Four Decades of Performance Funding and Counting

Amy Y. Li

Contents

Introduction ..................................................................................... 486
Definition and Prevalence ............................................................... 486
Main Features of Performance Funding .......................................... 487
Brief History of Performance Funding ............................................. 488
Theoretical Logic of Performance Funding ...................................... 489
New Public Management .................................................................. 489
Resource Dependence Theory ......................................................... 490
Principal-Agent Theory ..................................................................... 490
Neoliberalism .............................................................................. 491
Ecology of Games ......................................................................... 492
Establishment of Performance Funding ........................................... 493
Adoption of Performance Funding Policies ..................................... 493
Advocates of Performance Funding and Evidence Usage ............... 494
Policy Implementation and Campus Responses .............................. 496
Intended Institutional Responses in Indiana, Ohio, and Tennessee .... 497
Unintended Institutional Responses in Indiana, Ohio, and Tennessee . 498
Challenges to Implementation in Indiana, Ohio, and Tennessee ...... 499
Policy Awareness and Campus Responses in Washington State ....... 500
Campus Implementation in Ohio and Pennsylvania ....................... 501
Campus Responses to Equity Metrics in Ohio and Pennsylvania ....... 502
Policy Evaluation .......................................................................... 502
Policy Evaluation Considerations .................................................. 503
Four-Year College Completion Outcomes ....................................... 516
Two-Year College Completion Outcomes ...................................... 519
Meta-Analysis on Outcomes and Access ......................................... 525
Outcomes Specific to Minority Serving Institutions ....................... 525
Access and Admissions Outcomes .................................................. 528
Equity Metrics and Enrollment Outcomes at Four-Year Colleges ....... 530
Performance Funding Impacts on Institutional Finances ................. 532

A. Y. Li (*)
Department of Educational Policy Studies, Florida International University, Miami, FL, USA
e-mail: AmLi@fiu.edu

© Springer Nature Switzerland AG 2021
L. W. Perna (ed.), Higher Education: Handbook of Theory and Research,
Higher Education: Handbook of Theory and Research 36,
https://doi.org/10.1007/978-3-030-44007-7_8
Abstract

Performance funding policies for higher education tie a portion of state revenue to public colleges according to student outcomes. This chapter is a comprehensive literature review of research studies on performance funding, highlighting the conceptual frameworks employed by authors and focusing on three areas of the policy: development, adoption, and design; campus implementation and responses; and ultimate impacts on retention and completion, access, student enrollment demographics, and institutional finances. This chapter pays attention to equity considerations of performance funding and discusses the policy’s impacts on colleges with different levels of resource capacity, selectivity, and student characteristics. The chapter concludes with considerations for future research on the policy.

Keywords

Performance funding · Performance-based funding · Outcomes-based funding · Policy implementation · Policy evaluation · Accountability · Equity metrics · Higher education finance

Introduction

Definition and Prevalence

The basic definition of performance funding (PF) is that it is a state-level funding formula that ties monies “directly and tightly to the performance of public campuses on individual indicators” or outcomes (Burke and Minassians 2003, p. 3). The policy differs from “traditional” forms of funding that rely more heavily on enrollment numbers to calculate state appropriations to higher education. PF is higher stakes than other possibilities, say performance reporting, which relies on outcome indicators as one consideration in determining state appropriations (Burke and Minassians 2004).

Performance funding at its core is a simple concept. It follows the basic public policy tools of “carrots” and “sticks,” a system of incentives and potential punitive consequences (Li 2018). Essentially, if public institutions are paid in part based on
graduation outcomes and other measures of student success deemed important, then public institutions should, intuitively, find ways in their internal practices to accomplish stated goals. However, colleges are complex organizations with many layers of operation, and the process to produce a graduate is not simplistic.

One challenge is that the perennial definitions of success in higher education are varied, depending on student goals and dependent upon institutional mission. One way to consider the outcomes of higher education is to view a student’s degree as providing a return on investment, where the student earns a higher salary than had they not earned a degree. Taking a purely economic approach to return on investment, students who complete college tend to have higher salaries than those who have a high school education. Furthermore, the taxpayer monies provided to public higher education generates a higher return on investment for students who graduate, versus those who do not. This rationale was an impetus behind the growth of accountability in higher education.

PF emerged from broader public concerns that too many students were entering college but not completing a degree. As of 2020, 62% of first-time, full-time students who entered 4-year colleges completed a bachelor’s degree in 6 years. At 2-year colleges, just 33% of first-time, full-time students completed an associate degree in 3 years (National Center for Education Statistics 2020b). These numbers exclude part-time students, who take even longer to complete a credential. Thus, college completion rates are a major reason why states have pursued PF as a public accountability tool.

Main Features of Performance Funding

A report from 2020 documented that 30 states had some form of PF in which 2- or 4-year public colleges received some portion of their state funding based on student outcomes (active and currently funded) (Rosinger et al. 2020). Two additional states (Idaho and Missouri) had PF policies in place but did not allocate funds to colleges according to the funding formula. Pennsylvania, which has a long history of PF, had, at the time of the report, paused its PF system while in the process of redesigning. Among the 30 states with actively funded policies, the majority (21 states) operated PF for 2- and 4-year colleges. One state (New Jersey) only operated PF for 4-year colleges, while the remaining eight states operated PF only for 2-year colleges.

PF policies across states universally incentivize graduation rates or degree completion (at the 2- and/or 4-year level). Other commonly used metrics include retention measures, such as first- to second-year retention or the completion of certain thresholds of credit hours (15 credits, 30 credits), transfer from 2- to 4-year colleges, and job placement or graduate wage outcomes. In addition, some states fund degree completion in STEM fields, which are considered to be a high demand area to fulfill state economic needs (Li 2018; Rosinger et al. 2020).

In recent policies, states have also created equity metrics to incentivize the retention and graduation of students from historically underserved populations. These typically include low-income (e.g. Pell grant recipients) and students of
color (most frequently defined as Black/African American and LatinX students, occasionally Native American students or Native Hawaiian students). As of FY 2020, more than two-thirds of PF states provide bonuses or have a separate metric for graduating low-income students, and half of the states explicitly recognize race as part of their funding formula (Rosinger et al. 2020). Less frequently, equity metrics offer extra funding for adult students aged 25 and older as well as first-generation students (Li 2018).

One feature of PF is the proportion of total state appropriations that is tied to performance on student outcomes. One could view PF policies as a drug treatment, with the percentage of operating revenues for an institution as the dosage of the treatment. States vary substantially in the percentage of funds tied to performance. As of FY 2020, Arkansas allocated 3% of funding according to outcomes, Nevada allocated 20%, while Ohio allocated 100% based on outcomes (Rosinger et al. 2020). While some states such as Ohio and Tennessee are known for allocating virtually all state appropriations to colleges based on a PF formula, the majority of states with the policy hover around the 5–10% amount (Li 2018).

**Brief History of Performance Funding**

In the United States, PF for higher education was first established in Tennessee (Banta et al. 1996). The state has gone through multiple revisions of the policy (Sanford and Hunter 2011), which has continued until today. Ohio also established one of the earliest versions of the policy in 1995 and 1997 (Dougherty and Reddy 2013) and, having gone through phases of revision, currently has a form of the policy.

PF has gone through periods of popularity and discontinuation, whereby some states began the policy, discontinued it, and then re-established it, while other states have continued the policy across time. A number of states adopted PF in the 1990s yet abandoned their policies during the “dot com” recession in the early 2000s (Dougherty and Natow 2015). Earlier versions of the policy, sometimes dubbed PF 1.0, operate as bonus funding, on top of regular state appropriations for higher education, in an amount typically about 1–5% of state funding (Burke 2002; Dougherty and Reddy 2013). In other words, colleges were offered an additional incentive above their base appropriations, if certain student outcome goals were met.

The second wave of policies (newly adopted or revised from previous iterations) was called PF 2.0. Rather than allocating performance-based funds as a bonus, PF 2.0 took existing base state funding and required institutions to earn back that funding by meeting certain performance goals. The portion of funding tended to be higher as well, upwards of 80–100%. The key feature of PF 2.0 is that colleges can actually lose funding from previous years if they do not meet outcome criteria designated by the funding formula, making base-funding PF policies a high-stakes accountability mechanism (Dougherty and Reddy 2013).

Policies adopted in the 2000s and especially 2007 onward tended to be considered PF 2.0, as concerns over efficiency and accountability in higher education
fueled the popularity of outcomes-based funding tools. Competition from other funding needs such as the increasing cost of healthcare, and K-12 appropriations, has also increased the desire of policymakers to use PF as a way to improve graduation outcomes at public colleges (Dougherty and Natow 2015). In conducting this literature review, I focus primarily on journal articles and books on PF. I recognize there are a number of studies published by research institutes and advocacy organizations, but I depend primarily on peer-reviewed articles and edited books. For an in-depth literature review of PF policies prior to 2013, see Dougherty and Reddy (2013). Reviewed in this chapter, Dougherty et al. (2016) offers detailed analyses of campus implementation of PF. Furthermore, see Jones et al. (2017) for discussions focused on equity and racial justice. For a summary of main components of PF, see the policy report by Li (2018). In this chapter, I provide more detailed literature reviews of the most recent studies on PF, particularly journal articles in the latter half of 2010s up to the most recent available at the time of writing this chapter.

Theoretical Logic of Performance Funding

There are several commonly applied theoretical frameworks in studies of PF. These include new public management, resource dependence theory, principal-agent theory, neoliberalism, and less commonly, the ecology of games. In this section, I survey the fundamentals of each theoretical perspective and describe how scholars have applied these frameworks to PF.

New Public Management

New public management, a theory that originated from applying business-like practices to public service organizations, assumes that market competition and contracting out government services will lower the cost of providing goods and services (Meier and O’Toole 2009). Transparency is also important in new public management, as performance data must be tracked and reported. New public management represents an effort to control public service professions by prioritizing the explicit measurement of output and outcomes and linking financial rewards to such outcomes (Frølich 2011). Previous studies have applied new public management ideas to the performance management goals of PF policies (Gándara and Rutherford 2018).

As intuitive as PF should be in theory, the research from new public management shows that among public organizations, paying for outcomes is rife with challenges. Focusing on a set of measurable outcomes is difficult when the process of achieving an outcome (e.g., graduation) is shaped by factors outside of the institution’s control, such as student academic readiness (Bell et al. 2018). An ongoing challenge in designing the weighting of performance metrics is to try and ensure that the pursuit of degree completion does not exacerbate persistent educational participation and attainment gaps across students of differing socioeconomic advantages.
**Resource Dependence Theory**

Resource dependence theory (RDT), which emerged from organizational theory and strategic management, has been used in multiple writings on PF. Because public colleges are dependent on funds from the state, their dependency on this material incentive will catapult some type of internal behavioral change aimed at maximizing such resources (Pfeffer and Salancik 1978). In other words, within states that have a PF policy, colleges will act in ways aimed at improving performance because colleges are provided material incentives from the state (Dougherty and Hong 2006; Dougherty and Reddy 2013). Studies incorporating a resource dependency perspective typically suggest that colleges that rely more heavily on state appropriations (such as community colleges) will be more likely to engage in the desired behaviors of improving stated student outcomes (Hu 2019; Li and Ortagus 2019).

**Principal-Agent Theory**

Another frequently cited conceptual framework behind PF is principal-agent theory (Li 2020; Tandberg and Hillman 2014). Using this theory, the state is considered the principal, and individual colleges are considered agents acting on the principal’s behalf to achieve a goal, in this case, college completion. The principal allocates funding according to a pre-determined formula, and expects the agent to deliver on goods, services, and/or goals aligned with the principal’s interests. However, oftentimes, the principal-agent problem emerges because principals and agents are presumed to have differing interests, and it is nearly impossible for the principal to strictly monitor all of the agent’s actions (Bohren 1998).

In other words, colleges have more knowledge about their internal workings, such as faculty teaching practices, student advising, and mentorship, than states have the ability to oversee. This imperfect awareness of internal workings is called information asymmetry (Kivistö 2008). Colleges may be tempted to maximize their own utility (take actions towards forwarding their own goals, such as increasing research productivity and prestige) rather than the stated goals of a PF policy (such as increasing the graduation rate of Pell grant recipients). Information asymmetry can be detrimental due to agency loss – the loss of efficiency due to the principal’s goals being different from the agent’s own self-interested actions (Holmstrom 1979).

Nearly all studies of PF incorporate some element of principal-agent theory to conceptualize the relationship between the state and its colleges. One complication that can emerge from this principal-agent relationship is colleges may “shirk” their responsibilities to the state (Holmstrom 1979; Lane 2007). That is, even though both colleges and states have the shared goal of improving degree completion, the process by which colleges achieve such goals may be nebulous. Instead, colleges may be tempted to use easier routes to confer more degrees, such as becoming more selective in admissions by accepting students who are more likely to graduate (Dougherty et al. 2016). Therefore, PF policy designers attempt to diminish “agency loss” by designing performance metrics collaboratively with colleges, to better align
principal-agent goals, or by phasing in PF policies to provide more time for colleges to adapt (Mckeown-Moak 2013).

Not only does principal-agent theory apply to the state and college relationship, but it also describes within-college relationships. High-level administrators at a college serve as agents when they interface with the external principal (the state) but also serve as principals when they direct the activities of mid-level administrators, faculty members, and support staff. For example, provosts must communicate PF goals to deans of individual colleges, and these deans must disseminate such information within their departments. Thus, the internal structure of the college also represents layers of the principal-agent relationship (Li 2017b).

Neoliberalism

Neoliberalism is another concept that arises in discussions of PF, such as in writings by Dougherty et al. (2016), and focuses on making public organizations more effective and economically efficient. A key feature of neoliberalism is that the organizations benefit from private enterprise, and a competitive marketplace, and government operations are made more efficient by privatization. If organizations cannot be fully privatized, governments need creative fiscal incentives for agents to compete with one another more efficiently and to better achieve the goals that principals designate (Dougherty and Natow 2019).

The desire for states to measure and calculate efficiency and to maximize quantifiable outputs represents the neoliberal components of PF for higher education (Letizia 2016). Neoliberal theories are used to frame a discussion in which institutions are more market responsive and better contributors to their state’s economic vitality. Neoliberalism shares commonalities with principal-agent theory. Principals such as states may be supporting neoliberal policies to forward their own agendas, yet institutional agents could also be proponents of neoliberal reform policies, although the motives of principals may differ from those of agents. Some studies show evidence that institutional actors supported PF as a way to better demonstrate to legislators and policymakers how much colleges were contributing to achieving state goals (Dougherty et al. 2011).

In the establishment of PF policies, neoliberalism also helps to explain the power that private businesses had to influence government agencies, by promoting the positives of business-like incentives to make public colleges more efficient. Even if businesses were not directly involved in crafting PF, they did promote neoliberal ideas through ideological influence, such as in South Carolina, Washington, Florida, and Missouri (Dougherty and Natow 2015).

Relatedly, businesses exerted power in the creation of PF because state policymakers were keener on policies that were attractive to businesses, even if businesses did not directly lobby for such policies (Dougherty and Natow 2015). For example, in Missouri, conservative businesses resisted additional public funding for higher education and contributed to an environment where PF seemed appealing to state policymakers.
While neoliberalism does capture the financial incentive and principal-agent piece of PF policy adoption, neoliberal ideas do not fully capture the way PF has been implemented. Monetary incentives are primary in applying neoliberal ideas via PF, but there are additional policy instruments used during the implementation of PF. One such incentive was persuasive communication, when the principal communicates the importance of a certain policy goal to better cultivate the interests of the agent (Dougherty and Natow 2019). For instance, in Pennsylvania, policy designers collaborated with institutional actors to design metrics consistent with the broader goal of greater college completion and further provided flexibility for colleges to choose metrics aligned with their own institutional strategic plans (Cavanaugh and Garland 2012; Zumeta and Li 2016).

A second implementation consideration that is not captured by neoliberal theories is the idea of capacity building, which entails providing or helping to facilitate the increase in resources that colleges need in order to be more successful at achieving stated PF goals (Dougherty and Natow 2019). Such resources could include greater analytical capacity, a technology system for using predictive analytics, “intrusive” advising, and resources to hire academic or student support staff. In the case of Ohio, institutions did invest in campus-level resources to be better positioned to achieve performance outcomes (Dougherty et al. 2016; Li and Zumeta 2016).

**Ecology of Games**

In explaining why PF policies have limited impact on student outcomes, Nisar (2015) points to the ecology of games perspective whereby universities are players engaged in multiple games. Games consist of interactions between institutional actors based on a set of rules. Colleges operate within a larger ecosystem and play multiple roles within this ecosystem and engage in games involving calculable strategies and tactics that give colleges a sense of success or failure.

For most colleges, the state funding game is perhaps not the primary game of interest. Rather, colleges may be interested in activities involving donors, in acquiring more prestige and influence, in seeking research and development funds, in moving up in the rankings, or in other engagements with higher payoffs and lower compliance costs. The relationship between the state and the colleges is not a linear cause and effect relationship that can be moderated using performance-based funds (Nisar 2015).

Additionally, the behavior of colleges depends on multiple games played internally, such as related to departmental resource allocations, tenure policies, and student recruitment. These behaviors can be unpredictable. Furthermore, the actors within a college change over time, along with their techniques and strategies in seeking goals beyond student retention and completion. Therefore, the establishment of a PF policy represents only one small portion of the ecology of higher education and is arguably not influential enough to fundamentally change the behavior of colleges. Nisar (2015) recommends that PF systems not be rigid in terms of design and time horizon. Since colleges and environments are actively changing, PF systems may need to be adjusted over time.
Establishment of Performance Funding

PF has grown to affect the majority of states, and the establishment of the policy has been a topic of research. Government agencies within the United States have been fundamental in establishing PF in various states, including Tennessee, the original innovator of PF (Dougherty and Natow 2015). A few studies have explored the factors related to the policy adoption and diffusion of PF across states and influences within states that contributed to the policy’s rise in popularity. This section outlines literature analyzing the origins of the policy, influential advocates of the policy, and the use of research evidence in crafting policies.

Adoption of Performance Funding Policies

Policy Innovation and Diffusion of Earlier PF Policies. One study examined the adoption of PF policies through a policy innovation and diffusion framework, from years 1979 to 2002. Similar to other diffusion studies examining how policies spread across contiguous states, this study excluded Alaska and Hawaii, because they have no neighbors, and similar to national studies of PF adoption and outcomes, this study excluded Nebraska because of the state’s non-partisan legislature (McLendon et al. 2006). Using event history analysis techniques, specifically a discrete-time logit model for nonrepeatable events, the authors determined which internal state characteristics, and neighboring influences, were associated with the adoption of PF policies as well as performance budgeting. While PF ties state appropriations to outcomes using a formula, performance budgeting differs in that it allows state officials “to consider campus achievement on performance indicators as one factor in determining allocations for public campuses” (Burke and Minassians 2003, p. 3). The literature recognizes performance budgeting as a lower-stakes strategy for accountability (Burke 2002).

McLendon et al. (2006) showed that a higher percentage of Republican legislators in a state increased the probability that a state adopted PF. The authors theorized that Republicans in the state legislature may have viewed PF policies more favorably than Democrats because PF policies are a form of high-stakes accountability aimed at monitoring the public bureaucracy of higher education.

McLendon et al. (2006) also found that states with a less centralized governance structure for higher education (a coordinating board or a planning board) were more likely to adopt PF, compared to states with more centralized governance structures (in the form of a consolidated governing board). The researchers theorized that consolidated governance boards may better represent the interests of higher education actors and serve to protect institutions from more rigid external accountability structures such as PF, whereas a less powerful board may not be able to (McLendon et al. 2006).

Interestingly, the researchers found that other state characteristics examined did not predict the adoption of PF, specifically educational attainment, percent change in gross state product, legislative professionalism, governor’s powers, the party of the
governor, percent changes in tuition, and percent changes in public enrollment (McLendon et al. 2006).

**Policy Learning and Diffusion of PF 2.0 Policies.** Since PF policies have especially expanded in popularity in the last decade, a more recent study on state adoptions of PF examined the adoption patterns among the 47 states between 2000 and 2013, focusing on PF 2.0 (Li 2017a). Using a Cox proportional hazards model, the author found that contrary to policy theories predicting that states emulate one another in adopting policy innovations (Berry and Berry 1990), there was a reverse diffusion effect for PF 2.0. States that had more neighbors that operated PF 2.0 or were in the process of adopting PF 2.0 were less likely to adopt the policy in their own state, suggesting a policy learning effect where states delay their own adoption of PF 2.0 by perhaps waiting to see the effectiveness of the policy emerge in neighboring states (Li 2017a).

Li (2017a) also found that states with a Republican-majority legislature were more likely to adopt PF 2.0, but the party of the governor was unrelated to adoption patterns. On the other hand, states with more liberal electorates, measured by citizen ideology, and states with more professionalized legislatures, characterized by greater staff resources, were more likely to adopt the policy.

Other variables that increased the probability of adopting PF 2.0 included higher levels of state appropriations per student and more rapidly growing public enrollment numbers. On the other hand, states with higher educational attainment levels were less likely to adopt PF 2.0. Variables unrelated to PF 2.0 adoption included governor’s budget powers, the percent of enrollment at private colleges, the percent of 2-year enrollment, state grant aid per student, tuition rate changes, and outmigration (Li 2017a). This study showed that factors related to the adoption of higher-stakes PF 2.0 policies differed from those related to earlier policy adoptions in the PF 1.0 era. The Li (2017a) and McLendon et al. (2006) studies illustrated that factors including political conditions helped to explain the spread of PF across states.

**Advocates of Performance Funding and Evidence Usage**

Political leaders, foundations, and advocacy organizations have been influential players in the establishment of PF. Some advocacy organizations, in the interest of garnering legislative support for PF, may be less likely to highlight empirical studies that run counter to a PF advocacy goal. Rather than evaluating the entire breadth of empirical literature, proponents of PF may focus on studies that offer more promising evidence about the effectiveness of PF (Bell et al. 2018).

Gándara et al. (2017) explored the role of Complete College America (CCA) in influencing policy activity around college completion, including PF, in Georgia, Tennessee, and Texas. The researchers collected data in the form of interviews of policy actors, observations of policy events, and archival documents. Findings suggested that CCA utilized rewards to incentivize states into considering PF, in the form of financial support for implementation, such as a predictive analytics software program. Further, CCA employed punishments such as shaming
institutions for low graduation rates. CCA portrayed PF as a best practice. In addition, national groups such as state intermediaries were also instrumental for legislators and staff to learn more about PF and understand what other states were doing. CCA also engaged in a practice of discrediting research that showed limited benefits from PF. Respondents of interviews also noted that CCA’s marketing tactics, such as focusing on the singular goal of college completion, and their easy to understand publications, were effective at advocating for PF policy adoption (Gándara et al. 2017).

Miller and Morphew (2017) examined the policy publications and political rhetoric prominent in the adoption of PF in Florida, Massachusetts, and Montana and, in particular, the influence of agenda-setting organizations. Specifically, the authors considered documents published by the Gates Foundation and the Lumina foundation, along with three additional advocacy organizations: Complete College America, HCM Strategists, and Jobs for the Future. The agenda-setting organizations framed the need to adopt PF in order to address the problems of low persistence and graduation rates, and a lack of graduates in the STEM fields. Organizations framed PF as beneficial for students and for state systems, to prioritize important goals in higher education. Failures from previous PF policies were attributed to poor implementation, lack of data, poor design, or lack of political will. In short, advocacy organizations were influential in diagnosing a need for PF within enrollment-based funding formulas and promoting PF as a solution in creating greater accountability in higher education (Miller and Morphew 2017).

Gándara (2019) explored how evidence was used in the formulation and implementation of PF policies in Colorado and Texas. She construed evidence as academic research, policy reports, data, stakeholder information, and best practices. During 2014 to 2015, the author conducted interviews with policy actors and campus administrators, observed legislative and state higher education agency meetings, and reviewed documents, primarily consisting of newspaper articles, meeting minutes, and email correspondence. Findings suggested that demanders of information included legislators and their staff members. The state higher education agency, as well as campus administrators, occupied both the role of demander and supplier of information.

Sources of evidence utilized by participants in the Gándara (2019) study included within-state data, information from other states, guidance from intermediary organizations, anecdotal evidence, research evidence, common sense, legislative services, and public perceptions. The most common form of evidence use was conceptual – to enlighten the consumer of evidence and not necessarily address a specific problem. Information was also employed for political purposes, such as when campus representatives used internal modeling to advocate for funding formulas that benefited their own institutions. Less often, information was used for instrumental purposes – to inform future decisions or policy positions (Gándara 2019).

Interestingly, Gándara (2019) found that in Colorado, a financial bonus for underrepresented minority students was abandoned even though evidence illustrated its utility. The legislators retained a low-income student premium but did not want to
be involved with the politics of including a racial/ethnic minority premium (Gándara 2020). The author also found that institutions with higher levels of fiscal resources were more able to use evidence, and campus officials were more likely to use academic research. Legislators in particular preferred information that was definitive, relevant, and timely, related to the economy, and perceived as coming from a credible source.

Utilizing data collected during the same time period, Gándara (2020) focused her analysis on the design process of Colorado’s PF funding model (adopted in 2014), posing questions about target populations, how these populations were portrayed, the political power and influence of policy actors, and which groups benefited or were burdened. Her findings revealed three types of institutions that were the primary policy targets. Access institutions were 2- and 4-year colleges with nonselective or minimally selective admissions criteria. They gained funding under the new model and were construed as deserving of extra funding. Research institutions were portrayed as not deserving of state funding but still wound up retaining their funding levels under the revised model. Research universities held more political power and greater resources and were most influential in the design process. Rural institutions had experienced declining enrollments, were less vocal during the design process, and ultimately lost funding under the new model (Gándara 2020).

The Gándara (2019, 2020) studies offer an insightful look at the policy creation process of PF, with particular close analysis paid to the sources and usage of information, the actors involved in information usage, negotiation of specific funding model features, and the varying degrees of political power exercised by different types of institutions. The studies reviewed on the adoption and creation of PF policies help to improve the understanding of the ways these policies were developed and crafted.

Policy Implementation and Campus Responses

After PF policies are adopted, there is another set of research questions regarding how policies are actually put into practice on campuses and how campuses respond to policies. In a large study, Dougherty et al. (2016) conducted interviews with state officials, state-level political actors, and administrators and faculty at nine 2-year colleges and nine 4-year colleges in Indiana, Ohio, and Tennessee. Interviews were conducted between August 2012 and April 2014. State-level interviewees included higher education commission officials, gubernatorial and legislative advisors and staff, business leaders, researchers, and consultants. Institutional actors included senior administrators, academic and non-academic administrators, and department chairs. The research team also analyzed documents including public agency reports, newspapers, and academic studies (Dougherty et al. 2016). The next section describes the Dougherty et al. (2016) studies on campus responses, both intended and unintended, and challenges associated with campus implementation of PF.
Intended Institutional Responses in Indiana, Ohio, and Tennessee

Campus Awareness of Performance Funding. Dougherty et al. (2016) found that college personnel, from university leadership to faculty and mid-level administrators, were aware that student graduation outcomes impacted their institution’s bottom line. Participants recognized that PF was valuable in prioritizing the need to improve student outcomes.

Extensive efforts by state policymakers were made in Indiana, Ohio, and Tennessee to disseminate information about PF to campus stakeholders. Evidence suggested that while broad awareness was high across campuses, faculty and mid-level administrators were less likely to receive direct communication about PF goals. State officials discussed building up campus organizational capacity to respond to PF but gave much latitude to campuses themselves to determine how to improve student retention and completion (Dougherty et al. 2016).

Changes to Academic Programs and Curriculum. A number of positive impacts were described by interview participants in response to PF. However, note that PF was just one of several concurrent influences aimed at improving student outcomes. It was empirically impossible to determine that campus responses were catalyzed solely by PF. All in all, the authors suggested that the coincidence of PF with other initiatives produced synergy rather than competition between different initiatives (Dougherty et al. 2016).

Dougherty et al. (2016) found that changes to developmental education commonly occurred in response to PF, especially at 2-year colleges but at 4-year universities as well. The PF policies in Ohio and Tennessee each had a performance indicator for developmental education, and developmental education reform was a part of other state-level initiatives in all three states. Changes in developmental education included allowing students to enroll the summer before their first fall term or to simultaneously enroll in developmental and college-credit bearing courses. At some colleges, instructional changes were also made to STEM courses, such as bringing new technology into the classroom.

A second prevalent change was that campuses improved course articulation and transfer, especially between 2- and 4-year colleges (Dougherty et al. 2016). This made it easier for students to get credit at 4-year colleges for courses they took at 2-year colleges. Transfers were also a performance metric in Ohio and Tennessee. Additional changes to academic programs included the creation of cohort courses, in which the same students progressed through a set of courses together.

Interestingly, Dougherty et al. (2016) wrote that Tennessee respondents noted the creation of shorter-term general education certificate programs, which produced substantial increases in such certificates, detected in later quantitative studies (Hillman et al. 2018; Li and Ortagus 2019). Some institutions reduced the number of credits needed for completing a program, standardizing bachelor’s degrees to require 120 credits, and associate degrees to require 60 credits. Finally, the researchers found that a few colleges increased online instruction, which added convenience for distance learning students (Dougherty et al. 2016).
Changes to Student Services. Campuses also made changes in the area of student support services (Dougherty et al. 2016). Specifically, campuses added more academic advisors, created online advising systems, and involved faculty to a greater extent in advising responsibilities. Colleges also implemented early warning systems to notify advisors of students at risk of dropping out, known as intrusive advising. Advisors incorporated degree maps to keep students more apprised of their progress toward graduation. In addition, some colleges reported creating new tutoring centers or started an online tutoring program. Other colleges increased their programming for first-year students (Dougherty et al. 2016).

There were also instances of colleges creating new scholarship programs and providing tuition discounts. Some colleges began charging students the same tuition amount for 12 or 15 credits, as an incentive to take more credits each term. Interviews revealed that some colleges streamlined registration and graduation procedures. These responses included the removal of a fee to apply for graduation and simplifying the application process. Additionally, colleges altered procedures such as reducing the number of times a student could withdraw from courses during their academic career, to encourage course completion.

Dougherty et al. (2016) discovered that some institutions consolidated the departments involved with advising and enrollment, making it easier for students to get answers to registration questions at a central location. Additional campus responses were observed at a few colleges, including improving career services offerings, creating living-learning communities, starting mentoring programs, or enhancing student organizations (Dougherty et al. 2016).

Unintended Institutional Responses in Indiana, Ohio, and Tennessee

Whether in direct pursuit of the ultimate outcomes of increased degree completion, PF policies have indirectly or directly led to impacts that were likely unintended by policy designers. The large-scale implementation studies conducted by Dougherty et al. (2016) have illustrated a number of potentially problematic responses to PF in the studied states of Indiana, Ohio, and Tennessee.

Increased Selectivity. Among interview participants in Dougherty et al. (2016), a commonly described response to PF was the restriction of admissions. Institutions faced pressure to improve their graduation outcomes and were less likely to offer acceptances to students who displayed lower academic preparation. Some colleges decided to require higher standardized test scores and GPAs for students to gain admissions or decreased the number of conditionally admitted students.

Second, colleges strategically targeted recruitment at more academically prepared students, such as suburban, out-of-state, and international students. Colleges were also less likely to recruit from high schools in urban or poorer areas. Furthermore, colleges showed evidence of shifting institutional aid from need-based to merit-based, to attract students with better prior academic performance (Dougherty et al. 2016).
**Reduced Academic Rigor.** Among some respondents in Dougherty et al.’s (2016) studies, the focus on improving course completion numbers resulted in faculty feeling pressured to inflate grades and pass students, who, in the past, would not have earned a passing grade. Additional concern was raised about the easing of grade forgiveness policies, and that students were more frequently being steered into “easier” courses.

In addition, some colleges reduced the number of credits required to earn a degree or made it easier to substitute courses to better facilitate student graduation. While streamlining the college curriculum could be viewed as a positive response, the impetus to alter curricular requirements should not be driven solely by PF incentives to graduate more students.

**Compliance Costs.** In interviews conducted by Dougherty et al. (2016), participants further described the extra work required to track students and report outcomes, which required staff expertise in institutional research. Yet, the resources needed to adhere to PF reporting standards were not provided to the colleges. Further, participants noted the added workload placed on faculty who were more frequently being asked to conduct assessment.

**Institutional Competition and Lower Morale.** A final unintended consequence of PF found through Dougherty et al.’s (2016) interviews in Indiana, Ohio, and Tennessee was reduced cooperation among institutions in a state who were competing against one another to earn performance funds. The sharing of best practices across institutions became stunted. Furthermore, some faculty and staff experienced declining morale and expressed frustration that their efforts were underappreciated (Dougherty et al. 2016).

### Challenges to Implementation in Indiana, Ohio, and Tennessee

The Dougherty et al. (2016) studies of campus implementation in Indiana, Ohio, and Tennessee also found a number of challenges for campuses to respond to PF. The first was that open-access institutions enrolled students from marginalized academic, economic, and social backgrounds who faced barriers in persistence and graduation. The student composition at some colleges made it difficult to report good performance on PF metrics.

At 2-year colleges, campus participants mentioned that some students who enrolled did not intend to ultimately graduate with an associate degree and were instead upgrading their knowledge or skill set by taking select courses. Some college respondents indicated it was unfair to be judged based on graduation outcomes when some entering students had no intention to earn a degree.

Campus respondents also pointed to limited institutional capacity at their institutions, specifically understaffing in institutional research and data systems that were outdated or tedious. Dougherty et al. (2016) found variations across institutional responses depending on type and capacity (measured by per-student revenue, data analysis capabilities, and enrollment percentage of historically marginalized
students). Lower-capacity colleges reported more challenges in meeting PF goals, especially due to enrolling a less privileged student population and having lesser developed data analytic ability.

According to Dougherty et al. (2016), because PF focused on aggregate outcomes that often could not be traced back, for instance, to the actions of individual faculty in classrooms or academic support staff during advising sessions, the policy has limited enforcement power and could reduce the moral of the very actors the policy seeks to incentivize. Theoretically, agents (i.e., faculty, staff) may be tempted to shirk their duties, yet individuals who self-select into faculty and staff positions tend to be more intrinsically motivated (rather than extrinsically motivated by monetary rewards). Faculty and staff may view PF as punitive and unfair, decreasing their motivation to promote institutional goals (Dougherty et al. 2016).

Policy Awareness and Campus Responses in Washington State

In a small-scale study, Li (2017b) interviewed faculty, administrators, and staff at a community college in Washington state to determine knowledge of, viewpoints toward, and responses to PF, complemented with document analysis. She found that the participants in the study were knowledgeable about the PF policy and were able to speak to the policy’s relevance to their individual departments, although support staff were less involved with goals related to the policy. The policy appeared to increase the visibility of student success goals, from developmental education, to retention and graduation, to transitions into the workforce.

Based on Li’s (2017b) findings, the original rollout of the Washington PF policy created some confusion about the metrics. It took a few years before the participants established a better understanding of how the formula calculated student progress. Some participants criticized the formula, that the metrics, since they focused on math and English, did not capture all that students were gaining from attending college. Participants recommended adding a measure for science course completion. Similar to findings about institutional perspectives of PF in other studies, those interviewed in Washington state criticized the policy for not accounting for student characteristics outside of the college’s control nor external job market conditions. Furthermore, participants criticized the data system, which at the time of the study, was cross-sectional and did not track the academic pathways of individual students. Faculty were somewhat skeptical of the policy in general and saw it as a critique of their existing teaching practices.

Nevertheless, Li (2017b) found that positive campus responses to the policy included greater collaboration between faculty and administrators in assessing student learning. For example, an English department explored student learning objectives across courses, especially since English course completions were incentivized within the Washington PF formula. Additionally, the college set up a transitions team to explore ways to improve academic programs and student support services (Li 2017b).
Campus Implementation in Ohio and Pennsylvania

Li and Zumeta (2016) conducted interviews in September 2015 and January 2016 with policymakers and campus officials in Ohio and Pennsylvania, to better understand how PF policies were anticipated to impact student outcomes, whether state policymakers were responsive to institutional feedback, and how campus administrators, faculty, and support staff have responded to PF. The authors interviewed officials from the higher education agencies in each state, university administrators (both academic and non-academic), faculty members, and support staff (Li and Zumeta 2016; Zumeta and Li 2016). Within both state PF policies, there was an explicit emphasis on increasing STEM degree attainment.

Li and Zumeta (2016) found that in Pennsylvania, the financial planning around PF was somewhat curtailed because allocation amounts were perceived as unpredictable. Nevertheless, institutional participants did appreciate the ability to propose metrics aligned with their strategic plans. Strong faculty unions in Pennsylvania limited the flexibility that administrators had in hiring faculty for academic programs based on shifting student interests. Findings also revealed that faculty were encouraged to experiment with high-impact teaching practices, especially in STEM fields (Li and Zumeta 2016).

Li and Zumeta (2016) discovered that institutions in both Ohio and Pennsylvania responded to PF by improving software systems and using data analytics to track student progress and identify areas of intervention. Participants noted that especially for students of color, it was important to trace where students may be falling through the cracks. Participants also discussed peer mentoring and intrusive advising (Li and Zumeta 2016).

While faculty showed varying levels of support for the idea of PF, they were generally supportive and offered no strong critiques (Zumeta and Li 2016). The studies also showed that focusing on the bottom financial line was a necessary reality for all institutions, especially in an era of declining state resources for higher education (Li and Zumeta 2016). Despite the existence of PF, institutions still derived the majority of their revenue from tuition and emphasized the importance of enrollment numbers.

Zumeta and Li (2016) further found that high-resourced institutions in Ohio and Pennsylvania were more able to attract non-state appropriations resources by recruiting more academically prepared students. The same institutions relying less on state revenues were also those selective enough to successfully recruit out-of-state and full tuition payers. PF appeared to exacerbate the gaps between “have” and “have not” institutions. Higher resourced, selective colleges benefited from stronger student applicant pools. The authors expressed concern that PF may be penalizing “poorer-performing” colleges, which also tend to be the ones serving a higher proportion of students from historically marginalized identities. These same colleges have a challenging time improving their performance because they are left with fewer resources to serve a high-need population. The researchers recommended additional resources be provided to institutions with more limited capacity,
especially those with historical access missions or those serving rural populations (Zumeta and Li 2016).

**Campus Responses to Equity Metrics in Ohio and Pennsylvania**

In part to reduce the inclination of colleges to enroll fewer students from marginalized identities, policymakers have incorporated equity metrics, as noted earlier, that provide additional funding for course completions or degree completions by students considered underserved in the state.

Li (2019) explored two states, Ohio and Pennsylvania, that incorporated equity metrics to serve students who were Pell grant recipients and underrepresented minority populations. The author interviewed policy actors at the higher education governing agencies in each state. She interviewed academic and non-academic administrators at five campuses which varied in their urbanicity, enrollment size, and primary mission in terms of teaching, access, or research.

Li (2019) found that in both states, the PF equity metrics played a role in the crafting of new programs for targeted populations and drawing more resources to existing programs. Multiple campuses engaged in the following in direct response to state emphasis on underserved populations: mentorship program for Black male students, greater involvement of faculty in academic advising, and the creation of scholarships for racial minorities or lower-income students. At the same time, participants at one institution did indicate that the funding formula was a factor in instituting higher admissions criteria, to enroll students who had higher entering GPAs and ACT scores. All institutions, and especially the historically access- and teaching-focused institutions, experienced pressure to improve the academic profile of its entering class (Li 2019).

While equity metrics did create positive institutional responses, Li (2019) also found that these metrics did not completely deter the inclination to become more selective. Again, it may be an anticipated response to become more selective, because it is the easiest way to secure better prepared and academically capable students. Findings from qualitative studies should offer caution that PF could cause further stratification among institutions – where more privileged students are concentrated at selective, higher-resourced institutions.

**Policy Evaluation**

Much of the research on PF has focused on whether these policies move the needle on degree outcomes. While PF has sometimes been adopted for political reasons, ultimately, these policies seek the best “bang for your buck” on public investment in higher education, to improve educational attainment. Attainment can be measured by raw numbers of associate and bachelor’s degrees, degrees per 100 FTE student enrollment, or by graduation rates, such as the percentage of student enrollees who graduate within 150% of the typical time to degree completion. Relatedly to fulfill
the ultimate goal of graduation, researchers have also examined first- to second-year retention.

In examining how institutions have responded to PF policies, a segment of the literature on PF policy evaluation has focused on minority-serving institutions. A portion of the research has also explored whether the policies have changed admissions rates, incoming student characteristics, and student demographic characteristics. Connected to the demographic composition of colleges, some studies have also analyzed whether the inclusion of equity metrics in PF policies shows any impacts. A last group of studies have considered institutional finances, and resource allocations that occur with the advent of PF policies, since these policies may have the power to shift institutional financial priorities and tuition setting behavior.

In the next section, I discuss elements of conducting PF policy evaluation that are relevant to all studies, describing methodological strategies and data analyses commonly employed. The literature review that follows is separated into the following sections: 4-year completion outcomes, 2-year completion outcomes, outcomes at minority serving institutions, access and admissions outcomes, equity metrics and related outcomes, and impacts on institutional finances. To offer some organization of the large number of studies, I list study features in Table 1. I display main findings from studies of 4-year completion and retention (Table 2), 2-year completion and retention (Table 3), access and student composition (Table 4), and institutional finances (Table 5). Studies on more than one set of outcomes are listed in multiple tables.

Policy Evaluation Considerations

Isolating the Policy Impact. Perhaps the most difficult challenge in analyzing the impacts of PF policies is isolating such impacts from the influence of other college completion goals. First, higher education administrators, faculty, and staff would tend to agree that student retention and graduation is one of their first and foremost priorities. Thus, the degree outcomes supposedly incentivized by PF should already be a goal of individual actors on a campus, so researchers must try to tease out what additional graduation increase might be due to the PF policy itself.

Second, a slew of other college retention and completion plans may be at play simultaneously. These plans could be at the institution level (e.g., mandatory academic advising before students can register for classes), at the system level (e.g., scholarships for financially needy students eligible for use at any college in the state’s system), at the state level (e.g., transfer articulation agreements to facilitate transfers between 2- and 4-year colleges), or at the national level (e.g., calls for college completion and affordability).

Analytical Designs. Methodologically, researchers have attempted to use rigorous statistical methods, with difference-in-differences (DD) being a highly popular method, to conduct policy impact studies of PF (for more details on DD, see Furquim et al. 2020). Also commonly employed are robustness checks, such as using different control groups and different time periods. These control groups are, naturally,
| Authors | State or national | End year | Begin year | Primary institutional data source | Level of analysis | Control groups | Method | Institution fixed effects | State fixed effects | Institution control variables | State control variables | Year fixed effects | PF treatment variable | PF treatment \( \times \) post-treatment years | \( \times \) duration of policy | \( \times \) STEM metric | \( \times \) duration of policy | County control variables |
|---------|------------------|----------|------------|-----------------------------------|-------------------|----------------|--------|--------------------------|-------------------|---------------------------|---------------------|-----------------|---------------------|---------------------------------|---------------------------------|---------------------------------|---------------------|---------------------------------|---------------------|
| Boland (2020) | National | 2016 | 1998–2006 | IPEDS | 4-year & 10-year | Institution | DD | No | Yes | Yes | No | No | Binary IPEDS; Public HBCUs in non-PF states; Private HBCUs in non-PF states | Binary IPEDS, IPEDS; \( \times \) duration of policy | Yes | No | Yes | No | Yes |
| Favero and Rutherford (2020) | National | 2013 | 1993 | IPEDS | 4-year | Institution | DD | No | Yes | No | Yes | Yes | Binary IPEDS; DD | \( \times \) duration of policy | No | Yes | No | Yes |
| Gándara and Rutherford (2020) | National | 2014 | 2001 | IPEDS, College Scorecard | 4-year | Institution | DD | No | Yes | No | Yes | Yes | \( \times \) duration of policy | No | Yes | Yes | No | Yes |
| Li (2020) | National | 2014 | 2003 | IPEDS | 4-year | Institution | DD | No | Yes | Yes | No | Yes | \( \times \) duration of policy | No | Yes | Yes | No | Yes |
| Hagood (2019) | National | 2013 | 1986 | IPEDS | 4-year | Institution | DD | Yes | Yes | No | Yes | Yes | \( \times \) duration of policy | No | Yes | Yes | No | Yes |
| Hu (2019) | Louisiana | 2016 | 2006 | NCES | 2-year | Institution in non-PF states | DD | No | Yes | Yes | No | Yes | \( \times \) duration of policy | No | Yes | Yes | No | Yes |
| Hu and Villarreal (2019) | Louisiana | 2005 | 2005 | IPEDS | 4-year | Institution in non-PF states | DD | Yes | Yes | Yes | No | Yes | \( \times \) duration of policy | No | Yes | Yes | No | Yes |
| Study                        | Institution | Year | Institution Type | Data Source | Data Duration | Institution Details                                                                 | Methodology | Pre-treatment Controls | Post-treatment Controls | Any Other Details |
|------------------------------|-------------|------|-------------------|-------------|---------------|--------------------------------------------------------------------------------------|-------------|------------------------|------------------------|--------------------|
| Li and Ortagus (2019)        | Tennessee   | 2001 | Institution       | IPEDS       | 2-year        | Institutions in non-PF SREB states; Institutions in non-PF bordering states; Institutions in non-PF states nationally | DD          | No                     | Yes                    | No                 |
| Birdsall (2018)              | Indiana     | 2001 | Institution       | IPEDS       | 4-year        | Institutions in non-PF states nationally; Institutions in non-PF bordering states | DD; Event study | Yes                    | Yes                    | No                 |
| Hillman et al. (2018)        | Ohio, Tennessee | 2005 | Institution       | IPEDS       | 2-year; 4-year | Institutions in non-PF MHEC states (for OH), Institutions in non-PF SREB states (for TN), Institutions in non-PF states nationally; Institutions in PF states nationally | DD; state-specific linear trend | Binary                | Yes                    | Yes                 |
| Hillman and Corral (2018)    | National    | 2005 | Institution       | IPEDS       | 4-year MSIs  | Institutions in non-PF states; MSIs in non-PF states; Non-MSIs in PF states           | DD          | No                     | Yes                    | No                 |
| Kelchen (2018)               | National    | 2004 | Institution       | IPEDS       | 4-year        | Institutions in non-PF states; Institutions in PF states without equity metrics     | DD          | No                     | Yes                    | No                 |
| Li, Gándara et al. (2018)    | Texas, Washington | 2004 | Institution       | IPEDS, State administrative datasets | 2-year MSIs | Non-MSIs in TX and Washington                                                      | Descriptive | NA                     | NA                     | NA                 |
| Li and Kennedy (2018)        | National    | 1990 | Institution       | IPEDS       | 2-year        | Institutions in non-PF states. Robustness checks: Institutions in coordinating/planning board non-PF states; Institutions in never-PF states; Year 1996 onwards; Regular sample excluding TN | DD          | Yes                    | Yes                    | No                 |

(continued)
| Author(s)               | State or national | Begin year | End year | Level of analysis | Primary institutional data source | Sector                  | Control group(s)                                                                 | PF treatment variable | Institutional control variables | State control variables | Institution fixed effects | State fixed effects | Year fixed effects |
|------------------------|-------------------|------------|----------|-------------------|-----------------------------------|-------------------------|--------------------------------------------------------------------------------|----------------------|-------------------------------|------------------------|----------------------|---------------------|---------------------|
| Umbricht et al. (2017) | Indiana           | 2003       | 2012     | Institution       | IPEDS                             | 4-year                  | Institutions non-PF states with similar demographics; Private institutions in Indiana; Public institutions in surrounding non-PF states | DD                   | Binary                        | Yes                    | Yes                  | No                  | Yes                 |
| Kelchen and Stedrak   | National          | 2003       | 2012     | Institution       | IPEDS                             | 2-year; 4-year          | Institutions in non-PF states                                                  | Two-way fixed effects | Binary                        | No                     | Yes                  | Yes                 | Yes                 |
| Hillman et al. (2015)  | Washington        | 2002       | 2012     | Institution       | IPEDS                             | 2-year                  | Institutions in non-PF WICHE states; Institutions in non-PF WICHE bordering states; Propensity score matched institutions in non-PF states nationally | DD                   | Binary; Binary interacted with individual post-treatment years | Yes, plus county controls | No                   | Yes                 | Yes                 |
| Hillman et al. (2014)  | Pennsylvania      | 1990       | 2010     | Institution       | IPEDS                             | 4-year                  | Master’s institutions in non-PF states; neighboring states; northeast region; all nationally; CEM full match; CEM first-year match | DD                   | Binary                        | Yes                    | No                   | Yes                 | Yes                 |
| Author(s)                        | Type    | Time Period 1 | Time Period 2 | Dataset | Institution Type | Institution Sample | Two-way Fixed Effects | Binary and Binary Interacted with Duration of Policy | No. of States | No. of States | No. of States | No. of States |
|---------------------------------|---------|---------------|---------------|---------|------------------|---------------------|----------------------|-----------------------------------------------------|--------------|--------------|--------------|--------------|
| Rutherford and Rabovsky (2014)  | National| 1993 (2003–2010 for retention rates) | 2010 (1993–2005 for 6-yr graduation rates) | IPEDS   | 4-year Institutions in non-PF states | Two-way fixed effects | Binary and binary interacted with duration of policy: Any PF; PF 1.0; PF 2.0 | Yes | Yes | No | Yes | Yes |
| Tandberg and Hillman (2014)     | National| 1990          | 2010          | State   | IPEDS 4-year Non-PF states; Non-PF bordering states; Non-PF states with coordinating/planning boards | DD Binary; Binary interacted with individual post-treatment years | No | Yes | No | Yes | Yes |
| Tandberg et al. (2014)          | National| 1990          | 2010          | State   | IPEDS 2-year Non-PF states; Non-PF bordering states; Non-PF states with coordinating/planning boards | DD Binary; Binary interacted with duration of policy; Binary for each state | Yes | Yes | No | Yes | Yes |

Notes: IPEDS = Integrated Postsecondary Education Data System. NCES = National Center for Education Statistics. HBCU = Historically Black Colleges and Universities. PWI = Predominantly White Institution. MSI = Minority-Serving Institution. SREB = Southern Regional Education Board. MHEC = Midwestern Higher Education Compact. WICHE = Western Interstate Commission for Higher Education. DD = Difference-in-Differences. DDD = Triple Difference. CEM = Coarsened Exact Matching. Comparison groups consist of public institutions, unless otherwise noted.
| Author(s)            | BA degrees                                                                 | Six-year BA graduation rate of FT, FT students | BA degrees per 100 FTE | Retention rate; Percentage of FT, FT students returning for 2nd year | Notes                                                                 |
|----------------------|----------------------------------------------------------------------------|-----------------------------------------------|-----------------------|--------------------------------------------------------------------|------------------------------------------------------------------------|
| Boland (2020)        | Decrease at HBCUs among PF 1.0 policies in 9th year; Decrease among PF 2.0 policies in 2nd year |                                               |                       |                                                                    | Separate analysis for PF 1.0 period (2000–2006) and PF 2.0 period (2007–2014) |
| Favero and Rutherford (2020) | Null for PF 1.0; Null for PF 2.0; Null on variance for PF 1.0; Increase in variance for PF 2.0 | Null for PF 2.0; Null on variance for PF 1.0; Increase in variance for PF 2.0 |                       | Null for PF 1.0; Null on variance for PF 1.0; Decrease in variance for PF 2.0 | Impacts differ across institutions of varying selectivity, and on HBCU versus not. Details not included in this table (see chapter) |
| Li (2020)            | Increase in STEM degrees; Increase in STEM degrees as a proportion of all degrees |                                               |                       |                                                                    |                                                                         |
| Hagood (2019)        |                                               |                                               |                       | Increase at bachelor’s granting institutions; Null at research and highly selective institutions | Additional analyses were conducted on the sample split by state political party control and higher education governance structure (see Hagood 2019) |
| Birdsall (2018)      | Mixed, conditional on institution’s percent of revenue from state funding |                                               |                       | Decrease                                                           |                                                                         |
| Study                          | Results                                                                 |
|-------------------------------|-------------------------------------------------------------------------|
| Hillman et al. (2018)         | Null                                                                    |
| Umbricht et al. (2017)        | Null (3-year average)                                                  |
| Hillman et al. (2014)         | Null in four of five models, increase in average effects for one model; Increase in two of five comparison groups in 4-year lag models |
| Rutherford and Rabovsky (2014)| Null for any PF; Null for PF 1.0; Null for PF 2.0; Null for duration on all three policy scenarios |
| Tandberg and Hillman (2014)   | Null on average; Null in all post-treatment years except increases in 7th, 8th, and 11th years |

Notes: FT, FT = first-time, full-time. FTE = full-time equivalent.
| Author(s) | AA degrees | Graduation rate within 150% time (3 years for AA; varies for certificates) | Short-term certificates (less than 1 year) | Long-term certificates (1 year or more) | Certificates (any length) | Certificates per FTE | Retention rate: Percentage of FT, FT students returning for 2nd year | Notes |
|-----------|------------|--------------------------------------------------------------------------|------------------------------------------|------------------------------------------|----------------------------|---------------------|---------------------------------------------------------------|-------|
| Hu (2019) | Null       | Null or increase depending on model (graduation rate includes AA and certificate seekers) | Increase | Increase | Increase | Null on part-time retention rates; Null on full-time retention rates | Additional analyses were conducted between minority-serving versus non-minority-serving institutions (per the author’s definition) and between low versus high-income serving institutions (see chapter) |
| Li and Ortagus (2019) | Null | Increase | Increase | | | | | Three additional control groups were used to analyze student subpopulations of institutions under PF policies: Without any equity metrics; with equity metrics except for adult students; with equity metrics except for low-income students (see Li and Ortagus 2019) |
| Author(s) | Description (2018) | Description (2015) | Other Description (2014) |
|-----------|------------------|------------------|------------------|
| Hillman et al. (2018) | Decrease compared to PF states nationally; Null compared to non-PF states within region; Null compared to non-PF states nationally | | Increase in Ohio compared to region, null otherwise; Increase in Tennessee for all comparison groups |
| Li and Kennedy (2018) | Null on average; Decrease for stronger policy types | Null on average; Increase for stronger policy types | Generally null across policy types |
| Hillman et al. (2015) | Generally null, increase in later years for AA degrees per 100 FTE | Null on total certificates awarded; Increase in per 100 FTE certificates | Decrease in total certificates; Decrease in per 100 FTE certificates |
| Tandberg et al. (2014) | Null | | Parameters were generated for each treated PF state in the sample (see chapter for more details, and see Tandberg et al. 2014) |

Notes: FT, FT = first-time, full-time. FTE = full-time equivalent.
### Table 4  Studies on access and student composition

| Author(s) | Access/admissions | Student composition |
|-----------|-------------------|---------------------|
|          | Acceptance rate   | 25th percentile     | 75th percentile    | Percent of minoritized students (Hispanic, Black) | Number of minoritized students (Hispanic, Black) | Percent on Pell Grant (ever during UG career) | Number of Pell Grant students | Percent first generation | Number of students by age | Notes |
| Gándara and Rutherford (2020) | Decrease on average; Decrease on low selectivity; decrease on high selectivity; Null for 1.0, decrease for 2.0 | Increase on average; Increase for low selectivity, null for high selectivity; Null for 1.0, increase for 2.0 | Null on average; Increase for low selectivity, null for high selectivity; Null for 1.0, increase for 2.0 | Decrease in Black students on average; Null for low selectivity, null for high selectivity; Decrease for 1.0, null for 2.0; Null on all models for Hispanic students | Null on average; Null for low selectivity, null for high selectivity; Null for 1.0, increase for 2.0 | Decrease on average; Null for low selectivity, increase for high selectivity; Decrease for 1.0, Decrease for 2.0 | | | | Sample was split into PF 1.0 and PF 2.0, and low versus high selectivity. Online tables show differences between PF policies with equity metrics and those without (see Gándara and Rutherford 2020) |
| Li and Ortagus (2019) | | | | Increase in number and in percentage enrolled (of FT, FT degree-seeking students receiving federal grant aid) | Decrease in number and in percentage of adult students (age 25 and older) | Decrease in number and in percentage of adult students (age 25 and older) | | | | Three additional control groups were used to analyze student subpopulations of institutions under PF policies: Without any equity metrics; with equity metrics except for adult students; with equity metrics except for low-income students (see Li and Ortagus 2019) |
| Birdsall (2018) | Decrease | Decrease using main comparison group; null in MHEC comparison group | | | |
|---|---|---|---|---|---|
| Kelchen (2018) | | Generally across models, increase in African-American enrollment among PF policies; Null for Hispanic students with any equity metric | Generally, decrease under any PF policy, and decrease under PF systems with equity metrics as well; Decreases observed at more selective and less selective colleges | Decrease in younger students (age 24 or younger) at less selective colleges under PF equity metrics compared to PF non-equity metrics | Directions noted in this table are general across multiple models; for more details on comparison groups, see Kelchen (2018) |
| Umbricht et al. (2017) | Decrease | Increase | Increase | Decrease after 2-year lag (although minority race categories not defined) | Increase in federal grant aid recipients among entering class (compared to private Indiana colleges); Null for other models |
| Author(s)               | State appropriations per student (or per FTE student)                                                                 | In-district tuition and fees for public 2-years | In-state tuition and fees for public 4-years | Out-of-state tuition and fees for public 4-years | Revenues in specific categories per FTE student | Notes                                                                                                                                                                                                 |
|------------------------|-------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|---------------------------------------------|------------------------------------------------|-----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Hagood (2019)          | Increase at research institutions, highly selective institutions; Decrease at master’s, bachelor’s, less selective, non-selective, and rural institutions |                                               |                                             |                                                 |                                               | Additional analyses were conducted on the sample split by state political party control and higher education governance structure (see Hagood 2019)                                                        |
| Hu and Villarreal (2019)|                                                                                                                        | Increase                                      | Increase                                    | Null                                           |                                               |                                                                                                                                                                                                     |
| Hillman and Corral (2018)| Null at MSIs in PF states compared to all institutions in non-PF states; Decline compared to MSIs in non-PF states; Null compared to non-MSIs in PF states |                                               |                                             |                                                 |                                               |                                                                                                                                                                                                     |
| Li, Gándara et al. (2018)| MSIs in Texas and Washington did not see disproportional declines in funding compared to non-MSIs in the respective states |                                               |                                             |                                                 |                                               | Performance on specific formula indicators was compared between MSIs and non-MSIs in the two states, respectively                                                                             |
| Kelchen and Stedrak (2016)|                                                                                                                        |                                               |                                             |                                                 | Decrease in Pell Grant revenue at 2-year and 4-year colleges; Increase in unfunded grants at 4-years; Decrease in auxiliary enterprise funding at 2-years | A large number of outcomes were examined. Null results not included in this table for the sake of brevity (see Kelchen and Stedrak 2016)                                      |
institutions, usually in other states that did not experience PF during the time period of study. These control groups often consist of institutions in neighboring states, states with similar higher education governance structures, states within the same higher education regional compact (e.g., Southern Regional Education Board), or institutions selected using matching techniques. The strategy is to find comparable institutions not subject to PF during the same time period, to serve as a counterfactual to model outcomes that would have been expected in the absence of PF.

Moreover, all studies use a set of control variables, most commonly at the college-level (e.g., student enrollment characteristics) or the state-level (e.g., revenue from state appropriations per student) and occasionally at the county-level (e.g., unemployment rate). These variables are intended to capture features within a college, system, or state that may affect the outcomes.

Further, researchers employ unit fixed effects (at either the college- or state-level) to account for unobservable features that are relatively constant across time (e.g., institutional prestige). Similarly, researchers use time fixed effects to account for trends within each year that affect all institutions in the dataset (e.g., a recession), to better control for factors impacting outcomes that are separate from a PF policy treatment.

**Policy Treatment Variable.** A summary of key features of studies are noted in Table 1. As seen, in evaluating PF as a treatment, researchers all have used a binary/dichotomous variable—designating an institution as being subject to PF in a specific year, or not. This coding format produces an estimate that captures the average treatment effect of PF on treated colleges, shown as an average policy impact across all posttreatment years.

Second, in a few studies (e.g., Li and Ortagus 2019; Rutherford and Rabovsky 2014), researchers also interacted the binary PF treatment variable with duration of time of the policy (number of operating years), designating the treatment variable as "1" during the first year of the policy, and as "2" during the second year of the policy, and so forth. This coding format generates one parameter capturing the policy impact on a single year posttreatment, which assumes a linear trend of the policy—the outcome changes by the amount of this parameter during each year posttreatment.

Third, in select studies, researchers (e.g., Hillman et al. 2018, 2015) have occasionally interacted the binary PF treatment variable with individual post-treatment years. This means that each year after the treatment produces its own parameter estimate, showing the policy impact during each year after the policy. This setup allows the treatment effect to vary during each year posttreatment. Although, caution is advised for interpreting such estimates, because oftentimes, policies have not been in place for more than 5 years and estimates for years that are further out are based on a smaller sample size of institutions.

Researchers (e.g., Kelchen 2018; Li and Ortagus 2019) have also employed lags between the PF treatment year and year in which the outcome gets measured, assuming that the impact of PF on an outcome is not seen until the next academic year. I do not discuss lags in detail in the following literature review because using 1- to 3-year lags are quite standard in policy evaluation studies.
Four-Year College Completion Outcomes

Earliest Studies. Earlier studies published in the mid to late 2000s decade found that PF policies were not associated with improvements on retention rates, degrees awarded, or graduation rates at public 4-year colleges (Sanford and Hunter 2011; Shin 2010; Shin and Milton 2004; Volkwein and Tandberg 2008). These studies have been discussed in earlier comprehensive literature summaries of PF, so I do not go into details here (see Dougherty and Reddy 2013, Table A2). The following studies mentioned are more recent (Table 2).

Studies on 4-Year Outcomes Published in the Mid-2010s. One critique of earliest versions of PF was that the policies had not been in place for many years. Thus, public universities had not had enough time to respond to policies. Therefore, more recent studies capture a longer time frame that, in theory, gives adequate time for institutions to make changes that influence college completions.

Using DD techniques on a national dataset from 1990 to 2010, Tandberg and Hillman (2014) found that institutions under PF did not show any significant change on the total number of bachelor’s degrees conferred, on average. After generating estimates for each year that the policy was in place, the authors found significant increases in bachelor’s degrees in the 7th, 8th, and 11th operating year of PF. However, the authors cautioned that a number of states studied under their time period did not continuously operate PF for 7 or more years (Tandberg and Hillman 2014).

Using similar methods and the same time frame of 1990 to 2010, Hillman et al. (2014) examined all institutions within the Pennsylvania State System of Higher Education, which at the time of the study, were exposed to one of the longest and most stable operating PF policies (Cavanaugh and Garland 2012). The authors examined the first iteration of Pennsylvania’s policy on bachelor’s degrees per 100 FTE enrollment using three control groups of geographically nearby institutions, and two control groups of institutions determined using Coarsened Exact Matching (CEM), which consisted of using covariates to create groups with similar features as the treated group.

Hillman et al. (2014) found that when estimating an average treatment effect, one of the five control groups yielded a positive effect on degrees per 100 FTE, while the remaining four showed null effects. The models that used a 4-year lag demonstrated an increase in degrees per 100 FTE among two of the three geographical nearby control groups. However, null impacts were found when using the two control groups generated by CEM matches – which the authors identified as more rigorously selected and exhibiting more similar pre-treatment trends. The authors concluded that Pennsylvania’s PF policy did not produce changes in degree outcomes different than what would have occurred in the absence of the policy (Hillman et al. 2014).

Rutherford and Rabovsky (2014) used a national sample to analyze three outcomes: bachelor’s degrees per 100 FTE students in years 1993 to 2010, 6-year bachelor’s graduation rates from years 1993 to 2005, and retention rates from 2003 to 2010. The smaller time period for retention rates was due to data availability from the Integrated Postsecondary Education Data System (IPEDS), and the smaller time
period for graduation rates was due to rates needing 6 years for reporting. The authors employed two-way fixed effects (fixed effects for each state and year) to analyze any PF policy, PF 1.0, PF 2.0, and the duration of these three policy scenarios. The authors described estimates that were significant in the absence of state fixed effects, which I note here, but I contend that the full models with both state and year fixed effects were more precise.

Rutherford and Rabovsky (2014) found null impacts on average graduation rates for all of the treatment conditions. In fact, they found a negative impact on the duration of PF 1.0 on graduation rates, suggesting that institutions under PF 1.0 may have seen declines across time, although this estimate was significant in the model without state fixed effects, but insignificant after state fixed effects were added (Rutherford and Rabovsky 2014).

On retention rates, Rutherford and Rabovsky (2014) found that no estimates were significant for any PF policy, PF 1.0, nor PF 2.0, nor for the duration of these three policy scenarios. Similarly, for bachelor’s degrees per 100 FTE, none of the policy scenarios were associated with changes in the outcome. The authors described negative effects of the policy on degree production, but these models were absent of state fixed effects and not as robust, so the overall impacts are basically null (Rutherford and Rabovsky 2014).

**Most Recent Studies on 4-Year Outcomes.** Li (2020) conducted a study that differed from other studies of PF on 4-year degree outcomes by examining STEM degrees specifically. Using a dataset from 2003 to 2014, the author analyzed STEM degree completions at institutions in states with a special STEM incentive embedded within their PF policies (and when relevant, PF incentives for health and related clinical science). These policies funded STEM completions as a separate performance metric or gave higher weights to the conferral of STEM degrees, versus degrees in non-STEM fields. STEM degrees were considered to be high-need fields, since salaries as well as the likelihood of being employed are usually higher compared to non-STEM disciplines. With the exception of math, STEM courses usually have a higher instructional cost (due to labs and higher faculty salaries), which is more reason to provide added funding in PF formulas for the completion of such degrees.

Li (2020) found that compared to institutions in states without PF STEM incentives (including those not under PF and those under general PF absent of STEM incentives), institutions under PF STEM incentives saw increases in total STEM bachelor’s degrees. This positive impact was found for the average treatment effect, as well as the operating years (duration of years) treatment variable, which assumes a linear increase in degrees during each posttreatment year. Compared to institutions under PF absent of STEM incentives, the treated group saw increases in STEM degrees. Furthermore, compared to institutions not under any PF system (general nor STEM), STEM bachelor’s degrees increased at treated institutions.

In addition, STEM bachelor’s degrees as a proportion of all bachelor’s degrees increased significantly at institutions under state PF STEM incentives, on average, and as an operating years impact, compared to each of the three control groups noted above (Li 2020). These results were robust to alternative specifications and control
groups. Interestingly, the number of non-STEM degrees did not decrease, suggesting that the increase in STEM degrees did not come at the expense of other fields.

The biggest challenge in Li’s (2020) study was differentiating STEM degree completions due to PF incentives versus institutional and statewide initiatives to promote STEM. Li’s (2020) is the only study (of the ones reviewed) that consistently showed increases in bachelor’s degrees in response to PF, although the study was specific to STEM degrees. The study underscores the need to perhaps consider not just total numbers of degrees across all fields, but to consider the discipline in which degrees were awarded.

Favero and Rutherford (2020) applied the idea of stacking the deck, when policymakers use policy structures (e.g., PF) to steer agency decisions, which winds up favoring some stakeholders over others. Policies become constructed in a way that benefits the organizations that are most prominent during the policymaking process or most visible to the general public. The authors applied a second conceptual viewpoint better known amongst PF researchers—stratification—in which the most advantaged students self-select into institutions that are already better resourced, feeding into the ability for such institutions to excel on performance metrics.

Favero and Rutherford (2020) put forward a formal model to predict how institutions responded to PF, using a set of theoretical assumptions, empirical models on how performance at institutions in each year were determined, and Monte Carlo simulations. In their simulations, the authors assessed low-end, median, and top-end institutional performers, and altered two parameters: the extent to which previous year’s performance affected current year’s performance based on the proportion of state revenues received and the extent to which institutions positively altered behavior to respond to PF incentives.

Favero and Rutherford (2020) explained that in theory, PF could have two expected outcomes on different types of institutions. First, the distance in performance between the best and the worst performers may widen. Second, the average performance across all institutions may increase. Both of these possibilities would likely improve the performance of already high-performing institutions.

In analyses using DD on data from 1993 to 2013, Favero and Rutherford (2020) showed that PF 1.0 produced no changes to bachelor’s degrees per 100 FTE, but that PF 2.0 policies produced increases. Overall, neither PF 1.0 nor PF 2.0 produced changes to 6-year bachelor’s graduation rates of first-time, full-time students, nor any impacts on first- to second-year retention rates.

When examining the effect of PF policies on variance across low-performing, median, and high-performing institutions, PF 1.0 was not associated with any changes (in standard deviation on the three outcomes, see Favero and Rutherford 2020 for more details). Yet, PF 2.0 increased the variance on graduation rates across these three levels of institutional performance but decreased the variance on degrees per 100 FTE. Null impacts were observed for retention rates.

Favero and Rutherford (2020) also analyzed effects of PF moderated by Barron’s selectivity rating. PF 1.0 policies appeared to positively impact degrees per 100 FTE at non-selective schools, while showing null to negative impacts on more selective
schools. Their study showed that PF 2.0 was not associated with any changes in degrees per 100 FTE across all selectivity categories. For graduation rates, PF 1.0 had positive effects on the least selective schools, yet negative effects on the most selective schools. PF 2.0 positively impacted the graduation rate for more selective schools. For retention rates, more selective (but not less selective) institutions saw a marginal improvement in the presence of both 1.0 policies and 2.0 policies (Favero and Rutherford 2020).

Finally, Favero and Rutherford (2020) compared PF effects on HBCUs versus non-HBCUs, under the assumption that HBCUs may be considered lower performing due to serving more marginalized students. Their study demonstrated that PF 1.0 had no impact on degrees nor graduation rates for non-HBCUs. However, PF 1.0 positively impacted degrees at HBCUs. On the other hand, PF 2.0 positively impacted degrees for non-HBCUs, although no impacts on degrees were seen by HBCUs. PF 2.0 negatively impacted graduation rates for HBCUs, but not for non-HBCUs. Lastly, neither PF 1.0 nor PF 2.0 had any impacts on retention rates, whether at HBCUs or non-HBCUs.

Favero and Rutherford (2020) concluded that the evidence of PF policies on high-versus low-performing institutions was mixed, and not all results aligned with expectations. Specifically, their study found that PF 1.0 did not generally exacerbate gaps between high and low institutional performers. However, PF 2.0 policies had fewer positive effects on historically lower-performing institutions, compared to higher-performing counterparts.

Favero and Rutherford (2020) presented novel ideas on PF’s potential to exacerbate existing stratification amongst institutions within states. The analysis on variance using standard deviations could be replicated by other researchers looking at different PF time periods. Extending the HBCU comparisons to other minority-serving institutions could also prove an interesting way to determine differential effects of PF on differently resourced institutions.

On a practical matter, to make performance allocations more equitably distributed across institutions, the authors suggested making performance funds based on changes in year-to-year performance rather than absolute levels (Favero and Rutherford 2020). This would better recognize a lower-performing institution’s ability to improve upon prior year performance, rather than having to compete with other institutions who are already higher performers on raw metrics such as graduation rates.

Two-Year College Completion Outcomes

Studies on 2-Year Outcomes Published in the Mid-2010s. Table 3 summarizes findings from studies of 2-year outcomes. Tandberg et al. (2014) conducted one of the earliest national studies of PF impacts on associate degrees conferred at 2-year colleges, although the number of associate degrees was aggregated at the state level, and not the institution level. The authors examined years 1990 to 2010, and incorporated a binary PF treatment variable, along with a duration policy variable
indicating the number of active years of a PF policy. The authors incorporated three
control groups of non-PF states, non-PF bordering states, and non-PF states with
coordinating or planning board states, similar to the Tandberg and Hillman (2014)
study on 4-year colleges.

Tandberg et al. (2014) show results from DD, which suggested that PF policies
were not associated with any changes in associate degree completions, on average,
when compared to any of the three control groups. The authors also generated
parameters for each of the 19 treated states in their sample. Within individual states
after PF, four states experienced increases in associate degrees (MN, MO, NJ, WA),
six states experienced declines (CO, ID, NM, SC, TX, VA), three states showed
mixed or inconsistent estimates across models (AR, KS, OK), and the six remaining
states showed null impacts (FL, IL, IN, KY, NC, OH) (Tandberg et al. 2014).

Tandberg et al.’s (2014) study underscored the importance of analyzing differ-
ences in outcomes across states, since average treatment effects masked heteroge-
neity. The authors surmised that the lack of consistent and positive effects on
associate degree completions may be due to limited institutional capacity (financial
resources, full-time faculty) to respond to external accountability demands.

Hillman et al. (2015) analyzed Washington state’s community and technical
colleges, which were under a PF policy that awarded points for a number of
outcomes that included first-year retention, and the completion of a degree, occupa-
tional certificate, or apprenticeship. The authors explored years 2002 to 2012 and
used DD techniques. They found no differences in part-time retention rates nor full-
time retention rates, on average, for Washington colleges compared to each of the
three control groups not subject to PF: community and technical colleges from states
that were members of the Western Interstate Commission for Higher Education
(WICHE, a regional compact); colleges in WICHE that shared a border with
Washington state; and a sample of colleges nationally with similar pre-treatment
retention and completion outcomes generated using propensity score matching.
When interacted with the number of treated years, increases were observed in
some years but decreases observed in other years, for both part-time and for full-
time retention rates, suggesting no clear pattern (Hillman et al. 2015).

For the total number of short-term certificates awarded (credentials requiring less
than 1 year of study), Hillman et al. (2015) found no average effects after PF. For
short-term certificates per 100 FTE, the authors’ study showed increases on average
and increases for individual year estimates. By contrast, total numbers of long-term
certificates (requiring more than 1 year of study) showed consistent declines on
average, and during individual posttreatment years. Parameter estimates for long-
term certificates per FTE also revealed some evidence of systematic declines. Lastly,
Hillman et al. (2015) found that associate degrees awarded, on average, did not
change after PF in Washington, although, during the 5th and 6th year of the policy,
there were observable increases in degrees per 100 FTE in select models.

In summary, Hillman et al. (2015) concluded that colleges may be influenced by
Washington’s PF policy to graduate more students in short-term certificates. Inter-
estingly, retention rates did not appear to increase, which warrants further study to
determine if PF policies have influence on this pivotal time period in a student’s
academic trajectory. Associate degrees did increase in later years after Washington’s policy, offering some encouraging evidence of possible longer-term PF impacts. Washington did later revise its funding formula to differentiate between short- and long-term certificates, and part of the initial increase in certificates could have been attributed to better data tracking.

**More Recent Studies on 2-Year Outcomes.** Li and Kennedy (2018) conducted a national study examining PF impacts at 2-year colleges from 1990 to 2013 and employed a DD research design. The authors analyzed short-term certificate completions (less than 1 year of study), medium-term certificates (1–2 years of study), and associate degrees. In reviewing literature to emphasize the importance of analyzing short-term versus longer-term credentials, the authors noted that graduates of short-term certificates rarely experienced improvements in wages, on average. Long-term certificates were generally more likely to create wage gains for women compared to men. By contrast, associate degree completions produced consistent wage gains, on average, although variation existed depending on the field of study.

A unique contribution of Li and Kennedy’s (2018) study was the use of a PF policy typology ranging from Type I to Type IV. Type I policies were defined as lower stakes, with only base funding tied to outcomes, lacking in mission differentiation, with less than 5% of state revenues tied to outcomes. Type IV was the most advanced, with 25% or more tied to outcomes, both 2- and 4-year sectors affected, mission differentiation, and importantly, sustained for two or more consecutive years (Snyder and Fox 2016). While imperfect, the typology system was used to distinguish between rudimentary versus advanced versions of PF across states. For more details on typology, see Li and Kennedy (2018); Li, Kennedy et al. (2018); and Snyder and Fox (2016).

Li and Kennedy (2018) found that on average, colleges under PF did not observe any changes in total counts of short-term certificates, medium-term certificates, nor associate degrees, when compared to 2-year colleges nationally that were not subject to PF. When individual posttreatment years were incorporated, associate degrees showed declines in the 2nd and the 4th operating year of PF. No changes in short- nor medium-term certificates were observed up to the 5th operating year. When differentiated by policy typology, Type III and Type IV policies produced, on average, an increase in short-term certificates. No change was observed for medium-term certificates. On associate degree completions, declines were associated with Type IV policies, but not with Type I through III policies.

The estimates for each operating year of PF tended to support the average effects (Li and Kennedy 2018). That is, the stronger policy typologies (III and IV) produced increases in short-term certificates in years following PF. Medium-term certificate changes were mixed and told an inconclusive story, though the authors surmised that medium-term certificates declined in later years, an effect that was accompanied by growth in short-term certificates. Associate degree completions declined in a couple select years among Type I and Type II policies, declined in three of the five operating years examined for Type III policies, and consistently decreased in all five operating years among Type IV policies (Li and Kennedy 2018).
Using alternative control groups, Li and Kennedy (2018) conducted robustness checks and analyzed average policy impacts on the three credential outcomes. Results of the robustness checks showed null average policy impacts compared to each of the following alternative control groups: non-PF states with coordinating or planning boards; states that never adopted PF; the same sample using year 1996 onward to focus on PF 2.0 policies; and excluding Tennessee (which started PF before the study time period). In summary, the authors cautioned that PF policies, especially those that tied a higher proportion of state funding to outcomes, caused colleges to focus on short-term certificates and possibly displaced focus from associate degree completions. One shortcoming of Li and Kennedy’s (2018) study was that a number of states began or re-started their PF systems in 2011, 2012, or 2013, and thus the years analyzed may not have accurately portrayed how colleges would respond multiple years after PF.

Li and Ortagus (2019) analyzed a major change in Tennessee’s PF policy, which increased the proportion of state funding to higher education from 5.45% to 85% and added a 40% funding premium for retention and completion by students considered adult or low-income. The authors used DD to compare outcomes at 2-year colleges in Tennessee to three control groups of institutions not under PF. For the enrollment of adult and low-income students, the authors used additional control groups of institutions under PF but without the relevant premium.

Estimates suggested that the Tennessee PF policy change decreased the number of adult students, measured as the total fall enrollment of students aged 25 and older. This finding was consistent compared to control groups of institutions in non-PF states that were members of the Southern Regional Education Board (SREB), non-PF states that border Tennessee, and in the national sample of non-PF states. The number of adult students also declined compared to institutions under PF policies that excluded premiums. Moreover, when analyzed as the percent of adult students enrolled, this outcome was shown to significantly decrease after the policy change (Li and Ortagus 2019).

The Tennessee policy change increased the number of low-income students, represented by first-time, full-time degree-seeking students receiving federal grant aid. This increase was observed when compared to all control groups explored, except for the sample of all non-PF states nationally. In addition, the percent of low-income students in Tennessee 2-year colleges increased significantly, a finding robust to all control groups utilized. Parameter estimates on the student subpopulation outcomes were consistent using a 1-year lag of the policy and for estimates using the duration of the policy.

Furthermore, Li and Ortagus (2019) found that the Tennessee PF policy revision produced no changes to the number of associate degrees awarded. On the other hand, there was an increase observed for short-term certificates requiring less than 1 year of study, as well as an increase in medium-term certificates requiring 1 year or more to complete. The findings on certificate completions were significant in the lagged model and when using the duration version of the policy variable.

The results of the Li and Ortagus (2019) study was puzzling in the disparate findings regarding adult student enrollment versus low-income student enrollment,
even though both populations were offered 40% premiums in the revised Tennessee PF model. It is promising that the funding formula perhaps incentivized community colleges to enroll additional low-income students. The decline in adult students was perplexing, and qualitative studies are needed to explore college recruitment and outreach strategies. It is possible that returning adult students responded differently to marketing techniques compared to their younger peers.

Hillman, Fryar, and Crespin-Trujillo (2018) analyzed both Tennessee and Ohio, two states with high proportions of funding tied to outcomes. Using DD on data from 2005 to 2014, the authors found that Ohio 2-year colleges experienced a significant increase in the number of certificates awarded, compared to institutions located in non-PF states that were members of the Midwestern Higher Education Commission (MHEC). However, when compared to institutions nationally not under PF, and with institutions nationally under PF, the number of certificate completions in Ohio did not change. In Tennessee, certificate completions were significantly higher at 2-year colleges after PF, compared to each of the three control groups (Southern Regional Education Board (SREB) states, non-PF states nationally, PF states nationally). Increases were also observed in all post-policy treatment years.

For associate degree completions among 2-year colleges in Ohio, an average decrease was observed after PF, along with annual decreases, when compared to the control group of PF states nationally (Hillman et al. 2018). Similarly, in Tennessee, associate degree completions declined when compared to other PF states nationally and in each treatment year. The use of regional and non-PF national control groups showed no significant changes to associate degrees at treated colleges.

Furthermore, Hillman et al. (2018) found that on average, bachelor’s degree completions did not change at 4-year colleges after Ohio started PF, for all three control groups. Similarly, in Tennessee, bachelor’s degree completions were no different after PF, compared to non-PF Southern Regional Education Board states, non-PF states nationally, nor PF states nationally.

Hillman et al. (2018) concluded that as a whole, PF policies in Ohio and Tennessee produced increases in the lower-order goal of certificate completions. However, PF policies were ineffective at increasing higher-order goals, specifically the number of associate degrees and bachelor’s degrees. The decline in associate degrees in select models could signify that colleges prioritized certificate completions over associate degree completions.

Hu (2019) examined PF in Louisiana and divided her dataset using the proportion of racially minoritized students at each institution. She designated each institution as a minority-serving institution (MSI) using a 25% enrollment cutoff of that particular group and a 50% enrollment threshold for all minorities combined. The author defined minorities as students who were Black/African American, American Indian/Alaska Native, Asian/Pacific Islander, and Hispanic. Further, Hu (2019) designated institutions that enrolled at least 50% of students who received federal financial aid as low-income-student-serving institutions. By creating this additional difference between institutions, she utilized a triple-difference strategy, where the first difference was pre- and post-PF, the second difference was between treated and non-treated colleges, and the third difference was between treated colleges that were
high versus low on enrolling minoritized students or high versus low on enrolling lower-income students, respectively.

Hu’s (2019) DD estimates suggested that due to PF, Louisiana 2-year colleges awarded more certificates compared to their neighboring and Southern Regional Education Board (SREB) counterparts. Similarly, Louisiana colleges awarded more certificates per FTE enrollment compared to surrounding states and to SREB states, although this positive effect wore off after 2 years post-PF implementation.

Hu (2019) also analyzed associate degrees and found no changes to the number of degrees awarded after Louisiana implemented PF. Moreover, on graduation rates (which was pooled across associate degree seekers and certificate seekers), the models showed a significant increase in the same-year and 1-year lag models, although not for the 2-year lag models. Lastly, in all but one model, the PF policy did not produce any changes in full-time nor part-time retention rates (Hu 2019).

Hu (2019) examined MSIs (as per the author’s definition) in Louisiana versus non-MSIs in non-PF states and found no significant differences for the number of certificates awarded, associate degrees awarded, graduation rates, part-time retention rates, nor full-time retention rates (in the majority of models). Generally, across most models in the study, Louisiana low-income-student-serving institutions did not exhibit differing outcomes compared to non-low-income-student-serving institutions in control group states. The exception was that low-income-student-serving institutions in Louisiana appeared to confer fewer associate degrees. The author cautioned that PF policy designers may wish to add metrics for low-income students as to not disincentivize colleges from enrolling such students (Hu 2019).

To summarize her findings, Hu (2019) concluded that Louisiana’s 2-year colleges increased certificate production but did not increase associate degree production nor retention rates, in response to PF. The increase in graduation rates detected was most likely attributable to certificate programs, rather than associate degree programs.

Increases in Short-Term Certificates Consistent Across Studies. Across the several studies reviewed above, a common thread is that at 2-year colleges, certificate completions appeared to have increased with the advent of PF (Hillman et al. 2018, 2015; Hu 2019; Li and Kennedy 2018; Li and Ortagus 2019). While this is not necessarily a negative outcome, it is worrisome in the context of findings across studies that associate degrees did not seem to increase. The studies suggest that perhaps colleges are resorting to easier to accomplish alternative goals – increasing certificate production, since these programs are faster for students to complete. This finding consistently showed up for states such as Tennessee with higher proportions of funding tied to outcomes (Hillman et al. 2018; Li and Ortagus 2019). However, associate degree completions provide greater wage benefits for students.

Thus, PF policy designers may wish to consider weighting certificates to a smaller degree and providing greater relative financial awards to associate degree completions. It is also possible that PF simply does not provide the capacity needed for open access colleges to steer their students through associate degree programs, more so than colleges are already seeking to do in the absence of PF.
**Meta-Analysis on Outcomes and Access**

Bell et al. (2018) conducted a meta-analysis of 11 studies of PF and separated the outcomes of interest into completion (degrees conferred, graduation rates, or retention rates), and access (percentage of low-income students, Pell grant recipients, Hispanic/LatinX or African American students). Using a random effects model, the authors found that without controlling for 2- and 4-year sector differences, the effect of PF on degree completion was indistinguishable from zero. When separated into 2- and 4-year sectors, the effect was still null for both sectors (Bell et al. 2018).

Bell et al. (2018) additionally analyzed access measures across studies within the 11 that included an outcome variable on the admission of students of color and/or low-income students. They found that PF policies showed no impact on access. While the average effect on access for 4-year colleges trended negative (suggesting enrollment of fewer lower-income and Hispanic/Latinx and African American students), the authors concluded that estimates were essentially null.

In looking across studies to determine which control variables were associated with PF, Bell et al. (2018) did recommend that researchers ought to control for the following characteristics: total levels (or per student levels) of state appropriations for higher education (as anticipated, since an institution’s dependence on state revenues could affect responses to PF, and because resources affect completion), the state unemployment rate, tuition, and importantly, the number of years that PF has been in place. In contrast, studies that controlled for political variables including citizen ideology and the centralization of the higher education governing board did not show strong associations with whether PF policy impact estimates were significant (Bell et al. 2018).

**Outcomes Specific to Minority Serving Institutions**

Several publications have examined how PF may affect minority-serving institutions in particular. These institutions have a historic mission to serve students of color, a population that is also more likely to be first-generation or lower-income. Qualitative evidence suggests that institutional actors at HBCUs tend to be skeptical of PF given historical inequities in state funding for HBCUs compared to their predominantly white colleges (Jones 2016). Administrators interviewed by Jones (2016) voiced concerns about their institutions being held to unrealistic standards, since HBCUs enroll more low-income students and students of color.

Jones et al. (2017) wrote that PF policies perpetuated existing funding inequities between MSIs and their predominantly White institution counterparts. Specifically, MSI institutions were more likely to experience declines in state revenues after PF. In addition, institutions with the highest per-student funding from state revenues also tended to have the largest endowments and non-state resources, which further illustrates the advantage that higher resourced institutions have. Another challenge faced by MSIs was lacking the authorization to offer degrees in higher-demand fields, which exacerbated existing inequities in funding provision, and limited the
ability of MSIs to boost performance according to PF funding formulas. The authors pointed to the following issues: historic underfunding of MSIs, discrimination against people (and institutions of color), and the reality that PF policies were not designed to account for the special, historic mission of MSIs. In other words, MSIs tended to fare worse financially under PF regimes (Jones et al. 2017), and the inclusion of equity metrics is unlikely to completely counteract such inequities.

Hillman and Corral (2018) examined MSIs (all types) across states with PF, using standard DD models as well as models with state-specific linear trends. Compared to all institutions in non-PF states, MSI institutions in PF states saw no changes in average state appropriations per FTE. However, MSIs in PF states did experience significantly lower appropriations per FTE compared to MSIs in non-PF states. Nevertheless, these treated colleges did not experience changes in state appropriations per FTE after PF compared to non-MSIs in the same PF states. The authors concluded that on average, PF did not lead to funding level declines in institutions overall but did reduce funding for MSIs specifically.

With respect to models using state-specific estimates, Hillman and Corral (2018) found that MSIs lost significant appropriations per FTE in some states, gained in other states, and showed no difference in the remaining states. Interestingly, MSIs located in two states with higher-stakes PF, Tennessee and Ohio, disproportionately lost funding per FTE after PF, compared to MSIs in non-PF states, and compared to non-MSIs in the same state (Hillman and Corral 2018). The authors concluded that PF regimes would do well to equalize financial resources for MSIs, since MSIs serve a disproportionately greater number of students of color and low-income students. Future research, notes the authors, could examine differences in resources provided to MSIs in PF states with equity metrics versus states that do not include equity metrics.

Boland (2020) analyzed PF impacts on bachelor’s degrees for his treatment group of public, 4-year HBCUs in states with PF. The author utilized five different control groups: public HBCUs in non-PF states; PWIs in PF states; PWIs in non-PF states; private 4-year HBCUs in PF states; and private 4-year HBCUs in non-PF states (which ultimately did not meet the parallel trends assumption for the outcome analyzed). The author used a binary indicator for PF treatment that interacted with the number of years of each state’s PF policy. The author sliced up his analysis to approximate the PF 1.0 period (1998 to 2006) and the PF 2.0 period (2007 to 2014).

For the years 1998 to 2006, among the four control groups, DD estimates showed a negative impact of PF on HBCU bachelor’s degree completions in the 9th year of PF policies, with years prior trending negative although mostly statistically insignificant (Boland 2020). For the years 2007 to 2014, estimates showed a negative impact in the second year of PF on bachelor’s degrees completed at HBCUs. While estimates were consistently negative in all years analyzed, they were only significant in a few isolated years up to the 8th year of PF. The study’s results should warrant caution for policymakers using PF at HBCUs, as HBCUs can actually confer fewer degrees after the policy, compared to their PWI counterparts.

Li, Gándara et al. (2018) explored Texas and Washington 2-year PF policies, both of which allocated state revenue based on student completion of developmental
education courses (non-credit bearing), pre-college math, and pre-college English courses. These two state’s PF metrics also consisted of retention measures (i.e., when students completed 15 or 30 credit hours towards a degree, certificate, or apprenticeship).

Li, Gándara et al. (2018) calculated median total state appropriations per FTE distributed to 2-year MSIs in Texas and in Washington and descriptively compared these to non-MSIs in the respective states. In Texas, in the decade before PF, MSIs received higher median amounts of state funding per FTE student compared to non-MSIs. After PF, MSIs earned less than non-MSIs, although the authors noted this change was likely not due to PF specifically but rather to large total enrollment increases at MSIs. In Washington, descriptive trends showed that MSIs were not disproportionately affected by PF in terms of overall state funding per FTE student.

The study by Li, Gándara et al. (2018) found that MSIs in Texas earned slightly more performance points per student than non-MSIs on students completing their first credit-bearing math course, and on transfer to 4-year colleges. Non-MSIs earned more points for students completing their first 15 semester hours and for critical field degree completions (STEM). The metrics for developmental education and gateway courses seemed to earn the same points for Texas MSIs and non-MSIs both. In Washington, MSIs tended to earn more points per student on developmental education (basic literacy and math). Yet, non-MSIs earned more for pre-college math and English course sequences and for degree, certificate, and apprenticeship completion.

The authors concluded that the “on-the-route” performance metrics helped to prevent disproportionate funding losses that MSIs sometimes incur under PF policies that focus primarily on graduation outcomes. This study, while descriptive, stresses the importance of analyzing MSIs on specific performance metrics and not just on overall completion or access metrics (Li, Gándara et al. 2018). Especially for state-specific studies, it would be interesting to compare how MSIs (and different types of MSIs) do on performance metrics focused on intermediary outcomes.

The research on PF and MSIs is still growing and generally points to the need to carefully consider the design of PF metrics. MSIs can be at risk of losing funding under PF, especially if metrics are not judiciously designed to consider the characteristics of students served by MSIs. Progress metrics can be protective in recognizing the steps students take on their way to earning a degree. The few studies examining MSIs show that findings on degree outcomes are quite consistent with findings that do not differentiate between minority and non-minority serving institutions. That is, the lack of positive impacts of PF on degree outcomes seems to emerge at MSIs as well (Boland 2020; Hu 2019).

McKinney and Hagedorn (2017) (not included in Table 1) analyzed student unit record data on a large community college district (consisting of multiple campuses) in Texas to determine the relationship between student demographic characteristics and performance-based funds earned. This study was unique and did not fit comfortably in any of the sections I created for PF evaluation studies. Thus, I include it here since it considers how colleges earn funds on individual performance metrics, similar to the Li, Gándara et al. (2018) study. McKinney and Hagedorn (2017) tracked a cohort of first-time students who enrolled during the fall 2007 semester
until the summer 2013 semester. The authors explored two subsamples: college-ready students enrolled in academic programs or technical/workforce programs, and students who were required to take developmental math (non-credit bearing).

McKinney and Hagedorn (2017) were interested in the cumulative amount of performance-based funds that each student earned for their college during their 6 years of enrollment, by meeting progression metrics. They conducted a regression analysis to predict the total PF funds each student earned, based on demographic and academic characteristics. Additionally, they conducted a logistic regression determining the likelihood that a student would earn no performance funds.

McKinney and Hagedorn (2017) showed that on average, students who earned the most PF funds for their college held the following characteristics: Asian, aged 19 or younger, completed a high school diploma (versus GED), attended college full-time, received Pell grants, and were assigned to developmental coursework one level below college level. Additionally, students with the following characteristics earned significantly less funds and were most likely to earn zero performance-based funds: African American, aged 20 or older, GED holder, enrolled part-time, and assigned to the lowest level of developmental math (McKinney and Hagedorn 2017).

Interestingly, students that started out in developmental coursework earned more funds than college-ready students, likely because the Texas model rewarded developmental course completion (McKinney and Hagedorn 2017). The authors’ results suggested that the Texas PF funding model may disincentivize colleges from enrolling underserved students, specifically those least prepared for college-level coursework. One suggestion they offered was for Texas to include equity metrics for underserved students.

**Access and Admissions Outcomes**

Umbricht et al. (2017) examined Indiana’s PF policy across years 2003 to 2012 using DD and three control groups: public 4-year institutions in non-PF states with similar demographic characteristics (KY, MO, WI); private 4-year institutions in Indiana; and public 4-year institutions in surrounding, non-PF states (IA, IL, KY, MO, MN, WI). The results of this study and others on access, admissions, and student composition outcomes are displayed in Table 4.

Across all models generated by Umbricht et al. (2017), there were no significant changes in the 3-year rolling average of the number of bachelor’s degrees awarded. The number of racial/ethnic minority FTE students for the entering class declined in the 2-year lag models among PF institutions in Indiana, compared to in-state privates, and to geographically nearby public colleges. The authors did not define how minority students were determined in their dataset, so one might assume these are students reported in IPEDS as any race other than white.

Compared to the private Indiana institutions control group, there was an increase in the number of entering students on a federal grant among public Indiana institutions, yet no change was detected when compared to the other two control groups. The authors noted that these results might be explained by the start of Indiana’s
Twenty-First Century Scholars program aimed at providing tuition-free college for low-income students attending an in-state public college (Umbricht et al. 2017).

Across all control groups and in different lagged models, Indiana public institutions observed an increase in the 25th percentile ACT score for the entering class, an increase in the 75th percentile ACT score, and a decline in the admissions rate. This is the most compelling take-away from the Umbricht et al. (2017) study. That is, institutions in Indiana appeared to engage in practices of increasing selectivity, which can be presumed as unintended by the designers of the PF policy.

In another study on PF in Indiana, Birdsall (2018) analyzed data among all public, 4-year institutions in Indiana from years 2001 to 2015 using DD. The unique component of this study was that the author separated institutions based on the percentage of operating revenue each institution received from the state. The main model showed that PF had small positive impacts on six-year graduation rates among first-time, full-time entering students seeking a bachelor’s degree, for institutions receiving less than 34% of their revenue from state funding (e.g., Indiana University-Bloomington, Purdue University). Yet, institutions receiving over 34% of revenue from state funding observed null and potentially negative impacts on graduation rates. Additionally, PF produced a negative effect on total bachelor’s degrees conferred per 100 full-time equivalent students for institutions receiving 34% or more of operating revenue from the state, but not for institutions receiving less than 34% revenue from the state (Birdsall 2018). In other words, institutions more dependent on the state for funding experienced more negative degree completion outcomes.

Further, Birdsall (2018) found that most Indiana institutions experienced a decline in acceptance rates after PF, compared to control institutions in other states not subject to PF, which is consistent with findings from Umbricht et al. (2017). Additionally, PF was associated with declines in the enrollment of minority students (defined by Birdsall as Hispanic and African American students), in the main models using a national control group. However, this finding did not hold in an event study analysis nor with the control group of institutions located in states that are members of the Midwestern Higher Education Commission. The decline in minority student enrollment was not associated with the institution’s level of dependence on state appropriations.

In summary, Birdsall (2018) inferred that resource dependent institutions were less able to change their practices to respond to performance outcomes, compared to those institutions with more ability to utilize alternative revenue sources. The study showed that Indiana’s more selective research institutions were able to improve graduation outcomes under PF, probably because these are the types of institutions most able to strategically manage their student inputs. These selective institutions (with less reliance on state appropriations) also became more selective compared to other institutions that were more dependent on state revenue (Birdsall 2018).

Birdsall’s (2018) finding suggests a stratification in higher education which gets intensified under PF. The more selective institutions have better capability to strategically manage enrollment under a PF policy and perhaps enroll fewer numbers of academically underprepared students, which may further limit the opportunity for
such students to attend institutions with more robust resources that produce better graduation outcomes.

**Equity Metrics and Enrollment Outcomes at Four-Year Colleges**

In addition to examining impacts of a general PF policy, a few studies have considered equity metrics embedded within policies. Kelchen (2018) explored 4-year institutions from years 2004 to 2014 and in addition to using a binary indicator for whether a state had PF, considered whether the formula had equity metrics to incentivize the enrollment of any of following underserved student groups: low-income, minoritized, and/or adult students. The author classified non-minoritized students as those who were White or Asian. Kelchen (2018) analyzed a series of undergraduate enrollment outcomes: minority students who were African American, Hispanic, Native American, and/or Multiracial; African American students; Hispanic students; adults aged 25 and older; and Pell grant recipients.

Kelchen (2018) found that compared to colleges not under PF, colleges under any PF system showed no differences in any of the outcomes in total undergraduate enrollment. However, compared to colleges not under PF, colleges under PF systems with equity metrics experienced an increase in African American students, although no changes in other enrollment outcomes. Compared to PF without equity metrics, colleges under PF with equity metrics experienced a decline in students aged 24 and younger. No differences in total undergraduate enrollment outcomes were observed between PF policies without equity metrics and the absence of PF (Kelchen 2018).

Examining first-time students, full-time FAFSA filers (and not all undergraduates) on federal grant aid, family incomes less than $30 K, and family incomes between $30 K and $75 K, Kelchen (2018) found no differences across all the control groups for PF policies. Among first-time, degree-seeking students, the enrollment of Hispanic students increased under PF policies versus no PF policies. The percent of African American students increased among policies with equity metrics versus no PF policies.

In separate analyses, Kelchen (2018) used Barron’s Educational Series competitive ratings to determine selectivity, designating colleges shown as “very competitive” or above being in the selective category, compared to all other colleges determined as less selective. Compared to no PF policy, the number of Pell grant recipients (all undergraduates) declined at more selective colleges under any PF policy and under any PF policy without equity metrics. The number of non-Pell grant recipients declined at more selective colleges under PF systems with equity metrics compared to no PF and declined at less selective colleges under PF systems with equity metrics compared to non-equity metric PF systems (Kelchen 2018).

Kelchen (2018) demonstrated that the enrollment of African American students increased at less selective colleges under any PF policy compared to no PF policy and at less selective colleges under equity metric PF systems compared to no PF policy. Lastly, the number of students aged 24 or younger amongst all
undergraduates declined at less selective colleges under PF equity metrics compared to PF systems without equity metrics (Kelchen 2018).

As a whole, Kelchen’s (2018) study showed an absence of systematic declines in Hispanic and African American students under PF policies, which may offer encouraging evidence that with the use of equity metrics, states can curtail colleges’ inclination to enroll fewer historically underserved students. The author inferred that PF policies with any type of equity metric produced increases in African American student enrollment in particular, especially for less selective institutions. More selective colleges under PF without equity metrics enrolled fewer Pell grant recipients, with colleges under equity metrics seemingly experiencing a smaller decline.

Gándara and Rutherford (2020) conducted a study on years 2001 to 2014, focusing on 4-year public institutions subject to PF and used DD to analyze PF impacts on the following outcomes: admissions rates; the 25th percentile and 75th percentile of standardized SAT test scores, in which the verbal/English and math scores were combined for each test (ACT scores were converted to SAT scores); and enrollment of underrepresented students. This study examined a number of results, split the sample by selectivity, split the sample by PF 1.0 versus PF 2.0, and considered the impact of PF policies with an equity metric to enroll underserved students.

After PF, institutions experienced a decline in the admissions rate, and the duration of policy time variables indicated that it would take 6–7 years for this negative relationship to normalize (Gándara and Rutherford 2020). Institutions under PF observed an increase in the 25th percentile SAT scores of first-time, full-time students. This effect wore off over time when considering the duration variable. No changes in the 75th percentile of SAT scores was evident.

Gándara and Rutherford (2020) also used a mean admissions rate of 70% to split their sample into two groups: institutions with low or with high selectivity. Among low-selectivity institutions, admissions rates declined after PF, and both the 25th and 75th percentile test scores were significantly higher. Among high selectivity institutions, the admissions rate also declined, although no changes were found for the 25th and 75th percentile test scores. Furthermore, the authors showed that the decline in admissions rates were driven by PF 2.0 policies, and not PF 1.0. Similarly, in analyses that split the treatment variable into PF 1.0 versus 2.0 policies revealed that the 25th and 75th percentile scores increased only under 2.0 policies (Gándara and Rutherford 2020).

In analyzing the number of underserved students total (not just entering students), Gándara and Rutherford (2020) found, on average, a decline in Black students, perhaps a decline in Hispanic students (estimate did not reach statistical significance), yet an increase in Hispanic students when using the duration variable. Their findings also revealed no changes in the percentage of students who ever received a Pell grant and a decline in the percentage of first-generation students. When split by low and high selectivity, the impact of PF on student composition showed null impacts on average, with only the percentage of first-generation students at high selectivity institutions experiencing an increase after PF. The findings on
underserved student enrollment were not consistent across models, so the authors advised caution in the interpretation of results.

When separated by 1.0 versus 2.0, PF 1.0 appeared to decrease the number of Black students and the percentage of first-generation students (Gándara and Rutherford 2020). PF 2.0 increased the percentage of Pell grant recipients (which the authors noted could be because the share of undergraduates receiving Pell increased after the Great Recession, which overlapped with PF 2.0 adoption) and decreased the percentage of first-generation students. Institutions under PF policies that had an equity metric exhibited increases in the percentage of Pell grant recipients yet showed declines in the number of Black students and in the share of first-generation students. Institutions under PF policies without equity metrics experienced declines in both Black and Hispanic enrollment (Gándara and Rutherford 2020).

From an array of interesting findings, Gándara and Rutherford (2020) concluded that PF, especially 2.0 versions, limited access for students because institutions became more selective, which is concerning because PF may increase the already existing racial gaps in standardized test score achievement. Students who are historically marginalized may be less able to enter institutions with higher selectivity that house the resources necessary to help facilitate retention and graduation. The authors also highlighted that declining enrollment of first-generation students after PF is consistent across models and emphasized that PF models rarely include first-generation as an equity metric (Gándara and Rutherford 2020).

In general, the institutions that were less selective appeared more likely to respond to PF incentives, especially later 2.0 ones, perhaps because they were more dependent on state revenue (Gándara and Rutherford 2020). This is in line with Birdsall’s (2018) finding that institutions more dependent on state funding were more likely to respond, in positive or negative ways, to PF policies.

**Performance Funding Impacts on Institutional Finances**

The following recent studies have explored how PF policies are associated with institution-level finances, such as tuition levels and revenue sources. Hu and Villarreal (2019) examined Louisiana, which is unique in that while the state legislature sets tuition levels for 2- and 4-year sectors, the PF policy (GRAD Act) allows institutions to increase their own tuition levels up to a cap of between 5% and 10%, if these institutions meet performance targets. The GRAD Act outlines target areas (degree outcomes, transfer, workforce, and economic development), although each institution gets to determine its own annual target levels (Hu and Villarreal 2019).

Using DD on data from 2005 to 2013, Hu and Villarreal (2019) compared the tuition rates of 2- and 4-year institutions in Louisiana to three control groups: institutions in non-PF states nationally; institutions in non-PF states belonging to the Southern Regional Education Board; and institutions nationally in non-PF states that held tuition-setting authority. Results from the study showed that among 2-year colleges in Louisiana, tuition and fee levels increased significantly higher after PF,
compared to institutions in each of the three control groups (Hu and Villarreal 2019). Propensity score matching confirmed the results that 2-year publics increased their tuition and fees at higher rates compared to counterfactual institutions not providing similar tuition increase mechanisms.

Similarly, 4-year institutions in Louisiana experienced increasing in-state tuition and fees after PF, compared to each of the three control groups (Hu and Villarreal 2019). On the other hand, Louisiana 4-year institutions did not see out-of-state tuition and fee levels differ significantly than what would have been observed in the absence of PF.

It would be interesting to explore whether enrollments shifted between 2- and 4-year institutions due to rapidly increasing in-state tuition. That is, perhaps students who otherwise may have attended a 4-year college in Louisiana opted for a community college with a lower sticker price (and perceived lower out-of-pocket expenses, whether this perception is accurate or not). Noted by the authors, a curious question would be whether the composition of in-state versus out-of-state students changed among 4-year institutions in Louisiana – whether the recruiting and financial aid strategies used by institutions changed in response to PF (Hu and Villarreal 2019). Further, it would be interesting to track student progress within Louisiana institutions, to determine whether the annual increase in tuition and fees may have caused students to drop out from one year to the next.

Kelchen and Stedrak (2016) examined the impact of PF on four groups of institutional financial components: per FTE student revenue categories (total, state/local appropriations, tuition and fees, auxiliary enterprises); per FTE expenditure categories (total, instruction, student services, institutional support, auxiliary enterprises); per FTE institutional grant aid categories (only for 4-years: funded grants, unfunded grants, and tuition discount rates); and institutional financial characteristics (listed tuition and fees, student loans per FTE, Pell grants per FTE, endowment per FTE (4 years only)). The authors conducted separate analyses across their 2-year college sample and their 4-year sample. Unlike many other policy evaluation studies, the authors used two-way fixed effects in the absence of a DD design setup.

In the full models, PF at 4-year colleges was not associated with any changes in per FTE revenues nor per FTE expenditures, within each individual category nor as a whole (Kelchen and Stedrak 2016). However, PF was associated with an increase in unfunded grants for students, within the institutional grant aid category, which the authors noted was often merit-aid. Most interestingly, colleges under PF received less per FTE in Pell grant revenue compared to colleges in non-PF states, indicating a shift toward enrolling students from higher-income families.

Among the full models for 2-year colleges, under PF, Kelchen and Stedrak (2016) found that colleges received less in revenue for auxiliary enterprises. Consistent with findings among 4-year colleges, 2-year colleges also received less per FTE in Pell grant dollars when operating under a PF policy, which the authors explained as colleges potentially enrolling fewer low-income students. Note that some results were reported even though estimates only meet the \( p < 0.10 \) threshold, which I excluded in this discussion.
More research, likely in the qualitative vein, is needed to determine whether and how 2-year colleges targeted students with less need, because as open access institutions, 2-year colleges are less able to modify admissions criteria to favor less needy students. Colleges may have altered their Pell grant revenue intake by focusing marketing and recruitment in wealthier high school districts. However, it does seem to me that it would be complex for a 2-year college to strategically recruit fewer Pell grant recipients in a way that would significantly alter the composition of institutional revenues, the way that the Kelchen and Stedrak (2016) study suggests.

Hagood (2019) examined the effects of PF on state appropriations per student (fall term enrollment, not FTE enrollment) at public, 4-year institutions across years 1986 to 2014. She included policies that operated between 1995 and 2014 and excluded any policies that operated before 1995. Using DD, the author analyzed subsections of her sample by splitting up her analyses into the following institutional types: all institutions; research universities; master’s comprehensive colleges; and bachelor’s granting colleges. In a second set of analyses, she split her full sample into highly selective; less selective; and nonselective colleges. Third, she split her sample between MSIs and non-MSIs and fourth, between rural and non-rural institutions.

Hagood (2019) found that PF was associated with increased state funding per student (measured in dollars and as a percent increase) at research institutions (also robust to the addition of state-specific linear trends), as well as highly selective institutions, compared to these same peer groups in non-PF states. Master’s, bachelor’s, less selective, nonselective, and rural institutions experienced declines in state revenue per student after PF (and in percent decreases, in alternative specifications), compared to each peer group of control institutions in non-PF states. By contrast, MSIs in PF states versus non-PF states did not exhibit any significant differences in revenues per student (Hagood 2019).

Hagood (2019) analyzed a more recent time period of PF policies, years 2004 to 2014, and found that the finding of significant increases in state funding per student did not hold for research institutions. By contrast, significant declines in funding per student after PF were observed at master’s colleges and at less selective colleges.

Hagood (2019) conducted a number of robustness checks and found that within states with coordinating boards, increases in state funding per student were observed at research and highly selective institutions, while declines in funding occurred more often at non-research/non-highly selective institutions. In states with consolidated governing boards, negative PF treatment effects were observed for less selective institutions only. The author concluded that within states with coordinating boards, PF directly benefited highly resourced institutions and burdened lower resourced institutions – perhaps due to weaker ability for institutions with less lobbying power to advance their own interests and advocate for greater state funding (Hagood 2019).

Further, research institutions and highly selective institutions both experienced increases in state funding per student in both Republican- and Democratic-controlled states (Hagood 2019). While other types of institutions were protected from funding losses in Democratic-controlled states, all groups of less selective institutions in
Republican-controlled states experienced declines. Lastly, analyzing bachelor’s degrees per 100 students as an outcome, Hagood (2019) found that after PF, bachelor’s institutions (but not research nor highly selective institutions) observed increases in degrees granted, even though these institutions did not see funding increases after PF.

Hagood (2019) assessed a number of outcomes and institutional types, and the main takeaways were that high-resourced institutions were more likely to benefit from PF policies in terms of state funding per student, whereby low-resourced institutions were more likely to see funding declines. The author concluded that PF policies redistributed funding to benefit high-resourced, politically connected institutions at the expense of lower resourced institutions that actually showed improvement in bachelor’s degree completion (Hagood 2019).

Summary of Literature on Performance Funding

Equity Considerations in Performance Funding

This section further highlights equity implications of PF policies. In more recent iterations, PF has certainly attempted to address racial inequality in higher education. As noted, states have different goals and strategies for designing policies that explicitly or implicitly address disparities in college participation and completion by race and/or ethnicity. A 2017 book found that 20 of 32 policies examined did address race disparities, although 6 of these 20 states made race metrics optional and not mandatory. States with racial metrics were more commonly located in the South (e.g., Arkansas, Florida, Virginia). Other states with optional racial-disparity metrics were Illinois, Kansas, and Massachusetts (Jones et al. 2017). The use of race-based metrics is likely necessary, given the evidence that PF appears to reduce access to colleges for historically underprivileged students.

PF policies were often designed and steered by the wealthiest and most selective institutions in the state (those with more resources, abundant lobbying power, prestige, and highly academically prepared student bodies) (Gándara 2019; Jones et al. 2017). Therefore, the resulting PF policies that are implemented may benefit the institutions that are in least need of funding support. Indeed, institutions with the least resources tend to have little representation on committees tasked with designing PF policies, further perpetuating an ingrained system of funding allocation in which colleges serving more underprivileged students are provided lesser resources (Jones et al. 2017).

**Increased Selectivity.** A number of studies examining the impact of PF on admissions to public institutions have found that institutions tend to admit fewer students from disadvantaged backgrounds, in order to report better graduation outcomes (Birdsall 2018; Dougherty et al. 2016; Gándara and Rutherford 2020; Jones et al. 2017; Li and Zumeta 2016; Umbricht et al. 2017). Colleges do so by reducing the percent of applicants who are admitted or by accepting students with higher standardized test scores. In the face of PF pressures, it may not be surprising that colleges
respond by becoming more selective, since this is a logical strategy, to enroll the
types of students with academic preparedness that would increase their likelihood of
graduation.

Colleges have also shifted recruitment strategies to target better academically
prepared students or more socially advantaged students (e.g., from wealthier
neighborhoods or school districts). Furthermore, colleges may have altered their
distribution of need- and merit-based institutional aid to recruit students who are
more academically prepared, such as focusing larger portions of their institutional
scholarships on merit-based aid (Dougherty et al. 2016; Kelchen and Stedrak
2016).

Due to the consistent findings of increasing selectivity, state policymakers
should carefully monitor the admissions and recruitment practices of institutions.
Since PF appears to reduce state revenue for less selective colleges (Hagood
2019), the policy does seem to exacerbate stratification in higher education, in
terms of resources and distribution of students. Students who are less prepared
academically wind up enrolling at less selective colleges (Li 2019), which makes
it even harder for such colleges to improve their graduation outcomes. In short,
evidence points to a risk that PF can exacerbate existing stratification in higher
education.

**Equity Metrics in Drawing Attention Toward Underserved Student Enrollment.**
As discussed, one segment of the PF literature has focused on how to tweak PF
policies in order to prevent or at least limit some of the temptation colleges face
(especially given the constraints colleges deal with) to increase selectivity in order
to boost performance. Equity metrics provide additional funds for retaining and
graduating students from historically underprivileged groups, such as lower-
income, students of color, returning adult students, veterans, and first-generation
students.

To highlight a few studies, colleges, especially those with more selective criteria
that operated under PF with equity metrics were more likely to enroll Black/African
American students compared to colleges under PF systems without such metrics.
However, the impact of equity metrics on Hispanic, low-income, and older students
was more mixed (Kelchen 2018). By contrast, another study found that the enroll-
ment of Black students decreased even when equity metrics were in place (Gándara
and Rutherford 2020). Yet another study found that the inclusion of equity metrics
led to increases in low-income students, yet decreases in adult students (Li and
Ortagus 2019).

Qualitative evidence does support that drawing attention to equity metrics can
boost institutional programming to retain underprivileged students (Li 2019). Also
encouraging is that using metrics for student progress (along with graduation out-
comes) can help prevent colleges that serve higher proportions of students of color
from being unduly “punished” under PF policies (Li, Gándara et al. 2018). More
research is needed on equity provisions in particular, to build a more consistent body
of literature on whether the inclusion of such measures reduces the inclination of
colleges to enroll less advantaged students. This is especially critical as more states
recognize the importance of including equity metrics.
General Comments on Research Design of Performance Funding Studies

The literature review shows that PF studies in the most recent years utilized more robustness checks. These included using event study techniques (Gándara and Rutherford 2020), propensity score matching (Hu 2019), and multiple sets of control groups to test for robustness across models (Kelchen 2018; Li and Ortagus 2019). As the literature base on PF continues to build, reviewers are clearly expecting more attention paid toward assumptions in DD models (such as parallel trends) and more explicit recognition of differences across policy designs.

More recent studies were also more likely to stratify within their samples to examine institutions with different missions, such as lesser resourced institutions (Hagood 2019) or minority-serving institutions (Hu 2019). This stratification of the data is especially important because PF has been shown qualitatively to create strain for less selective and more resource-dependent institutions (Dougherty et al. 2016; Li 2019). Therefore, quantitative analyses should complement such qualitative evidence by differentiating between institutions that serve different student subpopulations.

Relatedly, future national studies on 4-year institutions should, as some studies do, include control variables for institutional type (e.g., bachelor’s-granting, master’s comprehensive) (Hagood 2019) or selectivity (e.g. Barron’s, or admissions rates) (Gándara and Rutherford 2020; Kelchen 2018). Studies have evolved from earlier evaluations where all 4-year colleges were observed together, to more nuanced consideration of how PF impacts different institutions in disparate ways.

In addition, some studies, especially the most recent ones, have differentiated between the PF 1.0 time period versus the PF 2.0 time period (Boland 2020; Favero and Rutherford 2020; Gándara and Rutherford 2020; Rutherford and Rabovsky 2014). This is an important methodological design choice, because PF policy designs from earlier years were indeed lower stakes. One study incorporated a variable capturing design variation explicitly (Li and Kennedy 2018), and future research should continue to pay attention to design differences between policies when conducting large-scale evaluation studies.

Evidence of Limited Impact on Degree Completion

A prevailing theme across studies of 4-year colleges is worth nothing. An overall finding of null impacts of PF on bachelor’s degree completions was generally consistent across studies. With the exception of the Li (2018) study showing positive effects on STEM degrees, the general takeaway from most recent 4-year studies was that PF, in different forms and versions of the policy, has not improved bachelor’s degree completions (Birdsall 2018; Boland 2020; Favero and Rutherford 2020; Rutherford and Rabovsky 2014; Hillman et al. 2014; Rutherford and Rabovsky 2014; Tandberg and Hillman 2014; Umbricht et al. 2017).

There are potential explanations for this finding. At research institutions, degree completion outcomes are already quite good. Increases in state funding from
boosting graduation rates (especially relative to net tuition revenues) are not going to fundamentally change the way a university educates its students. In addition, many public 4-year colleges receive a majority of their funding from tuition and not the state, so state-instituted performance regimes may not steer institutions differently since the dependence on state revenues has decreased over time. Public higher education is increasingly privatized, and colleges must rely on an adequate number of full-paying tuition students. Thus, the money itself earned from a performance-based funding formula may not be enough as an incentive, or university leaders can make up this funding through fundraising without having to shift institutional goals (Bell et al. 2018).

Furthermore, the studies that analyzed outcomes at community colleges similarly point to no changes in degree outcomes after the advent of PF. Generally, PF has not altered college practices enough to move the needle on associate degree completions in state and national studies of 2-year colleges (Hillman et al. 2018, 2015; Hu 2019; Li and Kennedy 2018; Li and Ortagus 2019; Tandberg et al. 2014). Perhaps we should not expect measurable improvements in degree outcomes in a pay for performance scheme. Student characteristics, motivations, preparedness, and life circumstances substantially affect their academic trajectories in earning a degree or not. Even if colleges and states have aligned goals to increase degree completion, much of what factors into persistence and graduation is outside of the college’s control.

Relatedly, studies that examine retention rates at community colleges may be capturing student inputs and characteristics, and not the actual PF policy at play. Especially for part-time students, it is quite challenging to change the student experience to a level strong enough to increase overall retention rates. While the idea behind PF is intuitive and college completion is consistent with the purpose of higher education, it is possible that an external policy simply does not override the incoming student academic and motivational factors that influence completion.

Data Availability for Part-Time Students. The research on PF policy evaluation focuses on first-time, full-time students. Part-time students and transfer students are ignored in the analyses of PF impacts, which leaves out an entire subpopulation of students. Among fall 2018 students at 4-year colleges, 75% enrolled full-time, while 37% of students at 2-year colleges enrolled full-time (National Center for Education Statistics 2020a). This is both a data availability issue and a practical issue. IPEDS data does not allow researchers to analyze part-time student retention and completion, which minimizes institutional and policymaker focus on this population.

International Perspective on Performance Funding

Here, I briefly describe PF policies in countries outside of the United States, drawing primarily on Dougherty and Natow (2019). As of 2010, 19 European countries, Canada, and Australia have funded higher education systems in part based on output criteria including degrees earned, credits obtained, and especially research effort and productivity. According to Bell et al. (2018), at least 14 OECD countries have
operated performance-based research funding where research performance is measured according to metrics: external research income, bibliometric citations, doctoral degree completion rates, and peer assessments of research quality. Similar to the United States, government officials outside of the United States have been instrumental in adopting PF policies (Dougherty and Natow 2019).

The issue of gaming, or colleges behaving in ways that do not necessarily promote student success but that do help boost performance on outcomes, has been shown to be an issue internationally. Neoliberal theory, which focuses on agents shirking responsibility on purpose, would explain that “slippage” is to blame, a result that could be due to conflicting interpretations and not in purposeful gaming (Dougherty and Natow 2019). Therefore, any gaming could be caused not by malicious shirking (e.g., passing students who would otherwise fail gatekeeper courses) but by a lack of mutual understanding of PF workings between principal and agent.

Studies in Europe show a lack of impact of PF on student degree completion, similar to findings in US studies. However, studies in Europe do show that PF has produced higher rates of faculty research productivity in Denmark, the Netherlands, Norway, Switzerland, the United Kingdom, and Hong Kong, although these studies were not as vigilant about controlling for other external factors related to research productivity outcomes (Dougherty and Natow 2019).

In addition, studies in Europe and Australia, similarly to the United States, have shown that PF comes with substantial compliance costs in the form of management, administration, and reporting (Dougherty and Natow 2019). In addition, unintended consequences are also prevalent among countries subject to PF abroad. For instance, institutions may reduce their academic rigor to churn out more graduates. PF incentives for research productivity was associated with faculty trying to improve their publication statistics by slicing up publications into multiple articles with similar content or pursuing less challenging research questions. Further, in England, there exists evidence that PF for research has caused scholars to shift their research areas, abandoning some areas that may be less likely to secure performance funds (Dougherty and Natow 2019).

Increased stratification among colleges is a problem abroad as it is in the United States. PF appears to increase stratification by concentrating more research funding in Oxford, Cambridge, and the Russell Group institutions and disadvantaging other institutions that also need research funding. This type of favoring of more prestigious institutions is similar to what occurred in the United States (Dougherty and Natow 2019).

The Future of Performance Funding

With a four-decade-long time frame of operating in the American states and periods of popularity and discontinuation, PF certainly has an interesting history. It will be curious to see how policy designs change in the future. Despite evidence that shows a lack of a positive impact on bachelor’s and associate degree completions, the
policy is here to stay. More attention has been paid recently on aligning equity metrics to the access and equity goals of each state, and future research will continue to examine the impacts of policy features on historically underrepresented populations, a direction which is certainly important. As seen from the literature review, there are a multitude of ways that colleges have responded to the policy in terms of intended and unintended campus responses. The culmination of published works on the policy illustrates scholars’ enduring interest in the implementation and impacts of PF.

Performance Funding During an Economic Recession

Previous research points to the discontinuation of PF to be most likely during an economic recession (Dougherty et al. 2014). This makes intuitive sense—when state budgets are tight, and colleges are in need of resources, colleges would try to protect the existing base funding they have and advocate for the discontinuation of outcomes-based funding policies. When colleges are in need of funds and facing fiscal strain, they may require additional funding to keep operations afloat and funding to sustain student and academic support services.

With the recession of 2020 caused by the Covid-19 pandemic, it will be interesting to see how states react to the decline in taxpayer revenue and how much states will fund higher education. On the one hand, advocates of PF may argue that policies have largely been institutionalized and become part of campus culture. Especially during tight budgetary time periods, it is necessary to place pressure on colleges to closely examine how they expend resources. Therefore, it is important to enforce accountability policies that demand greater attention to completion, even in lean times.

A related consideration is that during recessions, taxpayers and policymakers may in fact demand greater transparency. Therefore, the recession of 2020 and the pandemic may result in greater scrutiny of colleges, especially public institutions. There is little evidence to suggest that the student experience of a college education has improved substantially in the last decade or so. Yet, the net tuition that students pay has increased disproportionately faster than incomes and inflation, and PF fits into an overall desire for accountability.

Opponents may argue that PF policies fail to close attainment gaps in higher education by race/ethnicity, wealth, and income level. Specifically, PF appears to harm the institutions in most need of resources. Therefore, some opponents of the policy may lobby for a discontinuation of PF altogether, citing the economic recession as a good time to evaluate existing practices.

The current economic conditions could cause “a reckoning” of higher education, as institutions must learn to cut “non-essential” workers and streamline processes to make sure organizations are as lean as possible. This may include cutting administration, support staff, faculty, or changing workload duties (e.g., higher teacher loads, greater expectation of external grant securement). Performance funding fits into this broader reckoning because it has required colleges to be more cognizant of capturing
and reporting student outcomes data. In this vein, the goals of colleges (i.e., administrators tasked with making the organization leaner) and the goals of PF from the perspective of state policymakers may be quite aligned.

**Areas of Future Research for Performance Funding**

**Diversity of Policy Adopters.** One ongoing challenge in higher education that might affect whether PF reflects equity and diversity goals are the diversity of the policy designers. Limited research on PF policy creation has focused on the demographics of policymakers and institutional leaders who design and legislate such policies. It would be interesting to explore whether the decision-makers involved with crafting PF hold demographic characteristics such as race, ethnicity, and gender that reflect the faculty and staff members at college campuses that PF aims to target.

Future research may seek to examine whether the diversity of policy designers is related to the effectiveness of policy designs in addressing the needs of MSIs and of regional comprehensive institutions. Perhaps, if policy designers include alumni of HBCUs or more individuals of color, then these individuals may be able to advocate on behalf of equity metrics or a funding formula that lessens the fiscal pressure PF places on under-resourced colleges. Thus far, PF has been unable to close equity gaps and may exacerbate the inequitable funding allocations across institutions of varying lobbying power and prestige. Research should consider the diversity and backgrounds of decision-makers involved with PF policy designs and revisions.

**Inputs Drive Outputs.** Evidence on the outcomes of PF highlight the reality that institutional-level outcomes are driven heavily by inputs. Student characteristics and academic preparedness levels are major factors for whether students persist, graduate, and find jobs. Thus, not accounting for different student demographics sells short the contribution institutions make with their existing student bodies. For example, graduates of predominantly Black institutions in the United States face racial discrimination in hiring practices that make it more challenging for colleges with these alumni to produce positive job placement figures (Jones et al. 2017). Future research on policies should look for ways to reward institutions that serve higher numbers of students of color, since it does not appear that equity metrics in and of themselves are enough to reduce unintended consequences such as increasing selectivity.

**Caution in Raising Dosage Amounts.** Based on the literature, PF has been viewed with some skepticism by faculty and support staff who are responsible for teaching and advising. These individuals can play a role in whether a student chooses to persist or chooses to drop out. It does not appear, from the literature, that states with higher “doses” of funding tied to outcomes are actually experiencing increased degree attainment. Quite the contrary, states with high proportions of funding tied to outcomes may have a higher likelihood of unintended consequences such as diminishing access and increases in short-term certificates. I would caution policy designers and state officials who are considering adding greater stakes to an existing policy – because the research shows that more money tied to outcomes
could create more undesirable policy impacts rather than providing actual policy solutions.

**Capacity for Campus Initiatives.** As the research shows, colleges may be well-intended in improving student outcomes. However, lower-capacity institutions have limited capability to collect and analyze data that would help pinpoint areas of intervention along a student’s academic trajectory (Dougherty et al. 2016; Dougherty and Natow 2019). Regardless of best efforts, lower-capacity colleges may not have the staff nor data analytic capacity to take advantage of useful student engagement data on involvement in clubs and organizations, leadership programs, and tutoring. Even if colleges do find areas to buffer student support, such as providing mentors who are people of color, colleges may not have enough capacity to launch a large-scale mentorship program that likely requires years of implementation before graduation outcomes are impacted.

Thus, additional research may examine the ways states can help build greater capacity at institutions needing more support. Analyzing rural and less selective colleges may be a good place to start. The PF literature would benefit from a better understanding of state and local investments needed to implement campus support initiatives aimed at improving retention and completion.

**Goal Alignment.** In some contexts, institutions may not behave in ways that support the outcomes incentivized in a PF policy but rather, behave in other ways that are more consistent with the institution’s own values. For instance, if an institution is located in a rural area and sees its mission at serving its local population, it may be less inclined to increase admissions standards to accept more academically qualified students and recruit from a broader geographic area. Correspondingly, researchers should consider the existing and perhaps shifting mission of institutions as they engage in strategic enrollment management.

**Methodological Considerations**

Numerous studies mentioned in this chapter use difference-in-differences regression techniques to analyze the aggregate policy impact of PF on specified outcomes. There are certain areas of methodological consideration when undertaking future studies.

**Private Colleges as a Feasible Control Group.** First, there are interesting ways to leverage the use of multiple differences in evaluation studies of PF. As described, the first difference is between outcomes before and after PF adoption, while the second difference is between colleges subject to PF versus similar colleges not subject to PF. Because PF affects only public institutions, it may be wise for more researchers to consider private colleges as a feasible control group, beyond the few studies that have (Boland 2020; Umbricht et al. 2017).

Some considerations are warranted in determining control group section criteria. The underlying student population is critical, and researchers would want to match colleges with similar missions. The comparison would be that colleges within a resource-dependence perspective, that is, public colleges dependent upon state
appropriations have a compelling reason to improve degree outcomes compared to private colleges, which depend more heavily on student tuition and fees. The challenge here would be justifying that public and private colleges are similar enough to serve as feasible control groups for one another. The budgeting procedures and timelines of private colleges are different and may affect their admissions and recruitment procedures.

**Performance Funding Impacts on Tuition Prices.** The adoption of PF could cause institutions to raise tuition to make up for potential lost revenue. This also could lead to fewer students believing they can afford higher education, especially first-generation and disadvantaged students. Thus, PF could indirectly harm such students by disincentivizing them to enroll. Further, PF changes the allocation of state appropriations, or perhaps the allocation of appropriations changes the design of PF (e.g., open-access institutions believe they are unfairly funded under existing funding models and advocate for equity metrics in revisions of the policy). Thus, a number of different financial choices are being made with respect to the state allocating funding for colleges and colleges allocating funding within departments. Future research may be needed to examine how allocations within colleges change after PF, especially toward instructional expenses and student service expenditures, and compare across different types of colleges to analyze variation in resource allocation.

**Accounting for Student Inputs.** One challenge in analyzing aggregate policy impacts is separating graduation outcomes and outputs from inputs – the characteristics of entering students. Research clearly demonstrates that student academic preparedness, race/ethnicity, first-generation status, and enrollment intensity are correlated with whether that student persists from year to year and whether that student ultimately graduates. Regardless of how much an institution invests in academic or student services, it is possible that certain students faced with serious socioeconomic disadvantages (such as attending a poorly funded high school) may be at risk of dropping out even with the availability of such resources.

If higher education is about expanding access, then PF creates a challenging environment where it is difficult for middle-tier public colleges to succeed in fulfilling their mission to provide access for first-generation, low-income, and students of color. Future research should consider ways that PF policies may be punishing colleges that have high populations of underserved students and perhaps consider modifying policy goals for such colleges. Indeed, the literature on PF advises caution in applying a one-size-fits-all approach to funding colleges based on outcomes.

**Conclusion**

The research on performance funding has been quite expansive, covering its adoption patterns and origins, to numerous studies on policy implementation and impacts, many concentrated in the last decade. These studies have applied a number of
theoretical frameworks, from new public management, resource dependence theory, principal-agent theory, neoliberalism, and the ecology of games. Studies have examined the adoption and policy design phase of PF, considering the state characteristics and political influences that gave rise to the policy. A number of studies have explored campus responses to the policy, finding that PF has resulted in anticipated and unanticipated changes in terms of curriculum, advising, programming, and recruitment. Furthermore, a wide range of studies have analyzed the policy’s impacts on retention, degree outcomes, minority-serving institutions, access, student demographics, and institutional finances. This chapter has surveyed the different elements of research on PF and offered some suggestions for future examination of PF.

References

Banta, T. W., Rudolph, L. B., Van Dyke, J., & Fisher, H. S. (1996). Performance funding comes of age in Tennessee. The Journal of Higher Education, 67(1), 23–45.

Bell, E., Fryar, A. H., & Hillman, N. (2018). When intuition misfires: A meta-analysis of research on performance-based funding in higher education. In E. Hazelkorn, A. McCormick, & H. Coates (Eds.), Research handbook on quality, performance and accountability (pp. 108–124). Northampton: Edward Elgar Publishing.

Berry, F. S., & Berry, W. D. (1990). State lottery adoption as policy innovation: An event history analysis. American Political Science Review, 84(2), 395–415.

Birdsall, C. (2018). Performance management in public higher education: Unintended consequences and the implications of organizational diversity. Public Performance and Management Review, 41(4), 669–695. https://doi.org/10.1080/15309576.2018.1481116.

Bohren, O. (1998). The agent’s ethics in the principal-agent model. Journal of Business Ethics, 17(7), 745–755.

Boland, W. C. (2020). Performance funding and historically black colleges and universities: An assessment of financial incentives and baccalaureate degree production. Educational Policy, 34(4), 644–673. https://doi.org/10.1002/cc.154.

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

Burke, J. C., & Minassians, H. P. (2003). Performance reporting: “Real” accountability or accountability “lite”: Seventh annual survey 2003. Albany: The Nelson A. Rockefeller Institute of Government.

Burke, J. C., & Minassians, H. P. (2004). Implications of state performance indicators for community college assessment. New Directions for Community Colleges, Summer, (126), 53–64. https://doi.org/10.1002/cc.154.

Cavanaugh, J. C., & Garland, P. (2012). Performance funding in Pennsylvania. Change, 44(3), 34–39.

Dougherty, K. J., & Hong, E. (2006). Performance accountability as imperfect panacea: The community college experience. In T. Bailey & V. S. Moster (Eds.), Defending the community college equity agenda (pp. 51–86). 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. (2019). Analysing neoliberalism in theory and practice: The case of performance-based funding for higher education (No. 44). London. Retrieved from www.researchcghe.org

Dougherty, K. J., & Reddy, V. (2013). Performance funding for higher education: What are the mechanisms? What are the impacts? ASHE Higher Education Report, 39(2), 1–135.
Dougherty, K. J., Natow, R., Hare, R., Jones, S., & Vega, B. (2011). The politics of performance funding in eight states: Origins, demise, and change: Final report to Lumina Foundation for Education. New York: Community College Research Center. Retrieved from http://eric.ed.gov/?id=ED517751

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. https://doi.org/10.1177/002716214541042.

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.

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. https://doi.org/10.1007/s11162-019-09551-1.

Frølich, N. (2011). Multi-layered accountability. Performance-based funding of universities. Public Administration, 89(3), 840–859. https://doi.org/10.1111/j.1467-9299.2010.01867.x.

Furquim, F., Corral, D., & Hillman, N. (2020). A primer for interpreting and designing difference-in-differences studies in higher education research. Higher Education: Handbook of Theory and Research, 35, 667–723.

Gándara, D. (2019). Does evidence matter? An analysis of evidence use in performance-funding policy design. The Review of Higher Education, 42(3), 991–1022. https://doi.org/10.1353/rhe.2019.0027.

Gándara, D. (2020). How the sausage is made: An examination of a state funding model design process. Journal of Higher Education, 91(2), 192–221. https://doi.org/10.1080/00221546.2019.1618782.

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. https://doi.org/10.1007/s11162-017-9483-x.

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, 1–14. https://doi.org/10.3102/0013189X20927386.

Gándara, 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. https://doi.org/10.1080/00221546.2016.1272089.

Hagood, L. P. (2019). The financial benefits and burdens of performance funding in higher education. Educational Evaluation and Policy Analysis, 41(2), 189–213. https://doi.org/10.3102/0162373719837318.

Hillman, N. W., & Corral, D. (2018). The equity implications of paying for performance in higher education. American Behavioral Scientist, 61(14), 1757–1772. https://doi.org/10.1177/0002764217744834.

Hillman, N. W., Tandberg, D. A., & Gross, J. P. K. (2014). Performance funding in higher education: Do financial incentives impact college completions? The Journal of Higher Education, 85(6), 826–857. https://doi.org/10.1353/jhe.2014.0031.

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. https://doi.org/10.3102/0162373714560224.

Hillman, N. W., Hicklin Fryar, A., & Crespin-Trujillo, V. (2018). Evaluating the impact of performance funding in Ohio and Tennessee. American Educational Research Journal, 55(1), 144–170. https://doi.org/10.3102/0028312117732951.

Holmstrom, B. (1979). Moral hazard and observability. The Bell Journal of Economics, 10(1), 74–91.

Hu, X. (2019). Efficiency for whom? Varying impact of performance-based funding on community colleges in Louisiana. Community College Review, 47(4), 323–359. https://doi.org/10.1177/0015552119864409.
Hu, X., & Villarreal, P., III. (2019). Public tuition on the rise: Estimating the effects of Louisiana’s performance-based funding policy on institutional tuition levels. *Research in Higher Education, 60*(5), 636–669. https://doi.org/10.1007/s11162-018-9526-y.

Jones, T. (2016). A historical mission in the accountability era: A public HBCU and state performance funding. *Educational Policy, 30*(7), 999–1041. https://doi.org/10.1177/0895904815586852.

Jones, T., Jones, S. M., Elliott, K. C., Owens, L. R., Assalone, A., & Gandara, D. (2017). *Outcomes-based funding and race in higher education: Can equity be bought?* New York: Palgrave Macmillan.

Kelchen, R. (2018). Do performance-based funding policies affect underrepresented student enrollment? *The Journal of Higher Education, 89*(5), 702–727. https://doi.org/10.1080/00221546.2018.1434282.

Kelchen, R., & Stedrak, L. J. (2016). Does performance-based funding affect colleges’ financial priorities? *Journal of Education Finance, 41*(3), 302–321.

Kivistö, J. (2008). An assessment of agency theory as a framework for the government–university relationship. *Journal of Higher Education Policy and Management, 30*(4), 339–350. https://doi.org/10.1080/13600800802383018.

Lane, J. E. (2007). The spider web of oversight: An analysis of external oversight of higher education. *The Journal of Higher Education, 78*(6), 615–644. https://doi.org/10.1353/jhe.2007.0038.

Letizia, A. J. (2016). The evolution of control: The convergence of neoliberalism and neoconservatism in performance based funding policies. *Critical Education, 7*(2), 1–18. https://doi.org/10.1007/978-1-137-42617-8_15.

Li, A. Y. (2017a). Covet thy neighbor or “reverse policy diffusion”? State adoption of performance funding 2.0. *Research in Higher Education, 58*(7), 746–771. https://doi.org/10.1007/s11162-016-9444-9.

Li, A. Y. (2017b). The point of the point: Washington’s student achievement initiative through the looking glass of a community college. *Community College Journal of Research and Practice, 41*(3), 183–202. https://doi.org/10.1080/10668926.2016.1179601.

Li, A. Y. (2018). *Lessons learned: A case study of performance funding in higher education.* Washington, DC: Third Way. Retrieved from https://www.thirdway.org/report/lessons-learned-a-case-study-of-performance-funding-in-higher-education

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. (2020). Performance funding policy impacts on STEM degree attainment. *Educational Policy, 34*(2), 312–349. https://doi.org/10.1177/0895904818755455.

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.

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. https://doi.org/10.1353/rhe.2019.0097.

Li, A. Y., & Zumeta, W. (2016). *Performance funding on the ground: Campus responses and perspectives in two states.* New York: TIAA Institute. Retrieved from https://www.tiaainstitute.org/publication/performance-funding-ground-campus-responses

Li, A. Y., Gándara, D., & Assalone, A. (2018). Equity or disparity: Do performance funding policies disadvantage two-year minority-serving institutions? *Community College Review, 46*(3), 288–315. https://doi.org/10.1177/0091552118778776.

Li, A. Y., Kennedy, A. I., & Sebastian, M. L. (2018). *Policy design matters: The impact of performance funding policies on credential completion at community colleges.* Madison: Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE). Retrieved from https://wiscape.wisc.edu/wiscape/publications/policy-briefs/pb028
Mckeown-Moak, M. P. (2013). The “new” performance funding in higher education. *Educational Considerations, 40*(2), 3–12.

McKinney, L., & Hagedorn, L. S. (2017). Performance-based funding for community colleges: Are colleges disadvantaged by serving the most disadvantaged students? *The Journal of Higher Education, 88*(2), 159–182. https://doi.org/10.1080/00221546.2016.1243948.

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. https://doi.org/10.3102/0162373702801001.

Meier, K. J., & O’Toole, L. J. (2009). The proverbs of new public management: Lessons from an evidence-based research agenda. *The American Review of Public Administration, 39*(1), 4–22. https://doi.org/10.1177/0275074008326312.

Miller, G. N. S., & Morphew, C. C. (2017). Merchants of optimism: Agenda-setting organizations and the framing of performance-based funding for higher education. *The Journal of Higher Education, 88*(5), 754–784. https://doi.org/10.1080/00221546.2017.1313084.

National Center for Education Statistics. (2020a). *Characteristics of postsecondary students*. Retrieved July 12, 2020, from https://nces.ed.gov/programs/coe/indicator_csb.asp

National Center for Education Statistics. (2020b). *Undergraduate retention and graduation rates*. Retrieved July 20, 2020, from https://nces.ed.gov/programs/coe/indicatorCtr.asp

Nisar, M. A. (2015). Higher education governance and performance based funding as an ecology of games. *Higher Education, 69*(2), 289–302. https://doi.org/10.1007/s10734-014-9775-4.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.

Rosinger, K., Ortagus, J., Kelchen, R., Cassell, A., & Voorhees, N. (2020). *The landscape of performance-based funding in 2020*. Retrieved from https://informedstates.org

Rutherford, A., & Rabovsky, T. M. (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. https://doi.org/10.1177/0002716214541048.

Sanford, T., & Hunter, J. (2011). Impact of performance-funding on retention and graduation rates. *Education Policy Analysis Archives, 19*(33), 1–30.

Shin, J. C. (2010). Impacts of performance-based accountability on institutional performance in the U.S. *Higher Education, 60*(1), 47–68. https://doi.org/10.1007/s10734-009-9285-y.

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, 12*(22), 1–26.

Snyder, M., & Fox, B. (2016). *Driving better outcomes: Fiscal year 2016 state status & typology update*. Washington, DC: HCM Strategists. Retrieved from http://hcmstrategists.com/drivingoutcomes/

Tandberg, D. A., & Hillman, N. W. (2014). State higher education performance funding: Data, outcomes, and policy implications. *Journal of Education Finance, 39*(1), 222–243. https://doi.org/10.1353/jef.2014.0007.

Tandberg, D. A., Hillman, N. W., & Barakat, M. (2014). State higher education performance funding for community colleges: Diverse effects and policy implications. *Teacher’s College Record, 116*(120307), 1–31.

Umbricht, M. R., Fernandez, F., & Ortagus, J. C. (2017). An examination of the (un)intended consequences of performance funding in higher education. *Educational Policy, 31*(5), 643–673. https://doi.org/10.1177/0895904815614398.

Volkwein, J. F., & Tandberg, D. A. (2008). Measuring up: Examining the connections among state structural characteristics, regulatory practices, and performance. *Research in Higher Education, 49*(2), 180–197. https://doi.org/10.1007/s11162-007-9066-3.

Zumeta, W., & Li, A. Y. (2016). *Assessing the underpinnings of performance funding 2.0: Will this dog hunt?* New York: TIAA Institute. Retrieved from https://www.tiaainstitute.org/public/pdf/ti_assessing_the_underpinnings_of_performance_funding_2.pdf
Amy Li is an Assistant Professor of Higher Education at Florida International University. Her research focuses on higher education finance and public policy, specifically performance funding, promise programs, student loan debt, state appropriations, and policy adoption. She studies the impact of local and state policies on college access and completion and is particularly interested in outcomes for historically underrepresented populations.