columns 4 and 5). The direction of the estimated effects for each group is also inconsistent. We do estimate that PBF 2.0 decreased the share of certificates awarded to URM students by 3.0 percentage points and 4.3 percentage points in the fifth and seventh year of adoption, respectively. However, given the inconsistent and insignificant pattern of results in the logged number of degrees, the results provide limited evidence that PBF 2.0 exacerbated race-based gaps in certificate production in Ohio.

[Table 4]

Like the estimated impacts on certificate conferrals, the pattern of results with respect to associate degree conferrals are different in the two states. PBF 2.0 did not impact associate degree production for URM students in Tennessee (Table 5, column 1). The point estimates are small, inconsistent in sign, and statistically insignificant. In contrast, the number of associate degrees awarded to non-URM students in Tennessee increased by 10-20 percent in each of the six years after the introduction of PBF 2.0. The effect estimates on the proportion of associate degrees

[21] Percent changes are calculated by exponentiating the effect estimates (i.e., $e^{\hat{\beta}}$).

awarded to URM students in Tennessee are also consistently negative and increase in magnitude over time, but none are statistically significant. Despite clear evidence that PBF 2.0 had differential impacts on associate degree production by race and suggestive evidence that this widened racial disparities in associate degree production, we are unable to reject that the race-based gap in associate degree conferrals remained unchanged after Tennessee introduced PBF 2.0.

In Ohio, the estimated impacts on associate degree conferrals to URM students (Table 5, column 4) are consistently negative, increase in magnitude over time, and are statistically significant in all but the first year after PBF 2.0 took effect. In 2015, seven years after the initial implementation and three years after most funds were allocated based on course/degree completion and equity provisions were in place, PBF 2.0 decreased the number of associate degrees conferred to URM students by 29 percent. By comparison, we estimate small, inconsistent, and imprecise coefficients on the number of associate degrees conferred to non-URM students in Ohio. We conclude that the new funding model in Ohio had no impact on associate degree production for non-URM students. Like in Tennessee, the differential impacts on associate degree production by race are suggestive of widening racial gaps in Ohio. We estimate consistently negative effects in Ohio (Table 5, column 6) similar in magnitude to the estimated impacts in Tennessee; however, those estimates are also statistically insignificant.

[Table 5]

The results also suggest that PBF 2.0 decreased the number of bachelor's degrees conferred to URM students in Tennessee by 2-5 percent per year (Table 6, column 1). The coefficients are consistently negative across all years and significant in years one and five post-treatment. The analogous estimates for non-URM students in column 2 are consistently positive and the coefficient is significant in the third post-treatment year. Again, the differential pattern of results is suggestive of widening racial disparities in bachelor's degree production in Tennessee after the adoption of PBF 2.0, but the estimated effects on the share of bachelor's degrees conferred to URM students are statistically insignificant.

PBF 2.0 had differential impacts on bachelor's degree production by race in Ohio. PBF 2.0 had no impact on bachelor's degree conferrals to URM students. The estimates are inconsistent in sign and small in magnitude over the post-treatment years (Table 6, column 4). Among non-URM students, however, the estimated effects are consistently positive, increasing in magnitude, and statistically significant beginning in the second post-treatment year. PBF 2.0 increased the number of bachelor's degrees conferred to non-URM students by 15 percent seven years after adoption. The effect estimates on the share of bachelor's degrees conferred to URM students in column 6 are negative in years four through seven after adoption, but only the coefficient in year seven is significant. We conclude that PBF 2.0 produced differential impacts on bachelor's degree production by race in Ohio, but those differences were too small in the initial years to shift the racial distribution of bachelor's degrees conferred. By 2015, PBF 2.0 decreased the share of bachelor's degrees awarded to URM students by three percentage points.

In summary, our results reveal that PBF 2.0 produced differential impacts on credential production by race in Tennessee and Ohio, albeit with distinct patterns in each state. Across both states, we generally estimate null or negative effects on credentials conferred to URM students and null or positive effects on credentials conferred to non-URM students.[22] In Tennessee, the strongest evidence is that PBF 2.0 produced sustained gains in certificate and associate degree production among non-URM students, but fleeting, if any gains at two-year colleges for URM students. In

[22] This pattern of effects appears to be common across the plurality of public institutions in Tennessee and Ohio; it is not explained by PBF 2.0 imposing disproportionate financial strain on the sole public HBCUs operating in each state. Estimated impacts are very similar when we exclude HBCUs and other predominantly minoring-serving institutions from both the treated and donor pool states. Results available from the authors upon request.

Ohio, the strongest evidence is that PBF 2.0 decreased associate degree production among URM students and increased bachelor's degree production among non-URM students.

[Table 6]

Robustness of the Effect Estimates

We examine the robustness of the effect estimates for URM and non-URM students in Tennessee and Ohio in Figures 2 and 3, respectively.[23] In each figure, we plot the outcome trend in the treated state (solid black line), the optimal synthetic state that minimizes the RMSPE in the pre-treatment period (dashed black line), and eight alternative counterfactuals that fit the data more poorly in the pre-treatment period (grey lines).[24] Comparing how closely the dashed black and grey lines align, and the gap between those lines and the solid black line in the post-treatment period, reveals whether the effect estimates are robust to alternative constructions of the counterfactual.

Many of the alternative counterfactuals track the observed pattern of less-than-two-year certificates conferred to URM students in Tennessee in the pre-treatment period (Figure 2, panel A). The alternative effect estimates generated by those counterfactuals are similar or greater in magnitude to the main effects we report. Likewise, most alternative counterfactuals demonstrate a reasonably good fit to the actual number of certificates conferred to non-URM students in the pre-treatment period; the effect estimates generated by the optimal synthetic state are similar to those estimated using alternative synthetic states. We conclude that our main estimates of increased race-based disparities in certificate production in Tennessee are robust to alternative constructions of the counterfactuals.

[23] Analogous results for the share of credentials conferred to URM students are presented in Appendix Figures A2 and A3.

[24] We plot alternative counterfactuals that best fit the data for at least one other outcome-by-group combination in Tennessee or Ohio.

Our associate and bachelor's degree effect estimates in Tennessee also appear robust to alternative counterfactuals that reasonably fit the outcome paths in Tennessee in the pre-treatment period (Figure 2, Panels B and C). For example, five of the nine synthetic controls reasonably approximate the observed, logged number of associate degrees conferred to non-URM students from the beginning of the pre-period. Eight of the nine counterfactuals continue to closely track each other in the post-policy period. The robustness checks are consistent with our finding of significant, positive impacts on the number of non-URM associate degrees conferred and null effects on the number of URM associate degrees conferred and the number of bachelor's degrees conferred to both URM and non-URM students.

The effect estimates in Ohio are also generally robust to alternative counterfactual constructions. For example, eight of the nine counterfactuals for estimating effects on associate degree production for non-URM students closely track each other in the pre- and post-treatment period, and the actual outcome path in Ohio deviates from all but one alternative synthetic state in the post-policy period. Likewise, the results in panel C of Figure 3 indicate that the positive impact on bachelor's degrees awarded to non-URM students is also evident using well-matched alternative counterfactuals. In Panel C of Figure A3, the optimal match provides the most conservative estimate of the decline in the share of bachelor's degrees conferred to URM students in Ohio compared to other well-matched alternative counterfactuals.

Discussion

In this study, we find that PBF 2.0 produced differential impacts by race/ethnicity, even when equity provisions are included in the funding model. We find that PBF policies exacerbated racial disparities in certificate completion in Tennessee and bachelor's degree conferrals in Ohio. We also find suggestive evidence that the policies exacerbated racial disparities in associate degree conferrals in both states. These heterogeneous impacts are consistent with previous research estimating small or null overall effects of PBF 2.0 on degree completion in Tennessee and Ohio (Hillman, et al., 2018; Ward & Ost, 2021). Our results within a particular degree outcome often consist of significant impacts for one group, and null effects for another. We interpret this as evidence that prior small or null effects in the aggregate may mask heterogeneity in the impacts of these policies on different student populations.

Because we analyze the impact of the PBF models in their entirety, we are unable to isolate the effects of specific policy components or the drivers of the differential effects across the two states, although our findings indicate that the heterogeneity of effects by race are not isolated to public HBCUs, but rather, experienced systemwide. From a policy perspective, identifying the effects of particular policy components and the mechanisms behind them is important as PBF 2.0 continues to evolve (Ortagus et al., 2020). We conclude by situating our results within the extant PBF literature and our understanding of Tennessee's and Ohio's PBF policies during the period of our analysis.

Our findings indicate that PBF 2.0 significantly, and disproportionately, impacted certificate production for URM and non-URM students in Tennessee, but not Ohio. These results echo prior estimates of PBF impacts on total certificate production in Tennessee and Ohio, which found that overall certificate production increased dramatically in Tennessee after the adoption of PBF 2.0, while the effect in Ohio was delayed and attenuated in comparison (Hillman et al., 2018). National analyses have also demonstrated that two-year institutions operating under PBF 2.0 policies were more likely to increase less-than-one-year certificates than other credentials (Li & Kennedy, 2018). It is possible that race-based gaps in certificate outcomes widened less in Ohio because, unlike Tennessee, Ohio did not tie state funding to less-than-one-year certificates and only explicitly incentivized longer-term certificate production at the end of our analytical timeframe. This may have reduced the pressure felt by two-year institutions in Ohio to improve certificate program performance.

It is also notable that in both Tennessee and Ohio, the pattern of differential impacts between URM and non-URM students holds across associate and baccalaureate degree outcomes, despite differences in the equity provisions across the two states. In Tennessee, students from low-income backgrounds and adult students were prioritized for additional funding during our analytic timeframe. In Ohio, the prioritization of URM student groups was phased in over the analytic time period across both the two- and four-year sectors.

One possibility, based on our findings, is that the equity provisions in the early iterations of PBF 2.0 were insufficiently linked to degree outcomes. If this were the case, we might expect the estimated effects to stabilize or attenuate in Ohio once the equity provisions were introduced in the two-year sector or once the hold harmless provisions expired in the four-year sector. However, we find no evidence of this. Instead, PBF consistently and negatively impacted the number of associate degrees conferred to URM students and had no impact on the number of baccalaureate degrees conferred to URM students throughout the post-treatment period. Our estimates are also robust to defining the post-treatment period only after the equity provisions were fully phased in.

We find little evidence that PBF 2.0 affected two-year institutions differently than four-year institutions. The estimated impacts on two-year outcomes in Tennessee are more similar to the pattern of four-year effects in Ohio, with null impacts on URM degree conferrals and positive, significant impacts on non-URM degree conferrals. The similarity of results across sectors suggests that the equity provisions in place during our analytic period were insufficient for mitigating disparate impacts across student populations. Qualitative work in Ohio and Pennsylvania suggests that the size of equity provisions may even be immaterial to PBF policies operating differently across student groups and institutional contexts (Li, 2019).

One potential explanation for the heterogeneous effects by race is variation in institution-level capacity to respond to PBF policies (Dougherty et al., 2016). Given long-standing inequalities in resources across institutions that cut across racial lines, broad-access public institutions that serve disproportionate shares of URM students may focus out of necessity on increasing attainment by any means. Prioritizing completion for historically underserved groups over traditionally advantaged groups may simply not be feasible at resource-constrained colleges when the stakes are high and the foundational conditions for raising performance are limited. More highly-resourced institutions may be better-equipped to respond to equity incentives, but because those institutions predominantly serve students from historically advantaged groups, racial gaps in credential attainment may nevertheless widen on balance. Which institutions can couple PBF policies with additional student supports, combined with the racial stratification of students into high- and low-resourced institutions may explain the differences in effects by race we observe in this study.

Given the mixed findings to date and the challenge of estimating impacts of complex, multi-layered policies, further research is needed to understand if and how institutional context moderates the effects of PBF 2.0 policies and if differential effects by race operate primarily through within- or between-college factors. Future work might also examine how PBF 2.0 policies affect institutional funding priorities and the adoption of complementary initiatives, such as targeted financial aid programs and/or diversity initiatives. Further exploration of how PBF 2.0

differentially affects students by race/ethnicity, income status, and age may also be worthwhile, as there may be additional heterogeneity within the broad subpopulations we examined.

We conclude with a final limitation. Because both states have revised their funding formulas since the period we study, future research should examine how the evolving nature of PBF 2.0 in Tennessee and Ohio has impacted educational equity in these states. Amidst the widespread adoption of PBF in higher education, developing a clearer understanding of the opportunities and limitations associated with these funding models is critical to financing public higher education systems in service of dismantling, rather than reinforcing, historical inequities.

References

Abadie, A. & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, *93*(1), 113–132. 10.1257/000282803321455188

Abadie, A., Diamond, A., & Hainmueller, J. (2010). 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.

Abadie, A., Diamond, A., & Hainmueller, J. (2014). Comparative politics and the synthetic control method. *American Journal of Political Science 59*(2), 495–510.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent." *Science, 344*(6186), 843–851.

Avery, C., & Turner, S. (2012). Student loans: Do college sudents borrow too much or not enough? *Journal of Economic Perspectives, 26*(1), 165-192.

Bailey, M. J., & Dynarski, S. M. (2011). Inequality in postsecondary education. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 117–131). New York, NY: Russell Sage Foundation.

Birdsall, C. (2018). Performance Management in Public Higher Education: Unintended Consequences and the Implications of Organizational Diversity. *Public Performance & Management Review*, *41*(4), 669-695. 10.1080/15309576.2018.1481116

Boelscher, S. & Snyder, M. (2019). Driving Better Outcomes: Fiscal Year 2019 State Status & Typology Update. Washington, D.C.: HCM Strategists.

Carnevale, A. P., Jayasundera, T. & Gulish, A. (2016). *America's Divided Recovery: College Haves and Have-Nots*. Washington, D.C.: Georgetown University Center on Education and the Workforce.

Chetty, R., Friedman, J., Saez, E., Turner, N., & Yagan, D. (2017). Mobility Report Cards: The Role of Colleges in Intergenerational Mobility (NBER Working Paper No. 23618). Cambridge, MA: National Bureau of Economic Research.

Dale, S. B., & Krueger, A. B. (2014). Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data. *Journal of Human Resources*, *49*(2), 323-358.

Davies, S., & Zarifa, D. (2012). The stratification of universities: Structural inequality in Canada and the United States. *Research in Social Stratification and Mobility*, *30*(2), 143-158.

Deming, D. J., & Walters, C. R. (2017). *The impact of price caps and spending cuts on US postsecondary attainment* (NBER Working Paper 23736). Cambridge: National Bureau of Economic Research.

Dowd, A. C., & Bensimon, E. M. (2015). *Engaging the "race question": Accountability and equity in U.S. higher education*. New York: Teachers College Press.

Dougherty, K. J., Jones, S. M., Lahr, H., Natow, R. S., Pheatt, L., & Reddy, V. (2016). *Performance Funding for Higher Education*. Baltimore: Johns Hopkins University Press.

Dougherty, K. J., & Natow, R. S. (2015). *The Politics of Performance Funding for Higher Education: Origins, Discontinuations, and Transformations*. Baltimore: Johns Hopkins University Press.

Dougherty, K. J., & Natow, R. S. (2020). Performance-based funding for higher education: How well does neoliberal theory capture neoliberal practice? *Higher Education*, *80*(3), 457-478. https://doi.org/10.1007/s10734-019-00491-4

Favero, N., & Rutherford, A. (2020). Will the tide lift all boats? Examining the equity effects of performance funding policies in U.S. higher education. *Research in Higher Education*, *61*, 1-25.

Gándara, D., & Rutherford, A. (2018). Mitigating unintended impacts? The effects of premiums for underserved populations in performance-funding policies for higher education. *Research in Higher Education*, *59*(6), 681–703.

Gándara, D., & Rutherford, A. (2020). Completion at the Expense of Access? The Relationship Between Performance-Funding Policies and Access to Public 4-Year Universities. *Educational Researcher*, *49*(5), 321–334. 10.3102/0013189X20927386

Goldrick-Rab, S. (2016). *Paying the Price: College Costs, Financial Aid, and the Betrayal of the American Dream*. University of Chicago Press.

Hagood, L. P. (2019). The financial benefits and burdens of performance funding in higher education. *Educational Evaluation and Policy Analysis*, *41*(2), 189-213.

Hillman, N. W., & Corral, D. (2018). The equity implications of paying for performance in higher education. *American Behavioral Scientist*, *61*(14), 1757-1772.

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.

Jones, T., Jones, S., Elliott, K. C., Russell Owens, L., Assalone, A. E., & Gándara, D. (2017). *Outcomes Based Funding and Race in Higher Education: Can Equity be Bought?* New York, NY: Springer International Publishing.

Kelchen, R., & Stedrak, L. (2016). Does Performance-Based Funding Affect Colleges' Financial Priorities? *Journal of Education Finance*, *41*(3), 302-321.

Kelchen, R. (2018) Do performance-based funding policies affect underrepresented student enrollment? *The Journal of Higher Education*, *89*(5), 702-727.

Kelchen, R. (2019). Exploring the relationship between performance-based funding design and underrepresented student enrollment at community colleges. *Community College Review*, *47*(4), 382–405. https://doi.org/10.1177/0091552119865611

Kelchen, R., Ortagus, J., Rosinger, K., Brown, L., & Voorhees, N. (2019). *Profiles of state performance- based funding policies: Missouri, New York, and Tennessee* (p. 9) [Policy Brief]. InformEd States. https://static1.squarespace.com/static/5d9f9fae6a122515ee074363/t/5e34488f82e01937d75e20a7/1580484753688/IS_Brief_PerformanceBasedFundingPolicies.pdf

Li, A. Y. (2019). The Weight of the Metric: Performance Funding and the Retention of Historically Underserved Students. *The Journal of Higher Education*, *90*(6), 965–991. https://doi.org/10.1080/00221546.2019.1602391

Li, A. Y., & Ortagus, J. C. (2019). Raising the Stakes: Impacts of the Complete College Tennessee Act on Underserved Student Enrollment and Sub-Baccalaureate Credentials. *The Review of Higher Education 43*(1), 295-333. doi:10.1353/rhe.2019.0097.

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. https://doi.org/10.1177/0091552117743790

McClelland, R., & Gault, S. (2017). *The synthetic control method as a tool to understand state policy*. Washington, DC: Urban Institute.

Morris, M. (2009). *LSC Greenbook: Analysis of the Enacted Budget, Ohio Board of Regents.* Columbus, OH: Legislative Service Commission.

Ness, E. C., Deupree, M. M., & Gándara, D. (2015). *Campus responses to outcomes-based funding in Tennessee: Robust, aligned, and contested.* New York, NY: Ford Foundation.

Ohio Board of Regents (2013). State Share of Instruction Handbook: Providing the Methodology for Allocating State Share of Instruction Funds for Fiscal Year 2014. Columbus, OH: Ohio Board of Regents.

Ortagus, J. C., Kelchen, R., Rosinger, K., & Voorhees, N. (2020). Performance-Based Funding in American Higher Education: A Systematic Synthesis of the Intended and Unintended Consequences. *Educational Evaluation and Policy Analysis*, *42*(4), 520–550. https://doi.org/10.3102/0162373720953128

Peri, G. & Yasenov, V. (2019). The labor market effects of a refugee wave: Synthetic control method meets the Mariel Boatlift. *Journal of Human Resources*, *54*(2), 267-309.

Rodriguez, A, Deane, K.C., & Daviss III, C.H. (2021). Towards a framework of racialized policymaking in higher education. In *Higher Education: Handbook of Theory and Resarch: Volume 37*, (pp. 1-81). Springer International Publishing.

Russell, L. (2019). Better outcomes without increased costs? Effects of Georgia's University System consolidations. *Economics of Education Review*, *68*, 122-135. https://doi.org/10.1016/j.econedurev.2018.12.006

State Higher Education Executive Officers Association (2019). "PBF Funding Policies by State". Unpublished raw data.

Snyder, M. (2015). "Driving better outcomes: Typology and principles to inform outcomes-based funding models". Washington, DC: HCM Strategists.

Snyder, M., & Fox, B. (2016). "Driving better outcomes: Fiscal year 2016 state status & typology update". Washington, DC: HCM Strategists.

Turocy, M. (2013). LSC Greenbook: Analysis of the Enacted Budget, Ohio Board of Regents. Columbus, OH: Legislative Service Commission.

Turocy, M., & Perch, M. (2013). LSC Redbook: Analysis of the Executive Budget Proposal, Ohio Board of Regents. Columbus, OH: Legislative Service Commission.

Umbricht, M. R., Fernandez F., & Ortagus J. C. (2017). An Examination of the (Un)Intended Consequences of Performance Funding in Higher Education. Educational Policy. 2017;31(5):643-673. doi:10.1177/0895904815614398.

Ward, J., & Ost, B. (2021). The effect of large-scale performance-based funding in higher education. *Education Finance and Policy*, *16*(1), 92–124.

Viano, S., & Baker, D. J. (2020). How Administrative Data Collection and Analysis Can Better Reflect Racial and Ethnic Identities. *Review of Research in Education*, *44*(1), 301–331. https://doi.org/10.3102/0091732X20903321

Zerquera, D., & Ziskin, M. (2020). Implications of performance-based funding on equity-based missions in U.S. higher education. *Higher Education*, *80*(6), 1153-1174.

Ziskin, M. B., Rabourn, K. E., & Hossler, D. (2018). Performance-based funding of higher education: Analyses of policy discourse across four case studies. *Journal for Critical Education Policy Studies*, *16*(2).

Zumeta, W., & Li, A. Y. (2016). Assessing the underpinnings of performance funding 2.0: Will this dog hunt. https://www.tiaainstitute.org/sites/ default/files/presentations/2017-02/ti_assessing_the_underpinnings_of_perfor mance_funding_2.pdf.

Tables & Figures

Table 1

Select Characteristics of Institutions in Tennessee, Ohio, and Donor Pool States Prior to PBF Adoption (Averaged Over 2004-2008)

	(1)	(2)	(3)	(4)	(5)	(6)
	Two-Year Institutions			Four-Year Institutions		
	Donor Pool States	Tennessee	Ohio	Donor Pool States	Tennessee	Ohio
Total Undergraduate Enrollment	96,895	64,405	119,963	62,636	101,765	200,965
	(196,579)	(2,052)	(2,532)	(80,778)	(3,017)	(3,631)
URM Undergraduate Enrollment	32,035	13,485	20,567	11,810	22,401	25,445
	(76,813)	(502)	(615)	(20,391)	(673)	(1,369)
Non-URM Undergraduate Enrollment	56,575	49,109	93,712	45,440	76,358	165,372
	(99,903)	(1,046)	(1,204)	(50,682)	(1,885)	(1,826)
Total Less-than-Two-Year Certificates Awarded	5,124	1,581	6,287			
	(9,261)	(51)	(1,114)			
Certificates Awarded to URM Students	1,400	301	803			
	(3,326)	(34)	(121)			
Certificates Awards to Non-URM Students	3,345	1,231	5,183			
	(5,316)	(91)	(869)			
Total Degrees Awarded	9,872	7,045	15,986	12,564	17,097	37,125
	(16,884)	(65)	(708)	(18,729)	(774)	(668)
Degrees Awarded to URM Students	2,516	1,043	1,869	2,014	3,040	3,279
	(5,793)	(68)	(93)	(3,989)	(134)	(185)
Degrees Awards to non-URM Students	6,535	5,871	13,405	9,347	13,668	31,717
	(9,356)	(43)	(532)	(12,090)	(540)	(412)
Average Net Price	$8,391	$7,633	$7,759	$14,560	$11,280	$18,479
	(3,454)	(559)	(144)	(2,377)	(186)	(581)

Educational Expenditures per FTE Student	$5,102	$4,925	$5,039	$10,817	$10,025	$10,407
	(975)	(202)	(91)	(2,848)	(214)	(395)
Share of Educational Expenditures Covered by	0.38	0.47	0.39	0.36	0.39	0.30
State Funds	(0.13)	(0.03)	(0.01)	(0.1)	(0.02)	(0.01)
Number of Institutions	19.8	13	22	5.9	9	20
	(22.6)			(4.5)		
Number of States	13	1	1	15	1	1

Notes: Means are reported with standard deviations in parentheses. The donor pool is comprised of states that never implemented performance-based funding from 2004-2015. See Table 2 for the list of states by sector included in the donor pool. Black, Latino, and American Indian/Alaskan native students are categorized as underrepresented racially minoritized (URM) students. White and Asian students are categorized as non-URM. students.

Source: 2004-2008 Integrated Postsecondary Education Data System

Table 2

Weights Assigned to Donor States for Each Outcome in the Tennessee Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Log Number of Credentials Conferred to URM Students			Log Number of Credentials Conferred to non-URM Students			Share of Credentials Conferred to URM Students		
State	Certificate	Associate Degree	Bachelor's Degree	Certificate	Associate Degree	Bachelor's Degree	Certificate	Associate Degree	Bachelor's Degree
AK	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AL	0.238	0.524	0.643	0.329	0.289	0.454	0.000	0.272	0.406
CA	0.000	0.108	0.195	0.203	0.227	0.329	0.325	0.000	0.347
CT	0.231	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000
DE	0.000	0.000	0.000	0.000	0.000	0.118	0.000	0.000	0.000
IA	0.360	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ID	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
KY	0.000	0.000	0.068	0.000	0.000	0.000	0.363	0.573	0.000
MD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NE	0.000	0.020	0.000	0.000	0.000	0.000	0.152	0.000	0.000
NH	0.000	0.000	0.000	0.027	0.044	0.000	0.000	0.000	0.000
NJ	0.147	0.000	0.000	0.186	0.082	0.000	0.000	0.156	0.000
RI	0.000	0.348	0.000	0.091	0.000	0.000	0.034	0.000	0.022
SC	0.000	0.000	0.000	0.000	0.148	0.000	0.000	0.000	0.000
VT	0.000	0.000	0.000	0.000	0.000	0.098	0.000	0.000	0.000
WV	0.025	0.000	0.094	0.164	0.209	0.000	0.111	0.000	0.225
Optimal Pre-Period Characteristics used to Create Synthetic Control									
Outcomes	Last 2 years	Last 2 years	Average	Average	Average	Last 3 years	Last 3 years	Last 3 years	Average

Covariates	Average of enrollments	Average of enrollments	Average of all covariates	Average of all covariates	Last 3 years of covariates	Average of enrollments	Average of all covariates	Last 3 years of enrollment; Average of all other covariates	Last 3 years of covariates

Notes: The donor pool is comprised of states that never implemented performance-based funding from 2004-2015. Reported weights minimize the RMSPE of the outcome in the pre-treatment period. Enrollment and non-enrollment covariates refer to the variables described in the Data section.

Table 3

Weights Assigned to Donor States for Each Outcome in the Ohio Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Log Number of Credentials Conferred to URM Students			Log Number of Credentials Conferred to Non-URM Students			Share of Credentials Conferred to URM Students		
State	Certificate	Associate Degree	Bachelor's Degree	Certificate	Associate Degree	Bachelor's Degree	Certificate	Associate Degree	Bachelor's Degree
AK	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AL	0.000	0.063	0.000	0.000	0.150	0.000	0.000	0.000	0.000
CA	0.392	0.372	0.168	0.043	0.294	0.688	0.217	0.132	0.250
CT	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IA	0.291	0.252	0.000	0.582	0.025	0.256	0.783	0.751	0.378
ID	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
KY	0.187	0.000	0.000	0.347	0.000	0.056	0.000	0.000	0.000
MD	0.000	0.000	0.035	0.000	0.000	0.000	0.000	0.000	0.000
NE	0.000	0.313	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.323
NJ	0.000	0.000	0.365	0.000	0.531	0.000	0.000	0.000	0.000
RI	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SC	0.000	0.000	0.432	0.000	0.000	0.000	0.000	0.117	0.049
VT	0.130	0.000	0.000	0.028	0.000	0.000	0.000	0.000	0.000
WV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Optimal Pre-Period Characteristics used to Create Synthetic Control									
Outcomes	Last 3 years	Last 2 years	Last 2 years	Last 3 years	Last 3 years	Average	Average	Last 3 years	Last 2 years

Covariates	Average of enrollments	Average of enrollments	Last 2 years of covariates	Average of enrollments	Average of enrollments	Average of all covariates	Average of enrollments	Average of all covariates	Last 3 years of covariates

Notes: The donor pool is comprised of states that never implemented performance-based funding from 2004-2015. Reported weights minimize the RMSPE of the outcome in the pre-treatment period. Enrollment and non-enrollment covariates refer to the variables described in the Data section.

Table 4

Estimated Effects of Performance-Based Funding on Certificate Production in Tennessee and Ohio, by URM Status and Year

	(1)	(2)	(3)	(4)	(5)	(6)
	Tennessee			Ohio		
	Log Number of Certificates Awarded		Share of Certificates Awarded to URM Students	Log Number of Certificates Awarded		Share of Certificates Awarded to URM Students
	URM Students	Non-URM Students		URM Students	Non-URM Students	
2009				-0.116	-0.064	-0.005
				[0.357]	[0.429]	[0.429]
2010	-0.125	**0.404*** **	**-0.046*** **	-0.285	-0.062	-0.014
	[0.143]	[0.000]	[0.000]	[0.143]	[0.357]	[0.286]
2011	**0.735*** **	**1.419*** **	**-0.072*** **	-0.025	-0.060	-0.006
	[0.000]	[0.000]	[0.000]	[0.714]	[0.571]	[0.500]
2012	**0.350***	**0.940*** **	**-0.060*** **	-0.104	0.000	-0.009
	[0.071]	[0.000]	[0.000]	[0.286]	[1.000]	[0.500]
2013	0.131	**0.879*** **	**-0.067*** **	-0.056	0.064	**-0.030*** **
	[0.357]	[0.000]	[0.000]	[0.714]	[0.500]	[0.000]
2014	0.140	**0.827*** **	**-0.074*** **	0.121	0.128	-0.028
	[0.286]	[0.000]	[0.000]	[0.214]	[0.143]	[0.143]
2015	-0.015	**0.718*** **	**-0.091*** **	0.263	0.252	**-0.043***
	[0.786]	[0.000]	[0.000]	[0.214]	[0.143]	[0.071]
Mean (unlogged) in year before policy adoption	408	1,699	0.194	807	4,995	0.139

*** p < 0.01, ** p < 0.05, * p < 0.10

Notes: Effects are estimated using the synthetic control method. Reported p-values (in brackets) are derived from placebo permutation tests and account for the quality of each placebo match in the pre-treatment period.

Table 5

Estimated Effects of Performance-Based Funding on Associate Degree Production in Tennessee and Ohio, by URM Status and Year

	(1)	(2)	(3)	(4)	(5)	(6)
	TN			OH		
	Log Number of Degrees Awarded		Share of Degrees Awarded to URM Students	Log Number of Degrees Awarded		Share of Degrees Awarded to URM Students
Year	URM Students	Non-URM Students		URM Students	Non-URM Students	
2009				-0.048	-0.024	-0.011
				[0.214]	[0.571]	[0.214]
2010	0.021	**0.094*** **	-0.012	**-0.041***	0.028	-0.008
	[0.929]	[0.000]	[0.357]	[0.071]	[0.571]	[0.429]
2011	0.064	**0.156*** **	-0.018	**-0.081***	0.090	-0.011
	[0.643]	[0.000]	[0.214]	[0.071]	[0.143]	[0.500]
2012	0.021	**0.171*** **	-0.020	**-0.180*** **	0.046	-0.014
	[0.929]	[0.000]	[0.143]	[0.000]	[0.500]	[0.429]
2013	-0.032	**0.129*** **	-0.019	**-0.179*** **	0.024	-0.015
	[0.786]	[0.000]	[0.357]	[0.000]	[0.714]	[0.429]
2014	-0.027	**0.129***	-0.023	**-0.256*** **	-0.014	-0.016
	[1.000]	[0.071]	[0.429]	[0.000]	[0.857]	[0.571]
2015	-0.032	**0.206*** **	-0.039	**-0.342*** **	-0.058	-0.027
	[0.714]	[0.000]	[0.214]	[0.000]	[0.571]	[0.357]
Mean (unlogged) in year before policy adoption	1,069	6,317	0.145	1,944	14,038	0.122

*** p < 0.01, ** p < 0.05, * p < 0.10

Notes: Effects are estimated using the synthetic control method. Reported p-values (in brackets) are derived from placebo permutation tests and account for the quality of each placebo match in the pre-treatment period.

Table 6

Estimated Effects of Performance-Based Funding on Bachelor's Degree Production in Tennessee and Ohio, by URM Status and Year

	(1)	(2)	(3)	(4)	(5)	(6)
	Tennessee			Ohio		
	Log Number of Degrees Awarded		Share of Degrees Awarded to URM Students	Log Number of Degrees Awarded		Share of Degrees Awarded to URM Students
Year	URM Students	Non-URM Students		URM Students	Non-URM Students	
2009				0.007	-0.008	0.006
				[0.813]	[0.688]	[0.375]
2010	**-0.037***	0.021	-0.006	0.048	**0.035*****	0.002
	[0.063]	[0.250]	[0.438]	[0.438]	[0.000]	[0.875]
2011	-0.012	0.048	-0.008	0.096	**0.096*****	0.002
	[0.438]	[0.250]	[0.375]	[0.188]	[0.000]	[0.750]
2012	-0.034	**0.072***	-0.018	0.065	**0.076*****	-0.002
	[0.375]	[0.063]	[0.250]	[0.438]	[0.000]	[0.750]
2013	-0.017	0.067	-0.016	-0.005	**0.114*****	-0.012
	[0.625]	[0.188]	[0.188]	[0.875]	[0.000]	[0.125]
2014	**-0.052***	0.030	-0.020	0.020	**0.123*****	-0.012
	[0.063]	[0.313]	[0.188]	[0.813]	[0.000]	[0.125]
2015	-0.054	0.036	-0.024	-0.049	**0.137*****	**-0.031***
	[0.250]	[0.313]	[0.313]	[0.438]	[0.000]	[0.063]
Mean in year before policy adoption	3,323	14,394	0.188	3,532	32,182	0.099

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

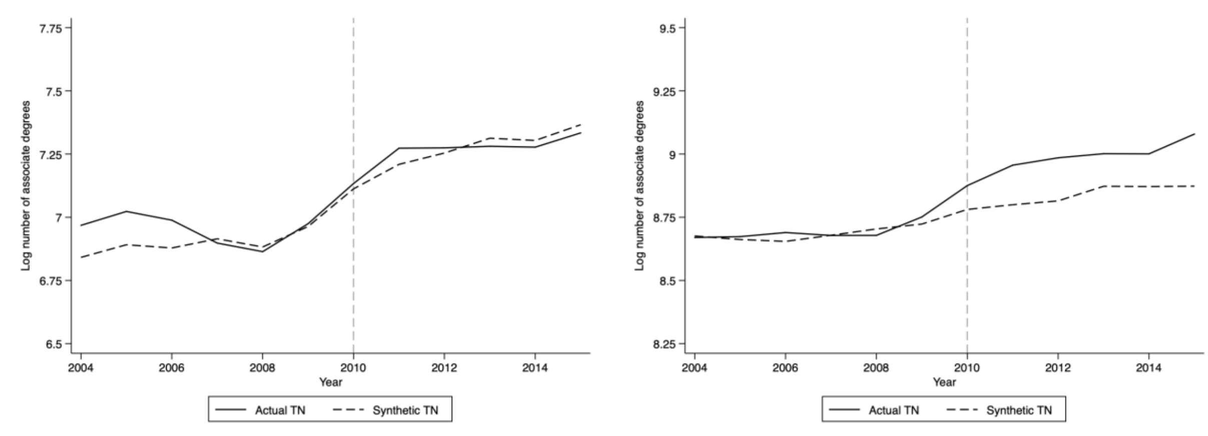
Notes: Effects are estimated using the synthetic control method. Reported p-values (in brackets) are derived from placebo permutation tests and account for the quality of each placebo match in the pre-treatment period.

Figure 1

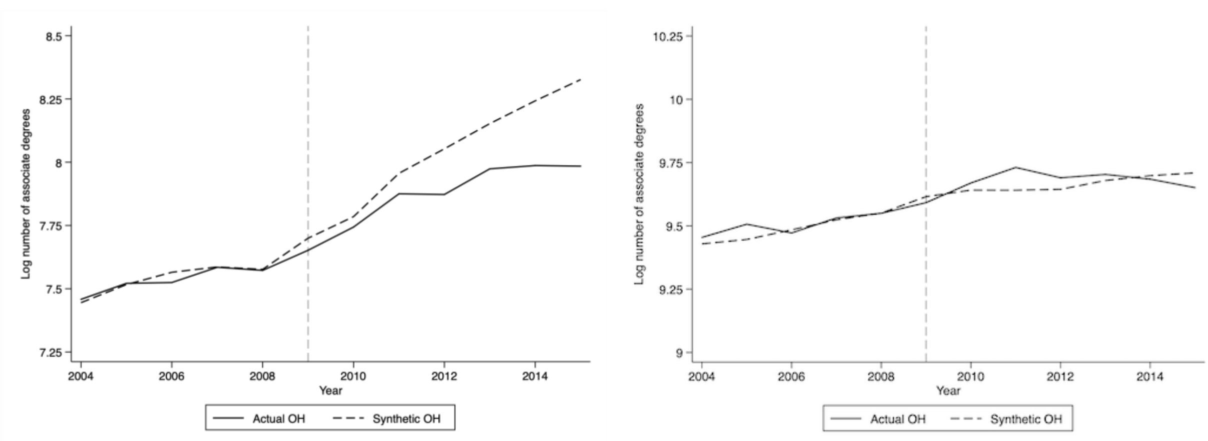
Trends in Associate Degree Conferrals to URM and Non-URM Students in Tennessee, Ohio, and Synthetic Control

URM Non-URM

A. Tennessee



B. Ohio



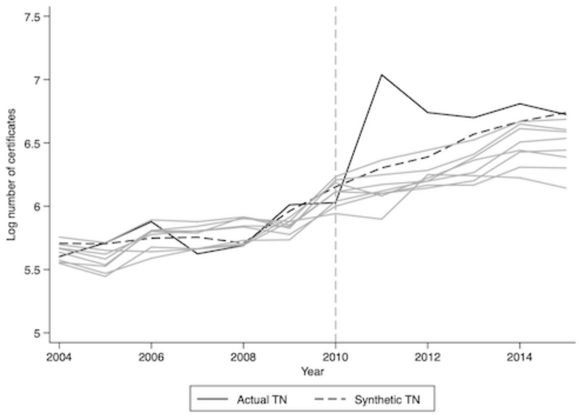
Notes: The synthetic control state is constructed by assigning weights to the donor pool that minimize the RMSPE of the outcome in the pre-treatment period.

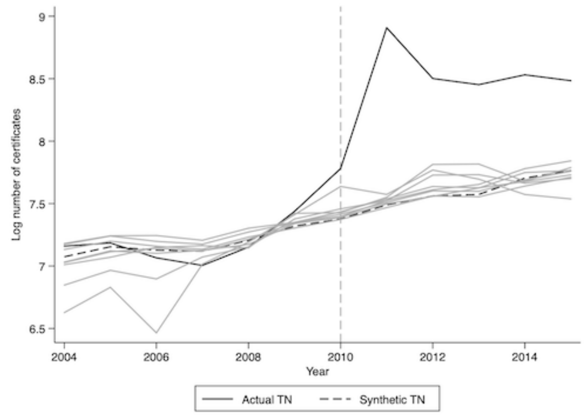
Figure 2

Robustness of the Effect Estimates in Tennessee for URM and non-URM students to Alternative Constructions of the Synthetic Control

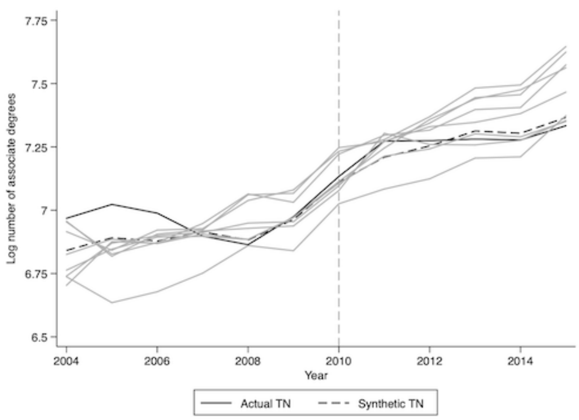
URM Non-URM

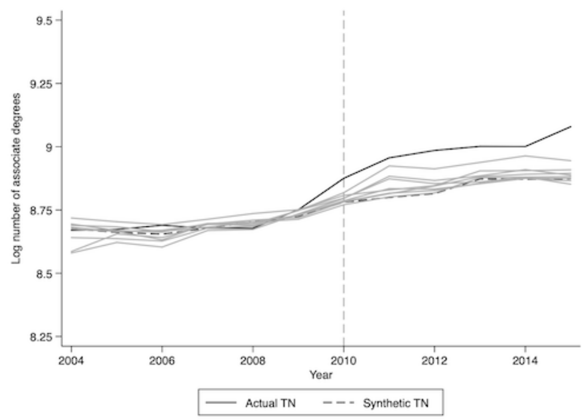
A. Log number of certificates awarded



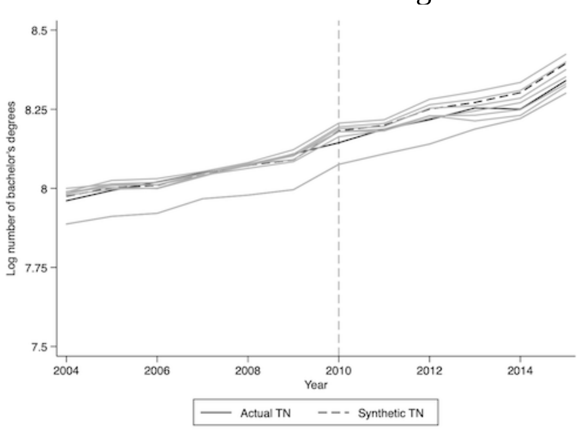


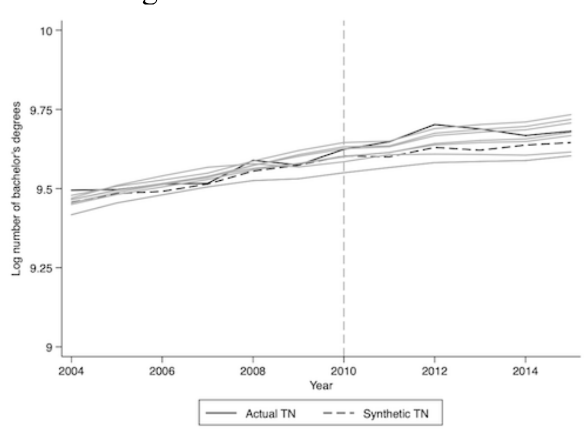
B. Log number of associate degrees awarded





C. Log number of bachelor’s degrees awarded





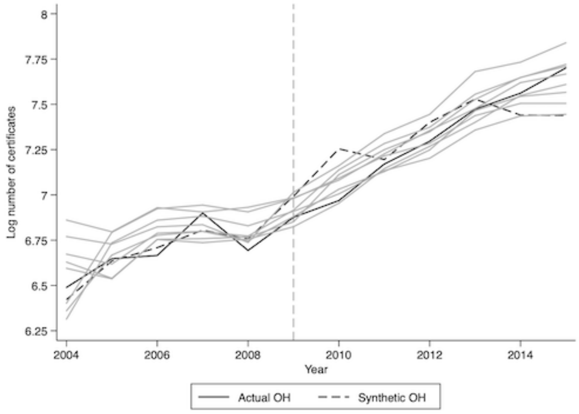
Notes: The dashed black line is the “optimal” counterfactual that minimizes the RMSPE of the outcome in the pre-treatment period. The grey lines depict alternative counterfactuals that best fit the data in the pre-treatment period for at least one other outcome-by-group combination in Ohio or Tennessee.

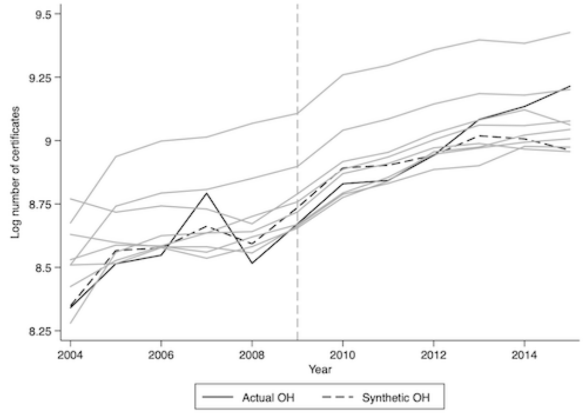
Figure 3

Robustness of the Effect Estimates in Ohio for URM and Non-URM Students to Alternative Constructions of the Synthetic Control

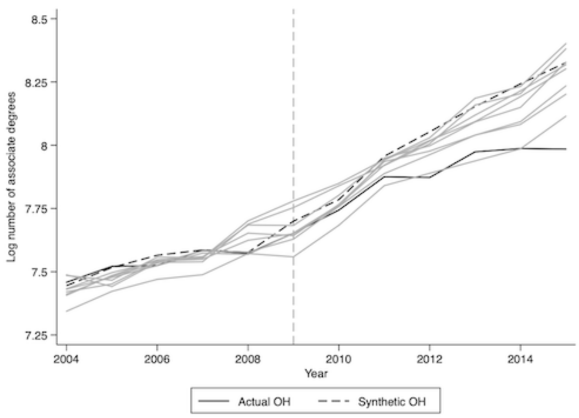
URM Non-URM

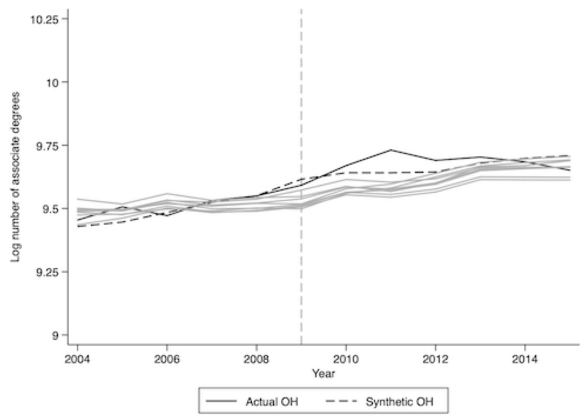
A. Log number of certificates awarded



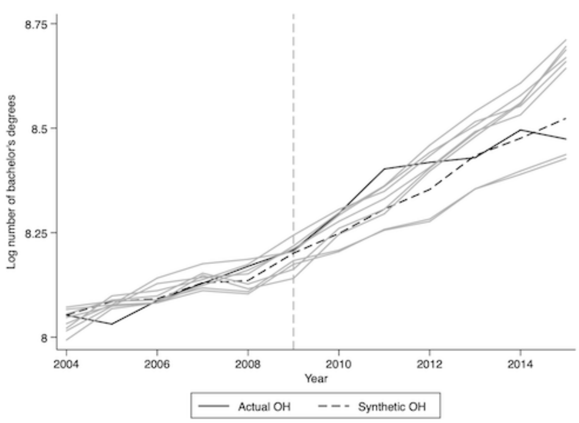


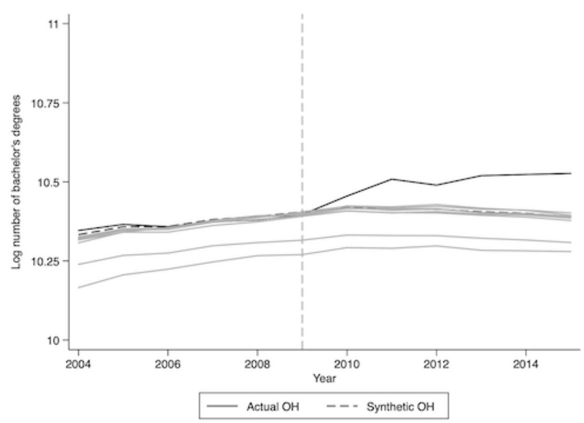
B. Log number of associate degrees awarded





C. Log number of bachelor's degrees awarded





Notes: The dashed black line is the "optimal" counterfactual that minimizes the RMSPE of the outcome in the pre-treatment period. The grey lines depict alternative counterfactuals that best fit the data in the pre-treatment period for at least one other outcome-by-group combination in Ohio or Tennessee.

Appendix

Appendix A: Identifying States with Performance-Based Funding

We identified 16 states as not having performance-based funding (PBF) from 2004-2015. Those states were Alaska, Alabama, California, Connecticut, Delaware, Iowa, Kentucky, Maryland, Nebraska, New Hampshire, New Jersey, Rhode Island, South Carolina, Vermont, and West Virginia. Nine of those states never adopted PBF as of fiscal year 2020 (Ortagus et al., 2020).

We relied on several sources to identify this set of non-PBF states. First, we examined existing studies of PBF and PBF with equity provisions (e.g. Gándara & Rutherford, 2018; Kelchen, 2018, 2019; Ward & Ost, 2021). Second, we consulted policy reports on PBF models published by policy advocates (e.g. Snyder, 2015; Snyder & Fox, 2016). Finally, we reviewed state policy documents available online from state agency webpages and information from the State Higher Education Executive Officers Association (SHEEO, 2019).

It is not straightforward to identify states with PBF policies or the specific features of their funding policies at any given time. PBF has been an active area of policy work in many states since the 2010s, and policies can change year to year. States have developed and implemented policies, developed but not implemented policies, or developed policies and partially implemented or phased in their policies over time (Snyder, 2015; Snyder & Fox, 2016). Arizona, for example, was recorded as having implemented PBF for some four-year institutions between 2012 and 2015 (Kelchen, 2018; Snyder, 2015); however, by 2016, Snyder and Fox noted that the state abandoned its plans to implement PBF across all public institutions. As another example, some researchers classify New York as a state that has adopted PBF in the two-year sector because the City University of New York (CUNY) system used PBF to appropriate funds to its community colleges (Kelchen, 2019). However, PBF advocates do not consider New York to be a PBF-adopting state given the absence of a statewide policy during our analytic time frame (Snyder, 2015; Snyder & Fox, 2016). Academic researchers have also raised questions about New York as a PBF-adopting state given the inconsistency between stated performance indicators and actual budgetary practices (Kelchen, Ortagus et al., 2019).

Given the dynamic policy environment and variation in interpretation, there are also inconsistencies among academic researchers regarding which states have adopted PBF in recent years. Kelchen (2018) identifies seven states as having PBF that are considered non-PBF states in Ward & Ost (2021). New Mexico has been described as implementing PBF for four-year institutions since 2004 in one study (Gándara & Rutherford, 2018) and since 2012 in another (Kelchen, 2018). There are also discrepancies between studies with respect to states that have implemented PBF policies with equity provisions. For instance, Indiana has been described as having PBF with equity provisions since 2004 (Kelchen, 2018), and in another study as having PBF (without equity provisions) since 2007 and equity provisions beginning in 2010 (Gándara & Rutherford, 2018).

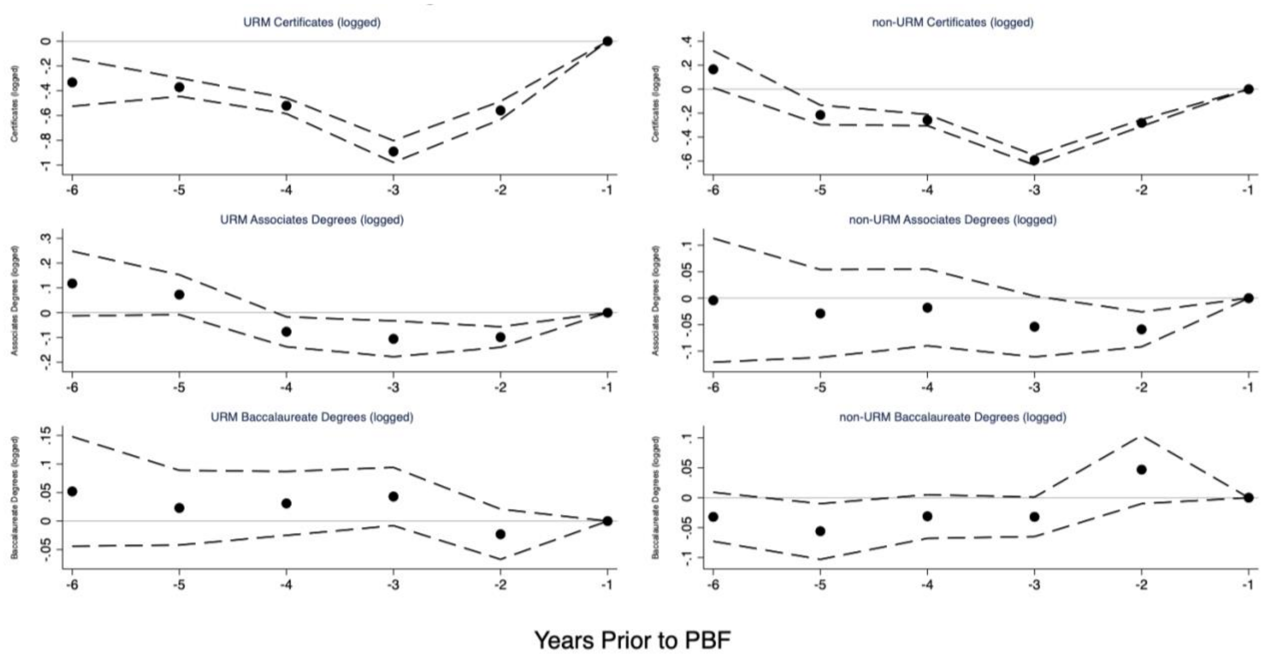
We attempted to address these discrepancies in the literature by creating the most restrictive control group possible. If a state was described as adopting PBF at some point in either or both the two- or four-year sector over our study period in at least one resource we consulted, we excluded it from the donor pool. We did this so that we could use a consistent set of non-PBF states to construct the counterfactuals for Tennessee and Ohio across all outcome-by-group combinations.

Figure A1

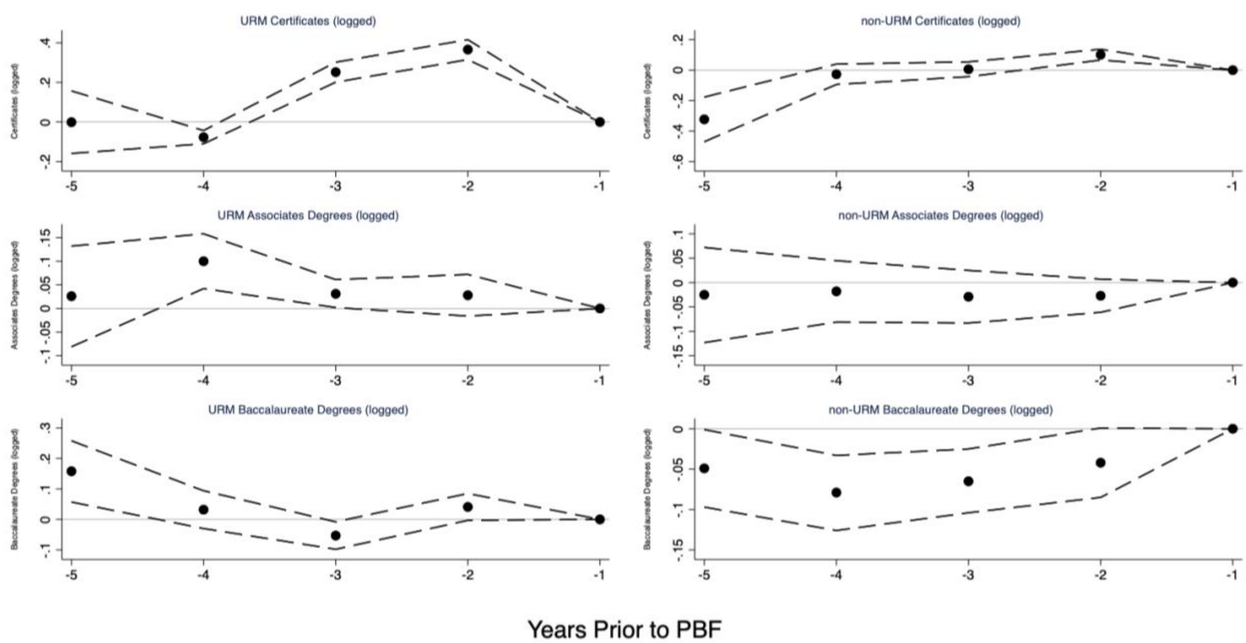
Evidence of Violation of Parallel Trends Assumption in Difference-in-Differences Research Design in Tennessee and Ohio

URM Non-URM

A. Tennessee



B. Ohio

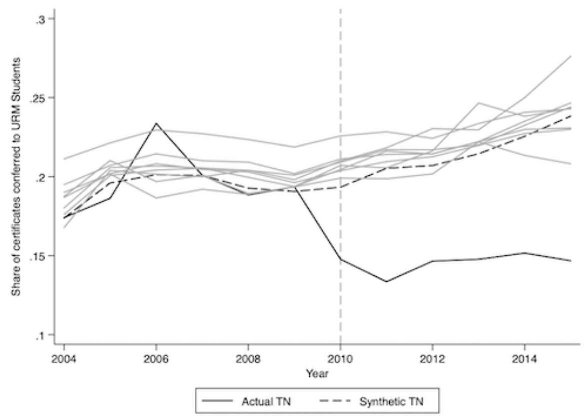


Notes: Each circle reports the estimated outcome difference between public institutions in Ohio (Tennessee) and comparison states in the years preceding PBF 2.0 adoption. The reference period is the year immediately preceding the passage of PBF 2.0 legislation. The dashed lines show the 95% confidence interval of the estimates.

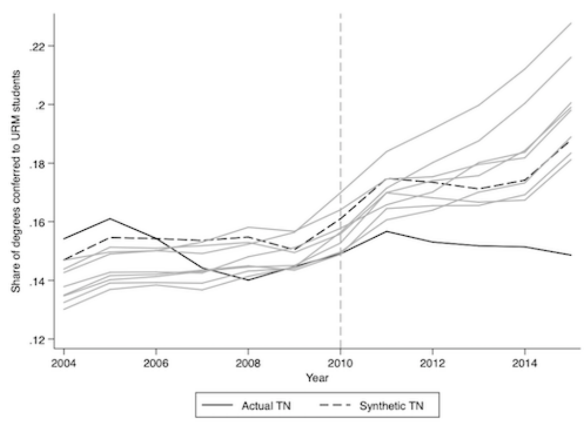
Figure A2

Robustness of the Share of Certificate and Degree Effect Estimates Conferred to URM Students in Tennessee to Alternative Constructions of the Synthetic Control

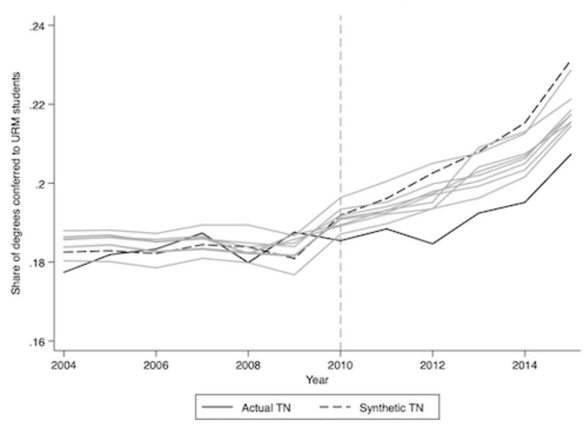
A. Less-than-Two-Year Certificates



B. Associate Degrees



C. Bachelor’s Degrees

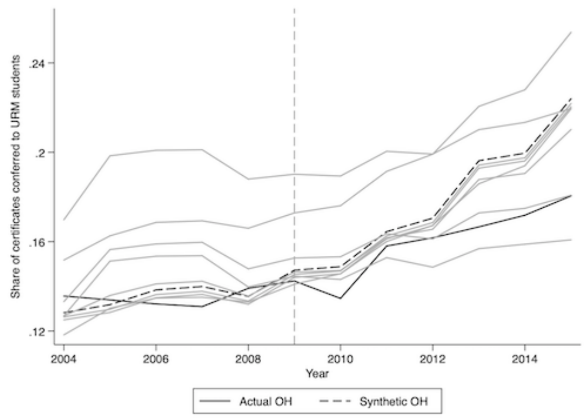


Notes: The dashed black line is the “optimal” counterfactual that minimizes the RMSPE of the outcome in the pre-treatment period. The grey lines depict alternative counterfactuals that best fit the data in the pre-treatment period for at least one other outcome-by-group combination in Ohio or Tennessee.

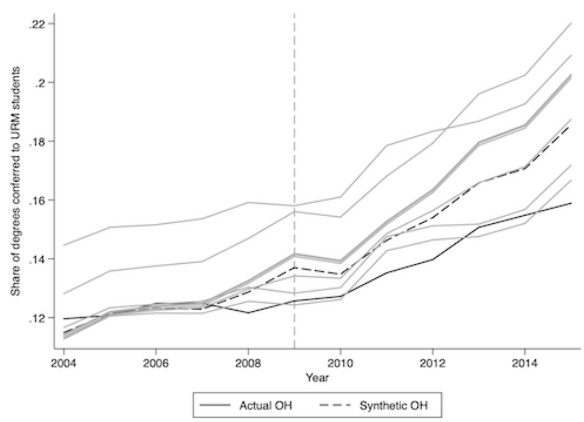
Figure A3

Robustness of the Share of Certificate and Degree Effect Estimates Conferred to URM Students in Ohio to Alternative Constructions of the Synthetic Control

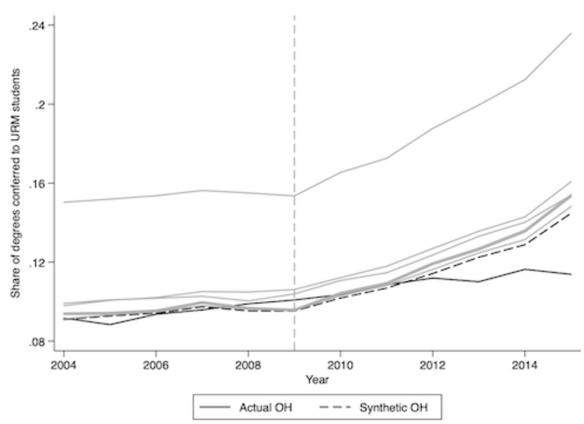
A. Less-than-Two-Year Certificates



B. Associate Degrees



C. Bachelor's Degrees



Notes: The dashed black line is the "optimal" counterfactual that minimizes the RMSPE of the outcome in the pre-treatment period. The grey lines depict alternative counterfactuals that best fit the data in the pre-treatment period for at least one other outcome-by-group combination in Ohio or Tennessee.