

**Diffusion of Policy Innovations: State Adoption of Value-Added Models in K-12
Education**

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Kirsten England
May 2022

Copyright © 2022 by Kirsten England.
All rights reserved.

DEDICATION

To my husband and close friends: You have always been my support system and I could not ask for better.

Also, to my beloved fur kids, Oliver, Maddie, Logan, and Theo, who bring me endless joy and stress relief. To Cameron, you may no longer be here physically, but you were my greatest happiness, and I miss you more than words can express.

ACKNOWLEDGEMENTS

The Political Science Department at the University of Tennessee has been so important in forming the researcher I am. My journey began like many adventures in my life, on a whim. I woke up one day and thought “I think I need my doctorate in order to make real change within a profession I love” and so I did just that. Along this journey, I must issue endless thanks to my dissertation chair Dr. David Houston who has been a source of endless support and guidance. Your patience has been all enduring through the bumpy road to bring this research to completion, and you have been an incredible mentor along the way.

Additionally, thank you so much to the rest of my committee – Dr. Michael “Doc” Fitzgerald for always being there for me by simply listening and giving amazing advice, Dr. John Scheb for having me always being frank and honest with suggestions, and Dr. Patrick Biddix for adding expertise and guidance to both deepen and broaden my policy base.

Along the path, I must thank those who always have helped me to achieve my goals. They include: Dr. David Folz, Dr. Joe Jarret, Mrs. Dianna Beeler and Mrs. Leslie Tolman thank you.

A very special thank you to Dr. Richard Pacelle who took me under his wing and agreed to work with me as a co-author and share his expertise. What I learned was invaluable and I cannot thank you enough.

To my friends Alexandra Rotzer, Erin Rowland, Rebecca Schoffman, and Melanie Hannan: you have seen me at my best and worst and stuck by me, encouraging me to strive for better. I very special thank you and acknowledgement is needed for Rachel Vincent who helped me to learn EHA when she did not even know it herself at the time. Rachel, this was done by “blind mice” and yet here I am! You are all incredibly intelligent, kind women who are, in many cases, how I was able to navigate my way through this crazy ride of grad school. I am forever in your debt.

Finally, thank you to Spencer, my husband. Simply put, I love you.

ABSTRACT

Schools chase test scores as a means of validating educational practices. Having imposed high-stakes testing, public schools were still not generating the student test scores under policies such as No Child Left Behind, the Department of Education thought would ensue. To hold accountability on the part of the classroom teacher, Race to the Top, and with it, merit pay was imposed by most states within the United States.

Using policy innovation and diffusion frameworks, this research looks at how over time states have come to adopt value-added policies for teacher evaluations. The research takes into consideration the sociopolitical systems that influence adoption of within the K-12 system.

I test six key hypotheses using the event history methodology and building off previous literature which focused on higher education, the research seeks to further the understanding of educational policy diffusion and the driving factors for value-added adoption. In the analysis of the data, it was found that both internal and external determinants were key in understanding how and why value-added policies are adopted by states.

TABLE OF CONTENTS

Chapter 1	1
Introduction.....	1
Internal versus External determinants.....	4
Diffusion of Meritocratic Policies and Performance-Based Funding.....	5
Proposed Research Contribution.....	9
Chapter 2.....	13
Teaching and Meritocracy	13
A Nation at Risk.....	13
Socio-economic Considerations.....	16
Micro-Credentials	24
Summary	27
Chapter 3.....	29
Performance-Based Funding in Higher Education	29
Performance-Based Funding Background	30
Principal/Agent Theory.....	33
Business and Performance-Based Funding.....	35
PBF 1.0 and 2.0.....	36
The Effect of Performance-Based Funding	39
Summary	46
Chapter 4.....	47
Value-Added Models	47
Space Race, the beginning of federally funded curriculum changes	48
Teacher Evaluation Systems	50
Additional Policies that Lead to Value Added Models	54
Assessments of Value-Added Models	67
Summary	77
Chapter 5.....	79
Innovation of Diffusion.....	79
What is Diffusion and Why it is Important.....	80
Mechanisms for Diffusion	83
Situation Framework.....	87
Model Building of Value-Added Policies.....	91
Regional Diffusion.....	92
Federalism.....	94
Internal and External Determinants	96
Additional Explanations for Adoption.....	99
Summary	100
Chapter 6.....	101
Data and Methods- Hypotheses for State Adoption of Value-Added Models.....	101
Hypotheses.....	102
Internal Factors	106
Data	111
Model Estimation Method	116

Chapter 7	118
Results	118
Results and Discussion	120
Conclusions	132
Chapter 8	136
Conclusion, Implications, and Further Work	136
Race to Top Coercion	139
Policy Entrepreneurs	142
Accountability Policies and Where We are Headed	144
What Can be Further Explored	148
Works Cited	150
Vita	175

LIST OF TABLES

Table 7.1	Policy Indicators.....	121
Table 7.2	Federal Circuit Courts.....	129
Table 7.3	Border State Diffusion.....	133

LIST OF FIGURES

Figure 7.1	Kaplan Meier Estimation of Diffusion.....	119
Figure 7.2	Value-Added Heat Map.....	122
Figure 7.3	Performance-Based Funding Heat Map.....	123
Figure 7.4	Regional Breakup.....	124

CHAPTER 1

Introduction

Throughout the last 30 years, the concept and practice of accountability in U.S. public education have undergone marked change. Early on, accountability often referred to the “design of statewide governance structures capable of accommodating the simultaneous need for institutional autonomy and external oversight of decision-making” (McLendon et al., page 1, 2006). Therefore, the question needing to be addressed by policy makers is: which activities and functions of educational institutions such as academic programs and learning objectives should be dictated by the state and which should be left for the schools or districts to decide (McLendon et al., 2006)? Accountability became a function of governance systems capable of both effectively and efficiently controlling resources and the decisions of campus leaders (Berdahl, 1971).

There are many reasons to study the determinants of state performance-accountability or meritocratic policies in education. This era of new accountability, based on meritocratic principles, represents a mile-marker in the dynamics of both federal-state and state-district relationships, both of which touch on important findings for the financing and governing of public education. The measurement of student performance, and by extension teacher performance, assumes that if individuals work hard enough, they can be successful at any endeavor. However, this ignores that factor that it is more than effort that influences performance. Context has a heavy hand in outcomes. These recent forms of accountability are based on measuring teacher performance through

student exam scores. This approach is based on meritocratic principles applied to education and a new avenue for achieving accountability in education.

Recent changes in how the public conceptualizes and implements practices of accountability have not been uniform across states. In recent years, the increased amount of meritocratic policies allows for researchers to test theories of both federal and local governmental behavior within the education domain (McLendon et al., 2006). However, not all states have implemented these meritocratic reforms or implemented identical “reform trajectories” (McLendon et al., 2006). Therefore, it begs the question as to which factors came into consideration for certain states to undertake the adoption of meritocratic-oriented approaches. Further, what characteristics within the states account for differences in the policy stances of the states? Put simply, why did states choose to adopt accountability policies based on meritocratic principles when they did?

This dissertation reports the results of an empirical investigation into these questions. I develop and test a theoretical model derived from previous state policy literature looking at state innovation and diffusion. My theoretical model focuses on how variations over time and across state sociopolitical systems influence policy adoption in the education policy domain (McLendon et al., 2006). I use event history analysis (EHA) to examine how certain state groupings and prior accountability policy adoption, as well as economic, organizational, and political characteristics that make up individual states, influenced the adoption of the meritocratic policy of value-added models during the time period of 2011-2016. In the chapters that follow, I present a theoretical framework, six

hypotheses, research design, analytic findings, and the conceptual implications of those findings.

The policy of value-added models was the focus of the diffusion research in this dissertation. Value-added models are an accountability tool used to measure the impact of an individual teacher's effect on student outcomes, in most cases, test scores. This is accomplished by using a variety of measures to predict each student's test score and then comparing these predicted scores to how the students actually scored on the test (Aslantas, 2021). Most states have implemented some form of value-added models for teacher evaluations.

Value-added models build off performance-based policies which were first seen in higher education. Performance-based funding is any funding that is received contingent on achieving specific outcomes. When concerning higher education, performance-based funding was first used by states to ensure higher graduation rates from institutions and tied state funds to the outcome. This ensured compliance with state directives and served to hold institutions accountable.

This study seeks to explain why these tools of accountability diffuse amongst states, specifically when looking at value-added models. Performance-based measures occurred over a thirty-year period, where-as value-added policies diffused quicker over a short six-year time period. I seek to attempt to explain the reasons for the rapid spread of this latter policy.

A number of scholarly studies have been conducted concerning state innovation and diffusion of public policies over the last five decades. Many publications that arose

throughout the 1970s and 1980s gave great insights into the determinants of innovativeness within various policy areas (Berry and Berry, 1991). Further, the central research question concerning state lead innovation was, “What causes a government to adopt a new program or policy?” (Berry and Berry, 1991; Walker, 1969; Downs, 1976; Grupp and Richard, 1975).

Internal versus External determinants

However, a critical weakness in the early state innovation literature is the delineation of internal versus external determinants for why innovation occurred (Berry and Berry, 1990). It was found that internal determinants in models would normally not specify a role for regional influence (Downs, 1976). On the other hand, regional diffusion models assumed that internal state characteristics have zero significant effects (Light, 1978). When both models for regional and internal determinants were examined together within a single study, their analyses were kept separate (Berry and Berry, 1990). The internal determinants model was used for analyses of the determinants of policy innovation, while regional diffusion models looked more at policy emulation or diffusion itself (Canon and Baum, 1981).

Further, neither a regional diffusion model nor an internal determinants model is a an explanation of state innovation in by itself, but instead, a combination of both is needed to best explain why policy diffusion occurred (Berry and Berry, 1990). Both internal and regional influences on a state’s likelihood to adopt a new policy can be predicted based on a theory by Mohr (1969) where he describes the propensity to

innovate as a function of the strength of obstacles against innovation versus the availability of resources for overcoming the same obstacles.

Diffusion of Meritocratic Policies and Performance-Based Funding

Through the recent accountability policies in education, it is assumed that what is being developed are people who have the knowledge and skills to successfully contribute to the nation's economy. This is a departure from the notion that the goal of education was to develop citizens. And that this more recent approach is based on the idea that what people achieve in society is based on merit. If citizens possess the ability and work hard enough, they will accomplish their goals and reap the deserved rewards. The architects of these recent policies apply to education. Additionally, if the appropriate measures of performance can be developed and enticing incentives are offered, teachers will be justly rewarded. If students perform well, teachers will be rewarded. If students perform poorly, teachers will not be rewarded and may even lose their jobs. Either way, the outcome is determined by teachers. They have control over the outcome that they face.

For meritocracy to work as intended, all who are within the system must start on a level playing field. Meritocracy, in the context of public schools, is based on the assumption that student educational performance can be measured using standardized metrics. Thus, the performance of teachers and schools can be measured by how well the students under their direction perform. It is also assumed that teachers and administrators are rational actors who respond to inducements, both positive and negative. Thus, performance-based funding has been offered as the inducement for teachers and schools to improve student performance. This approach to education is based on the assumption

that it is fair to use the same standards for all students and schools. This assumption, however, is problematic because of differing socioeconomic contexts and the fact that not all students begin in the same place.

Additionally, differences in the social relationships both among and within schools reflect the social backgrounds of the student body which will later influence their future economic positions (Bowles and Gintis, 1976). It has been found that working-class schools tend to emphasize behavioral control and rule-following, while schools in higher socio-economic suburbs instead have open systems that foster greater student participation, have less direct supervision, allow for more student electives, and stress internalized standards of control (Bowles and Gintis, 1976).

Differences in the social relationships within schools are consequently reinforced by inequalities in financial resources brought in by the institutions. The minute financial support for the education of those in the lower socio-economic echelons allows more resources to be set aside for the children of those with more outsized roles in the economy (Bowles and Gintis, 1976). Further, it forces upon the teachers and school administrators in the blue collar schools a type of social relationship that closely mirrors that of the factory (Bowles and Gintis, 1976). Financial considerations in poorly supported schools do not allow for smaller, more intimate classes, a multitude elective courses (including STEM), and specialized teachers who are able to personally address the needs of the students (Bowles and Gintis, 1976).

Well-funded schools, which generate large donations through the PTA and other sponsored events, are attended by the children of the elite. They can therefore offer

greater opportunities for the development of student autonomy using independent work and encourage other characteristics that will be required for adequate job performance in the upper levels of the occupational hierarchy (Bowles and Gintis, 1976).

Meritocratic practices of accountability are not a new concept and were first seen in education at the collegiate level. Within the past 40 years, the concept and practice of accountability in U.S. public higher education have undergone major change due to meritocratic policy measures enacted at the state level (McLendon, Heard, and Deaton, 2006). By 2003, 25 states had adopted a performance funding program (Burke and Minassians, 2001). The movement of this new accountability represents a milestone achievement in the changing campus and state relationship (McLendon, Heard, and Deaton, 2006). This relationship holds important implications for the budgeting and governing of schools of higher education, and for the ways in which elected officials decide the value of these institutions (McLendon, Heard, and Deaton, 2006). Not all states have undertaken these reforms, nor have all states sought to undertake identical reform paths. Due to these inconsistencies, questions arise as to which factors drove certain states to take the important step of adopting performance-oriented approaches (McLendon, Heard, and Deaton, 2006).

Having implemented performance-based funding practices in higher education, federal initiatives sought to implement similar meritocratic policies in K-12 in an attempt to effect accountability measures. Because of this, in the last two decades educational researchers have adopted value-added methods to evaluate student, teacher, and school performance at an increasing rate (Papay, 2011). It is by examining test scores over time,

that value-added models appear to isolate the contributions of individual teachers and schools to student achievement (Papay, 2011). Further, various states and districts have begun using these models to identify and reward high-performing teachers (McCaffrey et al., 2003). As value-added models have become increasingly widespread among states, and carry higher stakes, questions concerning both the validity and reliability of their results have grown in importance (Rivkin, 2005).

When considering the overall effectiveness of the meritocratic policies put into place, the research community is divided over the usefulness of value-added models (Rockoff, 2004). Some researchers state that the methodologies used to analyze overall effectiveness do support claims that specific teachers increase student achievement (Tekwe et al., 2004). Thus, compensation and accountability policies based on these estimates would be justified (Sanders, 2000). Other researchers, however, have argued that the many assumptions that underly these models make such claims have little basis and that low stakes uses may be most appropriate (Koretz, 2002).

This system, designed to create an even playing ground for all students and give minimum standards for achievement, has not only failed our students and staff, but created a form of a caste system with those at the bottom rungs rarely being able to propel themselves upward (Sandel, 2021). Additionally, there is disagreement over whether the policies work. It is this disconnection between policy and the push for policy adoption in the face of a lack of evidence supporting their effectiveness which begs the question: what explains the adoption of these policies?

Proposed Research Contribution

The models used in this research further the work of Walker (1969) who analyzed almost 90 different policies that had been adopted prior to 1965. He revealed an adoption pattern that explained the spread of public policies among states in which a regional leader (cue state) adopted a given policy first, followed by neighboring states within the same geographic region. Walker's (1969) findings served to redirect research at that time away from models that focused solely on the internal attributes of states and concentrate research on state-to-state diffusion.

The research in this dissertation also builds on Berry and Berry's (1992) research with the use of event history analysis. In two groundbreaking studies, Berry and Berry applied event history analysis in examining the likelihood that a state would adopt a lottery (1990) or a new tax (1992) within a given year. Berry and Berry's work invigorated the policy research community's interest in understanding the conditions that can lead to a state's propensity to adopt new policies. Additionally, Mintrom (1997) and Wong and Shen (2002) examined the spread of charter school policies. Mooney and Lee (1995) analyzed the adoption and spread of abortion regulation and death penalty laws. Hays (1996) also assessed the determinants of state reform of various civil and criminal statutes.

However, statistical evidence supporting the effects of diffusion are mixed and without consensus among the academic community. Mooney (2001) noted that only 50% of the 24 EHA models reported in studies of state policy diffusion that were published in top political science journals in the 1990s contained positively and statistically significant

coefficients for the diffusion variable. Mooney (2001) also found that the early studies conducted by Berry and Berry (1990 and 1992) contained a bias toward finding a positive diffusion effect. It was found that they failed to use time controls to stabilize the hazard rate.

I develop and test a theoretical model that was used in previous comparative state policy and politics literature. The model focuses on how variations over time and across state sociopolitical systems influence policy adoption in K-12 education. Event history analysis allows for the examination of how certain demographic, economic, organizational, and political characteristics of individual states, that influence the diffusion processes among states (McLendon, Heard, and Deaton, 2006). My research looks at catalysts for the adoption of meritocratic policies during the period of 2011-2016 by using event history analysis and to further explain why states quickly diffused the policy of value-added models for merit pay. If one exists, I should be able to see the diffusion pattern of the policy by states taking cues from border states.

The primary goal of the analyses within this research project is to compare the effects of the variables on the rate of policy adoption. Event history analysis is used as the analytical method (Cox, 1972). Event history analysis was originally found within medical research and used to study the timing of human deaths. It employs the use of longitudinal data to estimate and predict the likelihood of an event (adoption) occurring with time-varying variables (Blossfeld and Rohwer, 2002).

Event history analysis is increasingly being used in educational research as it provides better measures than logistic regression when handling longitudinal data, time-

varying variables, and censoring (Ampaw and Jaeger, 2012). Cross-sectional techniques, for example, logistic regression, assume variable effects remain constant across the event of interest (in this case policy adoption). Also, variables that change over the studied period of time have to be modeled as a different variable giving up degrees of freedom in the estimation (Ampaw and Jaeger, 2012). By using event history analysis, I can overcome these limitations by incorporating variables that change value across time (Ampaw and Jaeger, 2012). In previous education policy literature, event history analysis has been used in the research focused on higher education to model undergraduate student retention and institutional graduations rates (Chen, 2007; Gross, 2008; Ishitani and DesJardins, 2002). Further, I assume that only one outcome can occur (adoption) and that once outcome occurs the subject (state) is no longer at risk (since adoption has occurred, the state can no longer adopt the same policy). The assumption of a single occurrence can be relaxed within the framework set forth (Andersen et al., 1993).

My research builds upon the previously discussed comparative state policy and politics literature and policy emulation. Additionally, the research serves to highlight this fractured system and explain why policies, though glossy in their presentation, continue to be adopted. Additionally, I look at performance-based funding in higher education which served as the testing grounds for further merit-based education iteratives. This is followed by examining our system of high stakes testing couched within common core policies.

Further, I will look at previous policies which value-added pay models are based on. I will further explain the history of teacher-based merit pay within our system. It is by

understanding the history behind the policy, that I fully understand its advent and how it was meant to better our student populations, making them more competitive in the world market. The focus will be on policy adoption and diffusion in order to better understand why these education policies continue to perpetuate.

The research seeks to explain value-added policy adoption in an effort to shed empirical light on the rise and spread of these policies within states. I begin first by developing a conceptual framework based on previous literature and to build understanding of the value-added policy being diffused. Next, I present six hypotheses that emanate from the empirical models that explain value-added policy adoption. I then go on to describe the construction of a longitudinal data set containing indicators of the factors that I believe may have shaped state involvement in value-added policies. Next, I describe the research methods, which use event history analysis. Lastly, I will present empirical findings and discuss the conceptual and analytical implications of this and future work.

CHAPTER 2

Teaching and Meritocracy

Over the past four decades, there has been great concern over the quality of education that students have been receiving in public schools in the United States. Various explanations have been offered for why schools were “failing” beginning with a report under the Reagan administration that was entitled “A Nation at Risk” (The National Commission of Excellence, 1983). This report, along with many others like it were catalysts for why state and local governments chose to adopt innovative education policies in response to what they saw as ineptitudes within our current systems. The education system beginning in the 1980’s assumed the meritocratic notion that people will succeed if they are given the proper incentives to do so. It is by understanding meritocracy within our education system, that I can better understand the driving factors of why the policy diffuses between states.

In this chapter I will discuss socio-economic conditions and how they affect meritocratic education principles. I will then turn to micro-credentials in education and their proliferation. It is by understanding the background of meritocratic policies in education and initiatives such as micro-credentials that we have a full purview of why these policies continue to proliferate.

A Nation at Risk

The report “A Nation at Risk” (The National Commission of Excellence, 1983) laid out in detail concerns the federal government had for education in the US. Additionally, it looked at various studies and pointed to academic underachievement on

national and international scales (The National Commission of Excellence, 1983). A Nation at Risk stated that average SAT scores dropped an average of 50 points in the verbal section and about 40 points in the mathematics section throughout the period of 1963-1980 (The National Commission of Excellence, 1983). Further, nearly forty percent of 17-year-olds tested could not successfully draw inferences from written material, and only one-fifth could write a persuasive essay; while only one-third could solve a mathematics problem that had multiple steps (The National Commission of Excellence, 1983). The study highlighted unfavorable comparisons with students outside the United States and stated that American students never came in first or second and, in comparison with other industrialized nations, were often last (The National Commission of Excellence, 1983).

In response to these and similar problems, the commission made thirty-eight recommendations, divided them into five major categories which included: Content, Standards and Expectations, Time, Teaching, Leadership and Fiscal Support. The recommendations are as follows:

- For content, the commission recommended four years of English, three years of mathematics, three years of science, three years of social studies, and (at least one-half year of computer science for all high school students. The commission also recommends that students work toward being proficient in at least one foreign language beginning in the elementary grades.

- For standards and expectations, the commission warned against grade inflation and recommends that four-year colleges raise admissions standards and standardized tests of achievement such the SAT and ACT at major transition points from one level of schooling to another, especially from high school to college or work.
- For time, the commission recommended that school districts, along with State legislatures change to 7-hour school days, as well as a 200- to 220-day calendar school year.
- For teaching, the commission recommended that salaries for teachers be professionally competitive, market-sensitive, and performance-based. They also recommended that teachers needed demonstrate competence in an academic discipline.
- Lastly, for leadership and fiscal support, the commission pointed out that the Federal government plays an essential role in addressing student needs, especially those labeled gifted and talented, from disadvantaged socioeconomic backgrounds, minority and language minority students, and those under the special education heading. The commission also pointed out that the Federal government also must help ensure compliance with constitutional and civil rights and provide student financial assistance and research and graduate training.

The report pledged additional federal funding to help states meet expectations that were laid out. However, states were not required to adopt the recommendations from the

report. This report and ensuing directives were among the first steps taken by the federal government that attached monetary incentives for compliance with federal standards.

Socio-economic Considerations

Due to the “A Nation at Risk” (The National Commission of Excellence, 1983) report, many states chose to implement versions of the proposed recommendations. States sought to attract individuals that would build up their economic workforce and allow for the state to be competitive on a national scale. This was accomplished, in part, by having a top-ranked school system.

Building on this new competitive system, many schools seek to attract the “best and brightest” students to their programs (Crozier, 2018). This brings prestige to the institution. In turn, donors become attracted to the institution, which allows for the school and students to benefit from the additional funds, either in material ways such as upgraded technology or through human capital in the form of additional staff. This will also attract better trained staff which students will again be able to benefit from. This system of filtering students based on prior achievement gaps sets up a system hierarchy that students later down the chain can never benefit from (Sandel, 2020), creating larger achievement, and ultimately, wealth gaps with public K-12 schools competing for students through school choice initiatives.

It has been argued that education should be nationalist in its scope, committed to the national interest, looking to increasing national wealth, productivity of its citizenry and competitiveness, and to provide educational justice and equality of opportunity (Brown and Tannock, 2009). Meritocracy has been used to establish a link between

individual effort, which is then mediated through education (Souto-Otero, 2010). The working class and poor within the US are finding it increasingly difficult to access top level education, training, and employment (Brown and Tannock, 2009).

McNamee and Miller (2004) people assume that an education based on meritocratic principles rewards the bright, hard workers regardless of socioeconomic status, but most Americans also think that is how the system does work in practice. In the US, educational attainment and social mobility based on meritocratic principles is positively correlated with their socioeconomic status. Certain groups of children may be estimated as less worthy of advanced educational opportunities, even when their prior options have been limited through academic access (Meroe, 2014).

In contrast, children from families with higher socio-economic status enter school with greater cultural understandings often attend better equipped schools and therefore have access to a wider range of academic supports and opportunities (Meroe, 2014). Socio-economic backgrounds often overestimate the costs and underestimate the benefits of education and training, and because of this, attach a low value to education (Souto-Otero, 2010). In other words, those in low socio-economic brackets tend to not see the benefit in higher education in their long-term success. While proponents of the current educational system regard systemic discrimination as a fallacy. Social and economic inequality that exists is thought to be legitimate because it reflects the natural results of poor personal choices (Sunquist, 2002).

Meritocratic arguments can help coverup significant inequalities when it comes to access to high-caliber educational opportunities (Hostetler, Sengupta, and Hollett, 2018).

As a result, rather than improving the educational opportunities and quality for low-income students, studies have shown that low-income parents are less able to take advantage of choice options that are available (Hostetler, Sengupta, and Hollett, 2018). Low-income parents' choices are constrained by two factors: information and cost (Sunquist, 2002). This rhetoric further blames individual students and their cultural or family backgrounds for poor academic performances (Mills and Keddie, 2012). This takes the focus away from the fact that larger school and societal factors can negatively affect student achievement (Mills and Keddie, 2012). This is what Littler (2018) calls a socially erosive ethic that is competitively self-interested.

A meritocratic approach to education ignores that not all students start off on the same footing or are given the same advantages even in the public school system. In the United States, private and public educational institutions from K-12 through college, are expected to function as breeding grounds and identifiers of high talent and effort (Meroe, 2014). However, student investment is plagued with unequal distributions of funding and pupil spending which begins at the state level.

If I look at two states, for example, I find that a rich state, like Connecticut, spends on an average of \$18,000 per pupil per year, while a poorer state, like Mississippi, spends far less with average only \$8,000 per pupil per year (Markovits, 2020). This uneven distribution of spending can be seen even more acutely when looking at specific towns within rich versus poor states and the ability for a school's local PTA to funnel even more funds into already thriving school populations. The Scarsdale Union Free School District in New York has a median household income of \$238,000 and spent

about \$27,000 per student. In contrast, the Barbourville Independent School District in Kentucky with a median household income of \$16,000 only spends the minimum \$8,000 per student (Markovits, 2020). These trends tend to reflect gaps, surprisingly not between the middle and poor classes, but instead between the middle and elite classes in society (Markovits, 2020). School districts with higher household incomes have more options for pulling additional resources when needed. These districts can raise taxes on real estate or hold fundraisers which can bring in as much as \$400,000 in one night at a local Chicago elementary school in an elite part of the city (Markovits, 2020). The range of per pupil spending is aligned with income brackets: a poor student in public K-12 schools will have about \$8,000 in spending, a middle-class student about \$12,000, an upper middle-class student \$18,000, an elite student \$27,000, and an elite private school topping off at \$75,000 (Markovits, 2020). This is unique to the United States with it being only one of three nations in which public schools that serve more affluent students spend more per pupil and have lower student/teacher ratios than public schools that serve less affluent students (Markovits, 2020).

For any resemblance to a true meritocratic system to work within schools, funding allotments must be redistributed to all public schools within the system (Crozier, 2018). This would level the playing ground. Though some schools may distribute the funds differently, the overall amount of funds that a school receives would be stagnant and disallow parents from elite families to give an advantage their children in the public-school setting.

However, these additional funds received either through fundraising or additional per pupil spending serve to better prepare students of elite parents for careers past their K-12 schooling. The funds go to extracurricular enrichment such as science, math, coding and robotics camps, all of which are highly sought-after skill sets. These programs tend to complement, not just duplicate what is being done in the classroom (Markovits, 2020). It has been found that those who participate in these activities are 70 percent more likely to go on to college than students who are only occasionally involved in school clubs and 400 percent more likely than students who do not participate or have exposure to such activities (Markovits, 2020).

Because elite parents, who have benefitted from the meritocratic system themselves, they have the means to pay for further academic tutoring and high-stakes test prep. The income/achievement gap for SATs as a result is substantial. Students from families earning over \$200,000 per year (the top 5 percent of earners) score 388 points higher than students from families earning less than \$20,000 per year (the bottom 20 percent of earners), and middle class students only score 135 points higher on average than their poorer counterparts (Markovits, 2020). Academic attainment and performance can be deemed as the seeking and accumulation of academic and cultural human capital that is used to compete and trade later on the employment market (Meroe, 2014). This demonstrates that social justice does not require the elimination of all inequalities, but instead requires that the “differentially deserving” be “differentially rewarded” (Swift and Marshall, 1997).

The result was that as of 2016, 58 percent of Americans raised in households from the top quartile of the income distribution earned bachelor's degrees by age 24, with only 41 percent from the next quartile, 20 percent from the third quartile, and 11 percent from the bottom quartile (Markovits, 2020). At this level of achievement, children from low socioeconomic backgrounds are less likely than children from high socio-economic backgrounds to continue with furthering their education beyond high school (Souto-Otero, 2010).

To try to combat the effects of opportunity, some critical educators increase grades as an act of resistance to meritocracy. Further, some educators advocate for non-competitive forms of education, while other educators try to work within a meritocratic system to make meritocracy more equitable (Crozier, 2018).

As a public good, education in the United States is held as the principal means by which the youngest citizens are prepared to participate in our democratic society and eventually have the skills to negotiate the labor market (Meroe, 2014). The scarcity of high-quality educational resources and credentials can undermine the stores of human capital. This threatens the perpetuation of a sufficiently informed electorate and a competent labor force on a domestic and global scale (Meroe, 2014).

Students within low socio-economic strata often find that their years spent in “inferior” public schools did not prepare them for success. In addition, they are able to recognize that they lacked both the economic and social resources necessary to compete with the other privileged students (Sunquist, 2002). Although given the opportunity to engage, they were not provided with the tools necessary to succeed (Sunquist, 2002).

In his book, “How Meritocracy Works,” Markovits (2020) describes education as the labor market’s preferred sorting mechanism and goes further to state that the economic return from schools in terms of the labor market has become astronomical. This amounts to the better the schooling, the better the job and paycheck that follows. This creates a meritocratic loop where elite students, coming from advantaged backgrounds, are set up to go to and succeed within the nation’s top universities through a system of grooming set forth by the previous generation and worn like a badge by the current one. It has been found that the median college graduate with a BA (Bachelor of Arts), will earn more over their lifetime than 93 percent of workers without a high school degree and 86 percent of workers with only a high school degree (Markovits, 2020). Should that graduate go on and further their education with a professional graduate degree, they will over their lifetime make more than 99 percent of those without a high school degree, 98 percent more than those possessing only a high school degree, and 83 percent more than workers with only a BA (Markovits, 2020). Further, this has produced a minimal standard of education for the majority of students while reserving the best educational opportunities for the economic (and meritocratic) elite (Meroe, 2014).

In her 2020 study, Reay spoke with K-12 teachers and questioned their view specifically on meritocracy concerning standardized testing. The most common issue highlighted by the participants was the lack of a level playing field, with more than half feeling that the existing system was flawed because of the different socio-economic backgrounds of students. Interestingly, most of the participants (17 of 30) felt that the state had to be actively involved in levelling the playing field by redistributing

educational resources to disadvantaged groups. Less privileged students should first be given an advantage to have the same starting point as others. This was because she felt that it was not fair for these students to work harder just because of their less advantageous backgrounds (Reay, 2020).

Within the same study, in contrast, others were concerned about rewarding other forms of talent (Reay, 2020). They felt that the non-academically inclined students with different talents would not be rewarded or appreciated in our current education system that places great emphasis on standardized tests. The teachers' perspectives on meritocracy, including their understandings of the costs and restrictions of the concept as well as their inability to extend their resistance by creating alternative systems, appeared to be mostly canceled out by the discourse about the meritocratic system (Reay, 2020). The teachers' responses largely reflected their acceptance of these state-dominated initiatives, although a significant proportion of them asserted that the system resulted in the mimicking and legitimation of the interests of elites (Bourdieu and Passeron, 1977).

Ultimately, meritocratic arguments are as vital to justifying the perpetuation of failure and mediocrity for the majority as they are to rewarding the privileged and worthy few (Meroe, 2014). It is because of this, according to many scholars, education should be viewed as a non-merit-based aspect of society (McNamee and Miller, 2004). However, meritocratic principles, such as being in charge of your own educational trajectory, have become a part of public education through government policies. These policies including performance-based funding which can be found in higher education and federal policies such as No Child Left Behind, along with Race to the Top. The focus on individual

achievement and mobility is problematic because it can have a significant negative impact on relations within the larger community since not everyone starts off the same (Ho, 2021). Children connect educational success and failure to their own senses of individual self-worth and future prospects (Reay, 2020).

Micro-Credentials

With the understanding that all students do not begin on an even playing ground, states turned their attention on how best to address the issue of low-performing schools. One mechanism that was used in an attempt to attract better qualified teachers to lower socio-economic schools was micro-credentials. Micro-credentials are assessments of learning and demonstrations of skill and competency (Berry and Byrd, 2019). For example, micro-credentials can be used to support teachers in their development of skills needed to teach English language learners. Another benefit of micro-credential training is that it is often delivered online, making it more accessible to rural educators and those who need to engage in professional learning in the evening, on weekends, or during school breaks (Berry and Byrd, 2019). Additionally, states who use micro-credentials demonstrate their support for educators' abilities to understand what their own professional learning needs may be. Micro-credentials can be used to better understand what teachers know and can do, as well as a way to develop personalized systems of professional learning for any teacher or administrator (Berry and Byrd, 2019).

Micro-credentials are attractive to teachers because the credentials create a way for teachers to “free themselves from a one-size-fits-all” approach to professional development (Berry and Byrd, 2019). Further, micro-credentials are content-focused,

job-embedded, and they incorporate active learning (Berry and Byrd, 2019). Micro-credentials are a way for schools, districts, and states to create systems of professional learning that allow for teachers take control of their own learning, demonstrate evidence of impact, and demonstrate what they have accomplished (Hunt et al., 2020).

Additionally, about three in four US teachers already are engaged in some form of informal professional development or learning (Berry, 2017). While working with micro-credentials can be either formal or informal, they are a way for teachers to be recognized for the skills they have (LaMagna, 2017). As stated earlier, an online platform provides new ways for school districts to recognize which teachers and administrators are trained in teaching certain skills and/or leading improvement efforts as well as providing highly personalized professional learning to help individuals with specific needs (Olcott, 2022).

As an innovation, micro-credentials are still in an early stage. However, a growing number of states and districts, as well as the nation's largest teachers' union, are experimenting with them (Berry and Byrd, 2019). A growing number of districts are also piloting micro-credentials. Teachers who earn micro-credentials may receive a salary increase, a one-time stipend, continuing education credit hours and graduate credit (Hunt et al., 2020).

Even with this additional training through micro-credentials, the ailments that often plague our public schools are whittled down to the easiest and most convenient common denominator: the classroom teacher. Micro-credentials are a tool used to improve teacher quality, improve student outcomes, and meet the demand for teachers in specific subjects and locations. Additionally, districts and schools are implementing these

policies to increase teacher recruitment and retention (Podgursky, 2006; Reed, Rueben, and Barbour, 2006). Several states have recently enacted targeted teacher recruitment legislation in one or more of the following areas to attract teachers to high-need schools and subjects such as English and mathematics. These recruitment measures include convening work groups and collecting teacher supply and demand data. The measures give state and local education agencies the flexibility to design their own targeted teacher recruitment strategies which include four economic incentives: (1) Creating career pathways and grow-your-own programs for high school students. (2) Altering teacher preparation and licensure requirements. (3) Providing financial incentives such as scholarships/grants, loan forgiveness and hiring bonuses/salary increases. (4) Providing incentives for retired teachers to return to the profession (Aragon, 2019).

The four economic incentives listed in the prior paragraph may be ongoing or only enacted for a limited amount of time. The incentives may be awarded on receipt of a particular certification, or as a bonus to a teacher's salary for already holding a given credential (Strunk and Zeehandelaar, 2011). As an example, districts may choose to pay teachers more if they hold a particular credential in a high-needs subject area or if they choose to work in a target school (Murphy and DeArmond, 2003). Nationally, approximately 10 percent of school districts offer salary enhancements to teachers in high-need subject areas such as math, science, and special education (Murphy and DeArmond, 2003), with states like Massachusetts, Rhode Island, and Oklahoma having implemented bonuses for math and science teachers (Johnson, 2006).

Further, school districts may incentivize teaching in schools that are considered “hard-to-staff” because the geographic location or demographic makeup of the school may seem undesirable (Strunk and Zeehandelaar, 2011). The advent of incentivized wages for teachers in urban areas was first enacted in the late 1950’s (Conant, 1961). Additionally, large mostly urban districts like Los Angeles have been using teacher incentivization in order to attract high-caliber teachers to hard-to-staff schools (United Teachers of Los Angeles, 2009). Rural districts who also struggle to recruit and retain teachers due to social, cultural, and professional isolation and lower salaries, and are also using economic incentives to attract teachers (Collins, 1999).

However, there is some evidence that districts use targeted economic incentives to help mitigate their shortage needs and use fiscal incentive policies like teacher incentives as a means of accomplishing this goal (Kolbe and Strunk, 2010; Prince, 2003). Teacher incentivization is an attempt to address lower socio-economic school needs in hard-to-staff areas and bring a collective awareness of and commitment to public education from early childhood through secondary levels (Baum and McPherson, 2011).

Summary

In this chapter I looked at merit practices interact with socio-economic conditions of a state and further micro-credentials. Though meritocratic principles can allow for advancement, not everyone begins at the same point within the education system which skews how quickly individuals can advance. To help remedy this, micro-credentials are used to attract educators and allow them to have control over their professional

development and showcase the skills they have chosen to develop. Further these credentials allow for schools to incentivize educators to teach at hard-to-staff schools.

CHAPTER 3

Performance-Based Funding in Higher Education

The purpose of this chapter is to look at the first appearance of accountability practices in education. These principles first appeared in public education at the college/university level in the form of performance-based funding. Institutions of higher education in the United States have traditionally been funded based on a combination of prior appropriated amounts and on state budgets, along with the number of enrolled students (Layzell, 1999). The new approach is a departure from the traditional approach to funding. Then you can state that these policies may have served as the precursor to the adoption of value-added models in K-12 education.

In the following chapter I will be going briefly over the background of how performance-based funding came to a policy proliferated in higher education. Next, performance-based funding will be viewed through the lens of principal-agent theory and explain how institutions being agents had to answer to their principal being the state. The chapter will then turn to the how the private business sector further influences this policy. Lastly, the two waves of the policy with PBF 1.0 followed years later by PBF 2.0 and their driving factors will be examined.

In this study, I focus on how states that adopted performance-based funding also adopted the value-added model of teacher evaluations. This creates a system of accountability which was tested in higher education and trickled down to K-12. With higher education being the catalyst for further meritocratic practices, such as value-added models, I should see a trickle-down effect of policy diffusion.

Performance-Based Funding Background

In the 1960's and 70's, accountability in education often referred to the design of state policies (McLendon et al., 2006). The central question policymakers tried to answer was which activities and functions of institutions of higher education (such as academic programs, budgets, and tuition setting) should be dictated by the state and which should be left to the discretion of the leaders on campuses (McLendon et al., 2006). The accountability focus, therefore, was on the design of governance systems that made it possible to both effectively and efficiently regulate the flow of state resources to a campus (Berdahl, 1971).

This approach was promoted by the Reagan administration and its call to arms through the publication, "A Nation at Risk" (1983) which was briefly discussed in the previous chapter. The report stated that the educational foundations of our society were being eroded by a rising tide of mediocrity that threatened our very future as a Nation and a people (U.S. Department of Education, 1983).

Additionally, the report called for elected officials, educators, parents, and students to reform the public school system, which was deemed in urgent need of improvement (U.S. Department of Education, 1983). Numerous statistics were used to demonstrate the dire state of the current educational system and the inadequate quality of an American education. The main takeaway was that the nation was at risk, and the report further expressed concern that the once unchallenged superiority in commerce, industry, science, and technological innovation is being overtaken by US competitors (U.S. Department of Education, 1983).

The report suggested that there were specific things we, as a country, could do to reverse our current course and to bring American education to safety. This included several key points:

1. More rigor: In the educational sense, “rigor” means to make something more difficult. The report suggested that states increase the rigor of their curricula, challenging students to do better and graduate smarter (A Nation at Risk: Summary and Effects on Education, 2014).

2. New standards: Some states had already written standards, which stated that students would have to accomplish certain things to move through the grades and eventually graduate high school. The report was concerned that even those states that did have standards set the bar so low that the students really did not have to do much of anything in order to pass the course/grade (A Nation at Risk: Summary and Effects on Education, 2014).

3. Teacher preparation and pay: The report specifically noted that some teacher training programs did not provide adequate training for teachers. In addition, teachers were paid far less than other fields that required a college degree. As a result, there was a shortage of qualified teachers, especially in areas like science and math. Among other teacher-related recommendations, the report suggested that teacher training programs require more of prospective teachers and that teacher pay be tied directly to student achievement (A Nation at Risk: Summary and Effects on Education, 2014).

What was proposed as an answer to the posed problem was a system of performance funding, which differed from traditional enrollment-based funding (Burke,

2002). Performance-based funding entails allocating state dollars using a precise formula that ties appropriations to student performance measures at the institution level (Hiler, 2018). State officials believed in accountability, not solely in higher education but across governmental activities and funding (Dougherty and Natow, 2015).

These accountability measures called for a refocusing of attention on outcomes of campus activities, rather than the traditional focus on inputs alone (Zumeta, 1998). Key performance indicators were introduced in higher education and determined a certain amount of each institution's state appropriation (Jones, 2004). The goal was to increase quality, measured as efficiency and effectiveness, accountability, and self-regulation through effective top-down leadership (Dougherty and Natow, 2019).

Commonly used metrics with performance-based funding include the number of degrees completed, graduation rates, job placement rates, and milestone accomplishments such as year-to-year retention rates (Dougherty and Reddy, 2013). In scrutinizing outcomes, state policymakers have sought to influence institutional behavior for the purpose of improving institutional performance (McLendon et al., 2006). This approach links state appropriations to school performance on metrics like graduation rates (Burke and Minassians, 2003). Under these policies, the relationship between performance and funding is predetermined. If an institution meets a specified "performance target" it receives a specific amount or percentage of state funding (Burke and Minassians, 2003).

Reasons behind the shift in policy to a form of accountability through performance-based funding include globalization which brought pressure from the business community for campuses to maximize productivity and efficiency (Burke,

2005). There were recent shifts in the theory and practice of public-sector way of governing which placed valued on decentralization, entrepreneurship, and a greater emphasis on markets, competitiveness, and the measurement of performance (Burke, 2005). There were also extreme financial pressures placed on state governments throughout the 1990's, which served to call into question funding for higher education (Burke and Serban, 1997). Further, there were changes in state political leaders, with newer ones being more likely to implement policies that increased accountability in higher education (Burke, 2005). Lastly, there was a failure of the earlier "voluntary assessment" to demonstrate to elected officials that institutions of higher education were capable, without coercion, of meeting growing accountability demands (Burke and Serban, 1997).

Principal/Agent Theory

Performance-based funding also calls upon principal-agent theory, where the state is the principal and the agents are the educational institutions. The principal-agent model has been used successfully to understand a wide range of relationships between political institutions and other actors. It has also been used as a means of examining interactions between the states and the federal government (Fowler, 2020), the states and counties (Rasmussen, 2000), and the state and educational leaders at the local level (Seashore and Robinson, 2012). For instance, Enders, de Boer, and Weyer (2013) used a principal-agent model to examine the autonomy and performance of reforms within higher education and the ways in which the reforms were enacted.

Principal-agent theory examines the degree to which principals can exert control over the behavior of agents to stop the latter from shirking or following their own agenda. Principals use agents to accomplish their goals because the principal lacks the necessary resources or because the costs of achieving such goals are prohibitively high. The basis for any principal-agent relationship derives from the condition of mutual advantage (Eggertsson, 1990; Maltzman, 1997). In other words, principals use agents, and likewise, agents choose to be used because there exists some form of mutual payoff. The existence of such mutual advantages provides the principal with the ability to impose their goals on the agent, while reinforcing a relationship of dependency by attaching the agent's behavior to those goals (Eisner and Meier, 1990). In the case of performance-based funding, the mutual payoff would be that the educational institutions (agents) receive state funds and the state (principal) sees a high yield return on their investment in the form of increased graduation rates.

It is often found that agents are not committed to the goals of their principals and often shirk their duties. This leads to carrot and stick type policies that can erode the morale and commitment of employees who are motivated intrinsically (Andersen and Pallesen, 2008). For instance, a faculty member at a leading university described the situation of being responsible for graduation rates as “demoralizing” (Dougherty et al., 2016). The assumption of this type of funding is that educators are not working hard enough, and that they are not willing to change, nor improving (Dougherty et al., 2016). Similarly, the focus on the use of monetary incentives, as well as performance monitoring as the main motivators is used to move institutions in the directions desired by elected

leaders within states. Institutions in higher education are conceptualized as having interests that can run counter to those of the state leader's desires, material incentives and monitoring via performance metrics are seen as necessary to align both motives and to keep institutions in line (Broucker and DeWit, 2015). This is accomplished through monetary incentives (Dougherty et al., 2016). The state's ultimate goals are higher graduation rates, more research dollars coming in, and additional efficiency in the higher education system. If the institutions implement things that align with state goals and policies, then the state will monetarily support the effort. This boils down to a simple financial incentive model (Dougherty et al., 2016).

Business and Performance-Based Funding

In some states, business leaders also directly lobbied state governments to implement performance-based funding to secure and even design the system (Dougherty and Natow, 2019). In the state of Indiana, the leaders within the Chamber of Commerce testified in favor of performance-based funding and directly sent messages to fellow Chamber members supporting it (Dougherty and Natow, 2015). However, business also had an indirect effect, by shaping the context of ideas in which policymakers operated and by affecting the incentives government officials faced (Dougherty and Natow, 2019).

Performance-based funding is consistent with the privatization movement. By promoting the idea of improving efficiency through privatizing numerous governmental services and making government agencies use business-like approaches, colleges would be subjected to market-like constraint. Business made it more likely that state officials would view performance-based funding as a desirable and politically attractive policy for higher

education (Dougherty and Natow, 2019). State officials recognize that securing “business confidence” is vital to getting businesses to invest in the venture and further, to generate jobs and tax revenues, all of which politically benefit government officials and the state as a whole (Block, 1987).

Support for performance-based funding stems from two main motives. First, higher education supports performance-based funding because leaders of institutions share the belief of state government officials that it will improve higher education outcomes (Ness et al., 2015). This can be seen when a state community college leader in Florida stated, “I think the philosophy of performance funding is just a basic concept that should be embraced by everybody. I honestly believe that whether higher education institutions get more money or not, that they should be out there trying to get the best performance that they possibly can” (Dougherty and Natow, page 71, 2015). The second motivating factor is that many leaders of higher education institutions believe that performance-based funding provides them with a new way to legitimate themselves in the eyes of elected government officials and the public.

PBF 1.0 and 2.0

In 1979, Tennessee became the first state to adopt a performance-based funding policy in which a portion of state appropriations to public higher education was directly tied to student outcomes. Adoption and implementation took place over two waves with other states following Tennessee’s lead in the following decades.

The first wave of policy adoptions that took place in the 1990s (over a decade after first implementation) and is referred to in the literature as PBF 1.0 (Ortagus et al.,

2020). Initially, PBF programs consisted of pools of money that were used to reward colleges who met performance targets. As such, these funds were in addition to the state appropriation to a school. (Dougherty and Reddy, 2013). Many performance-based funding states suspended or discontinued the policies during tight fiscal times. This is because PBF 1.0 programs which offered a bonus or supplement were difficult to maintain when state governments were reducing expenditures (Dougherty et al., 2012). In addition to tight state budgets brought on by fluctuating economic conditions, other factors that attributed to the collapse of PBF 1.0 were an insufficient buy-in among key stakeholders and a loss of political support for these programs (Burke and Modarresi, 2000). From 1979 to 2010, 12 of 25 states that had adopted PBF shifted to lower-stakes policies, such as performance budgeting or performance reporting (Miller and Morpew, 2017). A performance budget reflects the input of resources and the output of services. The goal is to identify and score relative performance based on goal attainment for specified outcomes, in this case graduation rates. Performance reporting defines and measures overall success of the policies implemented. It provides an overview of how the entity is performing by collecting specific performance data, analyzing it, and provide suggestions to help in making decisions. The results were that states had not given up on accountability measures, they instead transitioned to tools that were less expensive.

The second wave of performance-based funding adoptions took place in the late 2000's and is referred to as PBF 2.0. The second wave included two key differences from the PBF 1.0 policy. First, instead of having bonus funds, PBF 2.0 placed a portion of base state appropriations in limbo to coerce institutional change and to make PBF recession-

proof (as the first one was not) (Ortagus et al., 2020). In this new wave of policy, institutions of higher education could potentially see a decline in state financial support if institutional performance did not meet specified goals. Second, states tied a larger portion of funding to outcome measures, with several states allocating more than 10% of higher education appropriations to institutional performance outcomes (Friedel et al., 2013). PBF 2.0 focused specifically on college attainment and completion, and used that often included metrics such as student retention and institutional degree productivity (Dougherty and Natow, 2015). PBF 2.0's increase in adoption coincided with the economic crisis and constrained state budgets for higher education (Miller and Morpew, 2017). Newly elected policymakers were giving increased attention to public financing of higher education, making it a salient agenda item (Dougherty and Natow, 2015). Additionally, President Obama issued a call to increase the number of college graduates and explored a number of policy options to reach this goal by 2020, which further served to fuel the diffusion of PBF 2.0 policies (Miller and Morpew, 2017).

For these new PBF policies, the amount of funding provided per full-time equivalent student in public universities is an indicator of a state's investment in the higher education of its students (Science and Engineering State Indicators, 2021). Funds allocated towards higher education are expected to produce a return in the form of human capital. Unlike previous methods of funding which based the allocation of state tax dollars on student enrollment, performance-based funding in higher education ties state funds given to public colleges and universities to educational outcomes. This serves to incentivize institutions and improve student outcomes while holding institutions

accountable for results. Institutions of higher education are heavily reliant on state funding, and therefore are likely to be responsive to performance-based funding systems since an important source of their revenue is at stake (Ortagus et al., 2020).

The Effect of Performance-Based Funding

Performance-based funding was seen as the solution to a funding mechanism that failed to reward colleges for actually graduating students. The shift to performance-based funding put a laser focus on college completion rates. It was found that only 59% of those who started college intent on earning a bachelor's degree did so within 6 years (Undergraduate Retention Rates, 2017). When concerning associate degrees and certificate programs, only a staggering 29% of students graduated within 3 years (Hiler, 2018). As previously stated, education beyond the high school level has become increasingly necessary to earn a living wage. Those states that had enrollment-based funding understood that funding increased college entry but did not translate to college completion. Under this previous approach to funding, institutions of higher education faced no consequences should students failed or simply withdrew from school. For this reason, states with lower rates of postsecondary educational attainment were more likely to pursue performance funding to try to increase graduation rates (Li, 2017).

The motivation of the policy is to reduce government funding at the same time as secure greater efficiency from various public agencies, including higher education institutions. In order to accomplish this, top government officials try to create pseudo-markets for those institutions by using market-like incentives to encourage competition for both clients and public support (Hood, 2006).

Institutions with more financial resources such as research universities in performance-based funding adopting states were more likely to use evidence to build internal models and identify where to focus their efforts to get a larger share of the funding (Ortagus et al., 2020). In contrast, institutions with fewer resources were less likely to come up with how best to allocate funds. Instead, these latter incentives rely on state directives to stay in compliance and ensure state funds would continue to be allocated to their institution (Thorton and Friedel, 2016).

Further, the nature of performance-based funding changes incentive structures for public colleges and universities (Snyder, 2015). However, states that have adopted these policies have demonstrated few spending changes in terms of funding at state or institutional levels. Results from various analyses of performance-based funding indicate that states do not necessarily give more money to colleges and universities based on outcome measures (Rabovsky, 2012).

Research on the implementation of performance-based funding finds extensive evidence of resistance and gaming on the part of higher education institutions (Rabovsky, 2012). US institutions frequently resist the demands of state governments that they pursue certain outcomes standards (Li, 2017). The institutions attempt to disguise their resistance by using such devices as reducing course and program demands on students so they can pass at higher rates (Rhodes and Sporn, 2002). This allows for institutions to be able to meet government officials' desire for higher graduation numbers without having to make many changes in their established organizational practices (Dougherty et al., 2016).

Studies of performance-based funding have also pointed to the fact that a major obstacle to the ability of higher education institutions to meet the demands of the policy is the actual make-up of their student bodies (Schulze-Cleven and Olsen, 2017). Colleges and universities find it difficult to produce as many graduates or to find jobs for their graduates as demanded if many of the students entering those colleges are not well prepared academically or have face other disadvantages (McKinney and Hagedorn, 2019). Higher education institutions are also delayed by the fact that the outcomes desired by state policy makers often take a time to occur. The median four-year college student takes six years to graduate so it takes a long time for institutions to demonstrate the effects that their political principals are expecting (Kelchen, 2018).

In the case of performance-based funding, the goals laid out by policy makers have been characterized as improved instructional, student-support, and research-administration efforts. Further, the goal is to have improved student outcomes, degree completion and faculty research productivity (de Boer et al., 2015). The most common campus-level changes include improved advising and counselling services, a retooling of developmental education, and streamlining transfer processes (Jenkins et al., 2009).

In the case of graduation numbers, the effects of performance-based funding were found to be weak or even nonexistent (Bell et al., 2018). Rutherford and Rabovsky (2014) reported that early performance-based programs were not correlated with improved student performance at public 4-year institutions. In their analysis of twelve US studies, Bell et al. (2018) found that the average impact of degree completion is not distinguishable from zero. Similarly, in a study of the performance-based funding

programs implemented in Ohio and Tennessee, Hillman et al. (2018) found no effect on baccalaureate and associate degree “production.” What was found was that there was a significant impact on degree production but argued that it came at the expense of greater effort to raise baccalaureate and associate degree completion (Hillman et al., 2018). This comes at the expense of considerable compliance costs, involving additional management and administration for colleges to carry out those policies (Kivisto, 2007; Gray et al., 2001).

Further, it has been found that proposed reforms may result in reductions in the quality of services as institutions scramble to keep up with heightened competition, as well as declining revenues, by reducing the quality of the services they provide (Kivisto, 2007). In order to improve graduation rates, institutions have resorted to weakening academic demands through the practice of inflating grades and reducing the number of difficult courses required for graduation (Dougherty et al., 2016).

To add to an already declining educational rigor, performance-based funding has been seen as reducing the intake of less advantaged students, constructing of institutional missions, stratification of institutions, growing stratification of the academic labor force, and a decrease in motivation on the part of higher education personnel (Dougherty and Natow, 2019). Institutions have become more restrictive in admissions which can be seen by their raising their academic requirements, focusing their recruitment activities on secondary schools with more advantaged students, and targeting financial aid awards towards students via “merit” as versus those with need-based aid (Dougherty et al., 2016). This is due to merit students being more likely than need-based students to

graduate. Additional studies of performance-based funding have identified the danger that this practice can increase inequality within the institution by financially penalizing poorer performing institutions, which further hinders their ability to improve their performance (Dougherty and Natow, 2019). For instance, in their study of community colleges in the state of Texas, McKinney and Hagedorn (2017) found that Black students, Hispanic students, adult students, and part-time students secured significantly less performance-based funding for community colleges relative to more advantaged students. This further demonstrates that these policies may contribute to a lack of resources at colleges and universities that serve large numbers of minority students and widens already-existing resource gaps within institutions of higher education.

Ortagus et al. (2020), in their study of 52 empirical articles published between 1998 and 2020, examined the intended and unintended consequences of performance-based funding policies in American higher education. Their study shows that most of the peer-reviewed academic research on the effectiveness of performance-based funding in the United States typically reveals null or modest effects of the policies on the institutional outcomes primarily targeted by the policies.

Thus, there is little evidence that that performance-based funding in higher education produces better student retention and graduation rates (Dougherty and Natow, 2019). Rather, performance-based funding produces many unintended impacts that are often ignored by institutions in order to stay in compliance with set policy (Naidoo, 2011). This, instead, points to compliance costs and drops in output quality but does nothing to anticipate other unintended consequences. These consequences include

reduced admission to less advantaged students, narrowing of institutional missions, rising inequality within the institutions, a growing stratification of the academic labor force, and lowered motivation of higher education personnel (Lane, 2008).

The potential benefits of performance-based funding can only come to fruition if efforts are made within institutions to systematically address the many obstacles that are encountered in realizing its intended impacts, as well as the unintended impacts (Morginson, 2009). The design of performance-based funding must be rooted in research designs and theories that rely on political science and social and organizational psychology (Pusser, 2008). This provides a better understanding of policy design and implementation than the economic formulations on which it currently relies (Pusser, 2008). Internal and external organizational factors, in addition to environmental ones are typically ignored by those who are designing these policies. They are therefore ignoring the varied needs, interests, goals, and “power resources” of multiple stakeholders (Davis et al., 1997). Also being ignored is how higher educational institutions are shaped by their communities, workforce, leadership, culture, organizational resources and processes, and the economy (Davis et al., 1997).

In an attempt to better understand the diffusion of performance based funding policies despite evidence that questions the policy’s effectiveness, Miller and Morpew (2017) looked at the policy through the lens of framing theory. Framing theory helps to identify competing schemas for understanding and communicating the same data to advance political objectives (Rein and Schon, 1996). Further, frames define policy problems and propose policy solutions by simplifying and making sense of information

understood by policymakers and the broader public (Miller and Morphew, 2017). They found that by framing information, political actors create narratives that advance their interests. A narrative describes the source policy problem source and offers a solution that potentially resolves the issue. Narratives additionally allow actors to restrict and control the decision making of public policymaking (Miller and Morphew, 2017).

Political actors incite action through what is known as motivational framing which is openly enlisting policymakers' support to enact change and influence the political process (Schulock and Snyder, 2013). Frames that rely upon rhetoric from broad public narratives are more likely to be picked up by fellow policymakers and later be enacted (Rein and Schon, 1996).

By defining policy problems that are resolvable through government intervention, state leaders are able to have sway over the policy agenda (Stone, 1989). Accordingly, Miller and Morphew (2017) used state leaders as agenda-setters because they work to generate and control political debate. What they found was that political actors submit a problem as the subject of policy decision. The leaders then influence the policy stream to control which policy ideas gain momentum and which policy ideas are then excluded (Kingdon, 1984). Therefore, policymakers believe there is a policy problem, the political actor's identified mechanism causes the policy problem, and the political actor's proposed solution will resolve the problem (Miller and Morphew, 2017).

The use of a framing of a problem to influence the policy agenda set the stage and made a salient case for the acceptance and implementation of performance-based funding policies within states. Further, these agenda setting organizations encourage

policymakers to adopt performance-based funding through motivational framing by associating the policies with already significant calls for accountability when facing constrained state budgets (Greenstein, 2015). This allows for the political buy-in of the narrative being put forth. Additionally, performance-based funding was equated with progressive thinking and innovation (Miller and Morpew, 2017). The message being broadcast was loud and it was clear, the states that chose to adopt performance-based funding policies were on the cutting edge of higher education policy and all other states choosing not to adopt the policies were falling behind. By coupling performance-based funding with progressive policy suggests that policy detractors are backward, out of touch, and interested in sustaining an older system (Davies, 2014).

Summary

This chapter has looked at the advent of performance-based funding and how the policy tied state funds to student graduation rates in higher education. The policy, having purse strings attached to state appropriations, mandated institutional compliance, though there was not a roadmap to follow in order to achieve the policy's goals. There have been two waves of policy implementation, with the first beginning with Tennessee in 1979 and a later wave in the 1990's. Both waves sought to enhance oversight of decision making by college administrators in light of declining graduation rates. The policy also had unintended effects on the type of students who were being admitted to colleges so that compliance could be met. This made higher education less attainable to many citizens, but still became a blueprint for further meritocratic education policies.

CHAPTER 4

Value-Added Models

Education policy over the past decade has shifted from emphasizing whole school performance to also focusing on individual teachers (Goldhaber, 2015). Policy makers had long sought to find a way to quantify the impact teachers had on students' learning within their classrooms. What were adopted were value-added models which are an objective measure that does not rely on human interpretation of teacher practices (Goldhaber, 2015). Instead, it is a system in which teachers are evaluated relative to one another rather than relative to an absolute standard (Goldhaber, 2015). This creates a distribution in which teachers can be ranked (Goldhaber, 2015). However, it is also a more novel way of evaluating teachers that many in the school system, at the time of adoption, were vastly unfamiliar with the method (Goldhaber, 2015). Yet, over the past decade, merit pay and value-added models for teacher evaluations have become important topics within many states and school districts in the last decade. Even though these models became popular, they were not without their critics. It is because of this that by only comparing average test scores, we have a very narrow understanding about the teacher and school's performance.

In this chapter I will examine federal policies under the space race that set the stage for federal involvement in state education curricula and accountability measures. I will then look at current teacher evaluation systems and how they have incorporated value-added models. Next, I will focus specifically on the adoption of value-added models by states and what economic principles led to policy adoption. From this point, I

will explore what set the stage for the proliferation of value-added policies, including prior federally mandated policies. Next, I will consider how student test scores affect teacher evaluations, which led to the expansion of value-added models. I will end the chapter with previous studies that have been conducted using value-added principles and teacher performance.

Space Race, the beginning of federally funded curriculum changes

Prior to the proliferation of the report “A Nation at Risk” and accountability measures such as value-added models, Congress in 1958 passed the National Defense Education Act. With its passing, the National Science Foundation began pouring money into curriculum development in K-12 schools (Hoff, 2020). This was to ensure our competitiveness of STEM (science, technology, engineering, and mathematics) type careers and to ensure science and math curricula were at the forefront of American education. The National Science Foundation, in total, spent \$500 million on curricula and teacher development (Hoff, 2020). In the mid 1970’s, there were federally funded curricula programs totaling fifteen for elementary schools, twelve for middle schools, and thirty-four for high schools (Hoff, 2020). The cross-content materials were meant to engage student and bridge subjects. For example, within the high school biology curriculum, there was a series of films, including ones on viruses, ecology, and social biology (Hoff, 2020).

Through the use of federal funds, the National Science Foundation had the goal of teaching the basic principles found in math and science through a hands-on approach and the notion of “learning by doing” (Hoff, 2020). It was solidifying the math and science

base, that the knowledge could be applied in the future through a variety of STEM careers (Review of Recommendations for NSF Project Management Reform, 2017).

This resulted in changes in the fundamentals seen in K-12. Students in K-6 started learning about sets and number theories based on groups other than 10 (Wooten, 1965). Additionally, high school students learning about physics now studied how waves travel in detailed and intricate experiments such as using tidal pools (Atkin, 1980).

Some materials developed such as those created and introduced by NASA (National Aeronautics and Space Administration) scientists were well-received at schools and widely used with the remnants still being found in classrooms today (Blake and Della-Piana, 2001). Other materials, like “New Math” introduced in the 1960s, along with a 5th grade anthropology program called Man: A Course of Study, fell flat and eventually led to Congress’ restriction of the National Science Foundation curriculum in the 1970’s (Review of Recommendations for NSF Project Management Reform, 2017).

What the National Science Foundation demonstrated was that the federal government can use grants to entice schools to adopt curricula (Review of Recommendations for NSF Project Management Reform, 2017). It is by tying federal funds to initiatives that states can become coerced into policy adoption that may or may not be beneficial. When considering the National Science Foundation funds, principles established by being at the top in the space race would be achieved because the education system set the groundwork in K-12. This groundwork served to develop students who were already exposed to STEM principles and could therefore further develop technologies needed to “win” the race at an accelerated pace.

The grants which drove this change, though they have changed in name, still exist in today's educational system. They are now known as Curriculum Innovation Proposals (CIP) and have diffused up to higher education. Curriculum Innovation Proposals are designed to support new, innovation curriculum development and the development of STEM higher education in areas bolstered by NASA (Curriculum development grants, 2020). The grants look favorably upon proposals of STEM curriculum that can be shared with other schools. Further, curriculum development is required to be in alignment with one of the following four NASA Mission Directorates: Aeronautics Research Mission Directorate, Human Exploration and Operations Mission Directorate, Science Mission Directorate, and Space Technologies Directorate (Curriculum development grants, 2020).

These grants were the first to set the stage for further federal involvement in setting state educational curricula. By tying federal funds to wanted educational changes, we saw state level compliance with federal directives. The federal government quickly learned that by incentivizing wanted policy changes, it was able to coerce compliance.

Teacher Evaluation Systems

In the last decade there have been significant shifts to teacher evaluation policies across the nation, and in most states the evaluation of teaching is no longer conducted just by school administrators (Berry and Byrd, 2019). Recent federal policies, particularly those promoted under Race to the Top, have promoted more classroom observations by trained reviewers as well as the use of student test score data in judging individual teachers (Berry and Byrd, 2019). However, the Education Commission of States has recently pointed out that states, districts, and schools often struggle to create, and further,

to implement the type of evaluation system that meaningfully differentiates teacher performance, and provide teachers with opportunities for specific support and development (Olcott, 2022).

Value-added models are intended to be an objective measure that looks at the amount of “value” that a teacher “adds” to, or in some cases detracts from, student learning and achievement from one school year to the next (Beardsley and Holloway, 2019). Value-added models are a statistical tool meant to measure the observable relationship between a teacher’s instruction and how that instruction aides in the student’s learning and achievement over time (Beardley and Holloway, 2019). This is done by measuring growth in achievement from one year in comparison to the next using student large-scale test scores, while controlling for students prior testing histories (Beardley and Holloway, 2019). Additionally, it controls for student-level demographics (such as race, poverty, English language proficiency, special education status), other student-level variables (such as attendance, suspension, retention), and other variables such as class size, average prior achievement (Beardley and Holloway, 2019).

Theoretically, policymakers design state and local tests to assess specific content that standards deem important (Papay, 2011). Most would hope that student progress on one achievement tested in one domain would equate to progress on additional tests. In other words, students deemed excellent readers on one test should also do well on another test within the same reading domain (Papay, 2011). Additionally, if a school district wanted to reward teachers based on their ability to improve their students’ reading scores, those within the district would have the understanding that the best teachers would raise

student scores not only on the state test, but also on additional reading assessments as well (Papay, 2011). Should there be conflicting scores from two similar tests, this would raise important questions about the validity of any teacher effectiveness estimates (Papay, 2011). The gains of achievement, needed to show student growth under the models, were measured through “high stakes testing” a cumulative standardized test given at the end of each school year beginning in third grade (Fiscal Year 2005 Budget Summary, 2004).

Student learning is the primary goal of any school, and value-added scores are deemed the most logical, cost-effective method for identifying teachers’ contribution to learning (Gordon et al., 2006). Studies have found that overall teacher effects explained 11% of the variation in student test score gains (Nye, Konstantopoulos, and Hedges, 2004).

With looming teacher shortages, states also needed to devise a plan to attract new teacher talent, while retaining exemplary teachers already within their system. Incentivization programs housed within value-added models would reward teachers for improved student test scores and serve to attract, as well as retain, high quality teachers. Studies have shown that the design of contracts, in this case teacher contracts, can have important effects on who opts to work in specific careers. Evidence from the private sector suggests that the choice of pay structure can affect both the productivity of individuals in their job and who opts to pursue employment at the establishment (Lazear and Shaw, 2007). The main idea that is more productive individuals may be more likely to enter and remain in a profession that rewards their talents, and less productive employees would be more likely to leave. Contrarily, if people are not rewarded for their

output of talents in a profession, then it would reason that the more talented people would be expected to leave if their talents are rewarded more elsewhere (Goldhaber, 2015).

Value-added models are also associated with long-term schooling with the focus being on student performance that would translate to increased numbers of students continuing on to higher education. This would also benefit employees in potential earnings, as well as nonacademic outcomes such as reduction in teen pregnancy (Chetty, Friedman, and Rockoff, 2014).

When considering economics-based principles, there seems to be a dominant view that complicated matters, like education, can be understood in both quantifiable and measurably objective ways (Beardsley and Holloway, 2019). Teachers and teacher quality are being seen in terms of a “production function model” where teachers produce a product that is taken in or “consumed” by students and is therefore measurable in terms of how much “value” a teacher “adds” to the production (Beardsley and Holloway, 2019). Given this new formulation, and the position of education based in economics, that society is made up of individuals that will make rational decisions given the right conditions, teachers are in a position as people who make good or bad decisions that result in adding or detracting value from student learning (Lazear, 1999).

Florida was the first state to adopt a value-added-based performance plan and implement the program in its counties through the 1990’s and early 2000’s. Additionally, Oklahoma and Colorado have both recently adopted pay-for-performance plans that include a value-added component; 16 other states have voluntarily implemented district-initiated programs (Hill, Kapitula, and Umland, 2011). This creates an atmosphere that

easily allows for states to adopt a policy of value-added models based on student achievement on standardized tests (Hill, Kapitula, and Umland, 2011). According to Also, by 2008, eight states used systems that based teacher pay in part on performance; however, not all schools in these states participated in these plans (Ballotpedia, 2019).

Furthering these diffusion of accountability measures, Race to the Top was implemented in 2010 and encouraged states to adopt value-added models for teachers with a total of 4.3 billion dollars in incentives should states comply with the mandate (Teacher Merit Pay, 2019). In 2012, U.S. Secretary of Education Arne Duncan stated that merit pay, as a part of teacher accountability under value-added models, was the U.S. Department of Education's "highest priority" (Teacher Merit Pay, 2019). In the plan, value-added models would hold teachers responsible for student testing outcomes and reward teacher success, which is quantified by student gains on standardized tests, through merit pay initiatives. These measures were due, in part, to the Great Recession and states needing to defray costs as much as possible (Paige et al., 2019). As of today, 44 states (88%) have adopted and have at least begun implementing value-added models to evaluate and, in many cases, make important decisions about teachers. This decision can include items such as teacher tenure, merit pay, as well as teacher termination should they not meet the minimum student growth requirements (Beardsley and Holloway, 2019).

Additional Policies that Lead to Value Added Models

Over the last two decades, new education policies have shaped our public schools. These policies, which target student achievement gaps, are designed to ensure that all

students are given equal opportunity to learn and that classroom teachers are held responsible directly for those outcomes regardless of outlying factors. The policies, which are supposed to streamline student objectives, are broken down by grade year and subject. For a student to show mastery and a teacher to demonstrate success, students need to achieve proficiency on the set objectives through various forms of high-stakes testing.

In the following section, I will provide an overview of three distinct policies: No Child Left Behind under the Bush Administration, the Teacher Incentive Fund, and Race to the Top under the Obama Administration. All three built upon one another and created the system of value-added model used for teacher evaluations that we currently see.

A key component of the No Child Left Behind policy is high-stakes testing. Though No Child Left Behind was enacted in 2002, the stage was being set in the mid 1990's. At that time, the testing industry began to experience a shift towards standards based high-stakes assessment as a necessary component of standards-based reforms (Jorgensen and Hoffman, 2003). The shift from test publishers' proprietary products to various versions of state-owned tests began and test publishers found themselves in the position of developing content, manufacturing tests, and scoring and reporting on assessments that were not originally part of their inventory (Jorgensen and Hoffman, 2003). Further, from 1994 to 2000, most states had instituted content standards that students should achieve throughout the year based on grade level. Performance standards, collection of longitudinal data, and use of secure test forms each year were used to track student achievement and inform instruction. Schools also recognized the importance of

quick test results reporting and thus by the end of the 1990s, states had become well informed about the design and implementation of large-scale assessment programs (Jorgensen and Hoffman, 2003).

On January 8, 2002, President George W. Bush signed into law the *No Child Left Behind Act of 2001* (Jorgensen and Hoffman, 2003). New mandates centered around accountability, local control, parental involvement, and funding (Jorgensen and Hoffman, 2003). If children are not learning, based on standardized test results, the law requires that the school find out why. If the schools are not meeting standards, options for improvement and help will be made available by the U.S. Department of Education (Jorgensen and Hoffman, 2003). It was also required that schools must ensure that all students learn the essential skills and knowledge defined by the state using grade-level standards and benchmarks (Jorgensen and Hoffman, 2003). No Child Left Behind requires that states have assessment systems that track the achievement of all students against a common set instructional standards that are implemented throughout the state (Jorgensen and Hoffman, 2003).

Further, states were required to assess third- through eighth-grade students annually in both the subjects of reading and mathematics of which the tests must be based on state standards that are challenging to the students and can therefore show mastery and growth (Jorgensen and Hoffman, 2003). The results of the standardized tests and subsequently, school grades, were made public so that anyone within the public can track the performance of any school in the US (Jorgensen and Hoffman, 2003). Improvement among lower socio-economic children must be demonstrated under the Adequate Yearly

Progress provisions detailed within No Child Left Behind (Jorgensen and Hoffman, 2003).

Lastly, schools unable to demonstrate adequate progress will be provided with assistance from the state and federal government and may be subject to possible corrective action should they still not demonstrate growth (Jorgensen and Hoffman, 2003). All states requesting federal funds will be required to submit plans that describe their achievement standards, aligned assessments, reporting procedures, and accountability systems (Jorgensen and Hoffman, 2003).

In exchange for greater accountability on the part of the states, the No Child Left Behind regulations would provide states with flexibility and control over how they use federal funds distributed as part of No Child Left Behind (Jorgensen and Hoffman, 2003). Schools, however, would be encouraged to use funds towards teacher retention, professional development, and technology training as they see best suited. States would also be given greater flexibility and control over their programs for English language learners (Jorgensen and Hoffman, 2003).

Schools that failed to make adequate yearly progress for two consecutive years were identified for “school improvement” (NCLB, New America, 2021). Schools were then required to draft an improvement plan which devoted at least 10 percent of federal funds provided under Title I to teacher professional development (NCLB and funding controversy, 2021). Further, No Child Left Behind required that school districts spend up to 20 percent of their federal Title I funds to provide immediate assistance to children

attending schools identified as in need of improvement, through either public school choice or supplemental educational services (NCLB and funding controversy, 2021).

No Child Left Behind created much controversy on how to implement its far-reaching policies and was deemed an “underfunded mandate”; however, states were not required to participate (NCLB and funding controversy, 2021). All requirements were a condition of federal funds that would be disbursed to the states through a block grant that was a component of No Child Left Behind. While a state may not be able to provide all recommended curricula without federal education funding, it could choose to opt out of No Child Left Behind and the requirements it included (NCLB and funding controversy, 2021). The law’s requirements only applied to those states that voluntarily elected to participate.

When No Child Left Behind was passed, the Department of Education assured states that federal funding would cover a significant portion of the costs associated with the law’s new requirements (NCLB and funding controversy, 2021). However, the costs of test administration, data collection, and school improvement reforms had been understated, and thus states argued that the federal government did not adequately support the reforms. It has been found that the estimate of additional expenditure needed to implement No Child Left Behind that was provided by the Government Accountability Office vastly underestimated. A Government Accountability Office report vastly underestimated the additional implementation expenditures states would see from 2002-2008, which varied depending on the type of test states used (NCLB and funding controversy, 2021).

States that implemented No Child Left Behind policies feared teachers would forego traditional teaching methods and begin to teach directly to the test standards in order to achieve the mandated gains. In a study conducted by Dee and Jacob (2010), it was found that in states where the exam results were used to hold teachers accountable, instruction was reported to have shifted to testing criteria which left out socio-emotional learning and resulted in behavioral problems on the part of the students (Dee and Jacob, 2010).

After seeing little to no improvement in student achievement, Congress began a new initiative to incentivize teachers and have better staff buy-in when it came to high-stakes results, especially in lower socio-economic areas where it was believed the largest student gains could be made. As a result, in 2006, Congress created the Teacher Incentive Fund as a means of supporting performance-based compensation for teachers and principals in high-needs schools (Evaluation of the teacher incentive fund, 2021). Grants totaling \$1.8 billion were sent to school districts that were willing to implement the four program components: measures of teacher/principal effectiveness involving student achievement gains; pay-for-performance bonuses; additional pay opportunities for extra responsibilities; and professional development aiming to improve performance (Evaluation of the teacher incentive fund, 2021).

The Teacher Incentive Fund program provides funding for projects that develop and implement performance-based teacher and principal initiatives (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021). Performance-based compensation systems must demonstrate gains in

student academic achievement, as well as have classroom evaluations conducted multiple times throughout each school year and provides educators with incentives to take on additional responsibilities and leadership roles (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021). The purpose of the program is to support the use of performance-based compensation for teachers, as well as to utilize other human capital strategies that enhance the status quo in an attempt to increase students' access to effective and highly-effective educators in lower socio-economic schools (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021).

Recipients of these federal funds use them to develop and fund teacher leadership positions and incentivize teachers to teach in lower socio-economic schools (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021). Projects that have been funded include: teacher career pathway programs that serve to diversify roles for teachers; teacher career pathways that recognize, develop, and reward excellent teaching; incentives for effective teachers who take on leadership roles within their schools; incentives that attract, support, reward, and retain the most effective teachers and administrators at high-need, target schools; rigorous, ongoing professional development training for teacher and principals, leadership roles for teachers specifically at school turnaround; and the creation of new salary structures based on teacher effectiveness (Teacher incentive fund, 2021).

A four-year study looked at the impacts of the Teacher Incentive Fund and bonuses to pay. The study measured the impact of pay-for-performance bonuses as part

of a comprehensive compensation system. The schools examined were to fully implement their performance-based compensation system. Control schools were to implement the same performance-based compensation system with one exception: pay-for-performance bonus component was replaced with a one percent bonus paid to all educators regardless of performance (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021). Schools that were studied were randomly assigned to one of the two groups. The report provided implementation and impact information and found that implementation was similar across the four years, with most districts implementing at least 3 of the 4 required components for teachers (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021). In a subset of 10 districts participating in the random assignment study, educators' understanding of performance measures did in fact improve over time, with many teachers being unaware that they were eligible for a bonus. Further, their understanding did not improve after the second year of implementation. Teachers were found to also underestimate the maximum amount they could earn. The teacher incentive fund and pay-for-performance bonus policy ultimately was found to have small, positive impacts on students' reading and math achievement (National Center for Education Evaluation and Regional Assistance: Evaluation of the Teacher Incentive Fund, 2021).

Broadening the impact of the Teacher Incentive Fund, the Obama administration introduced Race to the Top legislation in 2009. It is a competitive fund designed to promote school improvement on both a state and local school district level (What is race

to the top and how will it benefit public schools?, 2021). At its advent, \$4.35 billion was pledged in what was deemed the “largest ever federal investment in education reform” (What is race to the top and how will it benefit public schools?, page 11, 2021). State governments were asked to submit plans for education reform in order to be granted a portion of the funding set aside for the program (What is race to the top and how will it benefit public schools?, 2021).

In its press release, the White House laid out four key areas for reform in order to qualify for Race to the Top funding. These areas included, an improvement of assessments and more rigorous standards for schools, the turn-around of failing schools through increased emphasis and resources, increased support that allows teachers and staff to be more effective, and better methods for tracking progress of both students and teachers (i.e., value-added models).

Each state was given the opportunity to submit a plan directly to the U.S. Department of Education. In the end, 46 of 50 states submitted plans, though not all were accepted, but many states have received funding from the program to complete the initiative (What is race to the top and how will it benefit public schools?, 2021). However, regardless of the acceptance of a state’s plans, all states experienced a surge in the adoption of various education policies (Howell et al., 2020). Further, it was found that this surge was not an extension of past policy trends. Legislators from all states involved reported that Race to the Top directly affected policy deliberations within their states (Howell et al., 2020). The patterns of the policy adoptions along with legislator responses

corresponded with states' experiences in the Race to the Top competitions (McGuinn, 2012).

President Obama fashioned Race to the Top as a coercive measure aimed at state-level policymaking that aligned with his education objectives on college readiness, which was the ultimate goal of the policy (McGuinn, 2012). The legislation created new data systems, ways to track teacher effectiveness, and attempted to specifically target low-performing schools. Obama intended to incentivize excellence and stimulate reform and launch a race to the top in America's public schools (Howell et al., 2020).

Race to the Top was comprised of three phases of the competition. Phase 1 and Phase 2 included specific education-policy priorities on which each applicant would be specifically evaluated; states were asked to describe their current education status and then outline their future goals in meeting the four criteria in each of these categories. State applicants were scored across six categories that assessed the potential for a state's proposed policy changes would accomplish the goals of Race to the Top (Howell et al., 2020).

Further, to assist states in writing their applications, they were offered technical assistance workshops, webinars, and training materials (McGuinn, 2012). In addition to the workshops, organizations such as the National Council on Teacher Quality were asked to publish reports intended to help states maximize their likelihood of winning an award (McGuinn, 2012).

In all, 40 states submitted applications to Phase 1 of the competition (Howell et al., 2020) and both states' past accomplishments and their stated commitments to adopt

new policies informed the scores they received and hence their chances of winning federal funding (Howell et al., 2020). Finalists and winners were announced in March 2010, with phase 1 winners being Tennessee and Delaware. The two states were subsequently awarded roughly \$500 million and \$120 million, respectively, which amounted to 10 percent and 5.7 percent of the two respective states budgets for K–12 education for a single year (Howell et al., 2020). Thirty-five states submitted applications during Phase 2 of the competition which took place in June 2010. Nine winners who included Florida, New York, Georgia, North Carolina, Ohio, Maryland, Massachusetts, Hawaii, and Rhode Island were each awarded prizes between \$75 million and \$700 million in Phase 2 (Howell et al., 2020).

After phase 2 the funds set aside for the competition had been depleted. At the behest of the president, Congress allotted funds to support a third phase. A higher percentage of participating states won in Phase 3; however, the amounts of these grants were considerably smaller and ranged from \$17 million to \$43 million (Howell et al., 2020).

States that were awarded grants under Race to the Top were subject to a monitoring process that consisted of annual performance reports, accountability protocols, and actual site visits (Howell et al., 2020). After receiving an award letter, a state could immediately withdraw up to 12.5 percent of its overall award (Howell et al., 2020). The remaining balance of funds were only available to states after they received and approved by the state’s participating local education agencies, at which point a state’s drawdown of funds depended upon its ability to meet the specific goals and timelines

outlined (Howell et al., 2020). Like No Child Left Behind, Race to the Top was focused on only two learning outcomes, math and reading, which leaves all the social sciences (including economics, history, and the arts) and all of the sciences in the background (Onosko, 2011).

Further, under Race to the Top, as much as 50% of teachers' yearly evaluations is based on student test scores in math and language arts under value-added models (Onosko, 2011). Teachers who fail to improve student test scores at the projected rate for three consecutive years can then be fired or transferred from their position. Further, underperforming schools over a three-year period will see their administrators removed as well (Onosko, 2011).

A report conducted in 2016 by the Institute for Education Science (the research branch of the US Department of Education), looked at the effectiveness of the Race to the Top initiative. It was found that differences in student achievement between Race to the Top states and other states may be due to other factors and not the program (Klein, 2016). Race to the Top states differed from other states before receiving the grants, in addition to other reforms which may also have affected student achievement (Klein, 2016).

In the end, the report concluded, it is difficult to discern whether Race to the Top had any long-term impacts on state policy (Klein, 2016). There were differences between states that got the grants and those that did not, however, other factors could explain those differences since some states were already implementing policies outlined by Race to the Top even before the grants were given out (Klein, 2016).

What was found was that the original Race to the Top states were more likely to adopt policies and practices encouraged by the program such as rigorous standards and tests. This creates conditions for charter school success, and improving educator effectiveness (Klein, 2016). However, there were not significant differences between states that got the grants and those that did not receive awards when it came to creating data systems to measure student achievement and build state capacity (Klein, 2016). The major finding to note, was that states that got the smaller grants were more likely than non-Race to the Top states to adopt the program's policies in one area: teacher effectiveness in the form of value-added models (Klein, 2016).

With little gains in student achievement having been attained, it is surprising that protocols developed under Race to the Tops are still prevalent within our schools two presidential administrations later. In *Catching Up or Leading the Way: American Education in the Age of Globalization*, Zhao (2009) argues that America is at a crossroads and warns against the nationalization of education. Zhao (2009) expresses it will undermine the diversity of thought, creativity, individual talent, and expertise that previously made the American education system unique. Zhao (2009) further warns against a national reform agenda focused on test scores in only two subject areas (Language Arts and Math) and recommends a return to the tradition of local control. He states it will provide more diverse talents in students, rather than standardized laborers; more individuals with creative skills rather than good test-takers; and more individuals who will become entrepreneurs rather than simple, obedient employees (Zhao, 2009).

Assessments of Value-Added Models

Value-added models use student growth scores compiled at the teacher or classroom level to yield teacher-level value-added estimates (Paige et al., 2019). Under these new evaluation systems, potential teachers who believe they will be effective and rewarded for their effectiveness are more likely to see teaching as a desirable profession since there is now a monetary incentive (Goldhaber, 2015). Traditionally, teaching has been a very predictable occupation when it comes to pay and job security (Goldhaber, 2015). The use of value added represents a significant departure from this regime, and the increased risks associated with teaching could result in teaching being perceived as a less desirable occupation if the overall compensation level does not rise enough to offset perceived job security risks (Goldhaber, 2015).

Another point of contention is how to construct groups: whether teachers should be compared to others in the district, others in the same school, others in the same grade within school, or some combination of the above (Hill et al., 2011). In practice, however, it is hard to effectively isolate the effects of school, teacher, and grade (McCaffrey et al., 2003).

A number of concerns have arisen about using value-added models for teacher evaluation and accountability purposes include methodological concerns such as reliability, validity, bias; logistical concerns such as fairness, transparency, usability; and consequential concerns such as high-stakes testing attached to outcomes, and effects at the school level such as culture (Beardsley and Holloway, 2019). Recent research has found that teacher value-added scores frequently vary based on the specific student

assessment tool and the subtests used in constructing these tools (Lockwood, McCaffrey, Hamilton, et al., 2007). Hill et al. (2011) was critical of a reliance of test scores only. In their study, Hill et al. (2011) demonstrated that value-added scores do show convergent and, to a lesser degree, discriminant validity, since other scores carry a “signal” about the quality of classroom instruction and thus may be valuable and relatively inexpensive tools. Because of this, it is suggested that districts use value-added scores in combination with high-quality, discriminating observational systems or as a trigger for such observational systems. Hill et al. (2011) concluded that value-added scores alone are not enough to identify teachers for reward, remediation, or termination and further suggested that high value-added teachers did not necessarily have the best instruction. The data further suggested that teachers who should be the target of “remediation-type” interventions, would not be identified by value-added scores (Hill et al., 2011).

Most states, however, have chosen to adopt value-added models. This is due to purported positives that can come from the models such as information about teacher effectiveness could lead to direct action in terms of the type of professional development that teachers receive (Goldhaber, 2015). The identification of which teachers are effective might also encourage informal learning opportunities (Goldhaber, 2015). Hanushek (2009), for example, estimates that annually dismissing 5% to 10% of the least effective teachers in each school and replacing them with teachers of “average” effectiveness will boost the test performance of US students by about 50% of a standard deviation.

If prospective teachers and school districts have been informed about what makes for an effective teacher from the outset, the benefits of value-added models will rise.

Contrarily, if the models create disruptive effects in schools, such as mass teacher departures, or if value-added measures are not accurate enough to reflect true teacher contributions, the policies could harm student achievement (Goldhaber, 2015). A recent study (Ronfeldt, Loeb, and Wyckoff, 2012) shows there is a “disruptive effect” of teacher turnover. Additionally, it was found that attending a school where there is high teacher turnover appears to have a negative effect on student test scores in both math and English language arts (Ronfeldt, Loeb, and Wyckoff, 2012).

With a system of value-added models becoming increasingly widespread there should be additional research regarding the positive effect of these models on student performance. This, however, is not the case. What studies have found, including one conducted by RAND in New York City, is that three major components have not been met: Merit pay did not improve student achievement at any grade level, it did not affect teachers reported attitudes, perceptions, or behaviors, and many teachers reported that while the bonus was desirable, the program did not change their teaching practices (Evaluating the Effectiveness of Teacher Pay-for-Performance, 2019). What researchers did find was that merit pay tied to high stakes tests produced high teacher turnover, a growing number of inexperienced teachers filling classrooms, and more teaching to the test (McMurray, 2019). Additionally, it should not be assumed that a value-added score is a sign of teacher quality or effectiveness because these scores also represent a bias due to student selection, the effect of other resources on student achievement, and some measurement error (Amour-Garb, 2009).

To further ensure the best teachers are in the classrooms, credentials developed by each state attempt to ensure quality instruction by requiring teachers to obtain certain types of preparation prior to entering the classroom or, in some cases, to have certain characteristics such as on-the-job experience (Amour-Garb, 2009). Contrarily, performance measures should be based on observations of a teacher's instruction within the classroom by supervisors, peers, and/or external evaluators (Amour-Garb, 2009).

However, it has been found that value-added models are reliable within groups such as the same school or district (Paige et al., 2019). In these circumstances, value-added estimates of teacher effectiveness are mostly consistent from one year to the next, regardless of the type of students and subject areas (Paige et al., 2019). When considering value-added models, researchers have found that five issues need to be addressed: reliability, validity, bias, transparency, and fairness (Amrein-Beardsley et al., 2016). Reliability looks at how teachers are classified as effective one year and yet might be classified as ineffective the next, or vice versa (Paige et al., 2019). Given the stability of teachers' levels of effectiveness, these swings should not be seen within the data. Research has documented serious concerns when considering the reliability of value-added models. Teachers classified as effective one year might have a 25–59% chance of being classified as “ineffective” the next year, or vice versa (Paige et al., 2019).

Validity is the degree to which evidence and theory support the interpretations of test scores (Paige et al., 2019). It is found that there is no evidence that teacher value-added models are correlated with at least one other concurrent measure of teacher effectiveness (Amrein-Beardsley et al., 2016), such as a supervisor's observational

assessment or student survey-based assessment. Bias pertains to the validity of the inferences that stakeholders make from test-based scores (Paige et al., 2019). When it comes to bias, evidence suggests that teacher value-added estimates systematically differ because of the varying demographic characteristics of students nonrandomly assigned to their classrooms (Amrein-Beardsley et al., 2016). Because schools do not randomly assign teachers to grades, these variables are not controlled in order to mitigate bias (Amrein-Beardsley et al., 2016).

Transparency is defined as the extent to which something is both accessible and easily understandable (Paige et al., 2019). Previous research has found that teachers and administrators do not seem to comprehend the models being used to evaluate them, which would erode the extent to which they might use their value-added estimates to improve instruction or initiate reforms (Amrein-Beardsley et al., 2016).

Fairness is the impartiality of the way test scores are interpreted (Paige et al., 2019). Value-added models are, for the most part, only applicable to teachers who instruct in areas that are subjected to standardized tests (Paige et al., 2019). Current research suggests that predominantly it is teachers of math, reading, and language arts who are being held accountable using these systems. This leaves about 70% of all public-school teachers that are not counted and puts the weight of responsibility for both school and individual student achievement on the shoulder of only 30% of all teachers (Amrein-Beardsley et al., 2016). The ineligible teachers typically teach children in early childhood and high school grades as well as in “noncore” subject areas including social studies, science, art, music, and physical education (Amrein-Beardsley et al., 2016).

In a study conducted in a California school district, after having applied an economic-based analytical framework to a set of data, Hanushek (1970) argued that years of experience and education were not related to higher student achievement. Further, teacher effects did not explain Mexican American students' achievement of outcomes, and teacher effects did have any impact on White students' achievement outcomes, regardless of the student's socioeconomic status (Hanushek, 1970). Hanushek (1979), bolstered by the findings, called for additional "econometric instruments" to develop a model that could measure the amount of "value" that a teacher added or took away from student learning.

In addition, in his 2011 study, Papay found that using different achievement tests produces substantially different estimates of individual teacher effectiveness. The variation in teacher value-added estimates that is found from using different outcomes exceeded the variation that was introduced by implementing different model specifications (Papay, 2011). Although the tests can vary in their content, item format, scaling, and sample of students, none of these factors appear to explain the differences in teacher effects (Papay, 2011). The results he found suggested that test timing and inconsistency, like measurement error, play a much greater role (Papay, 2011). Particularly interesting was the finding that the timing of the test alone may produce substantial variation in teacher productivity estimates across outcome measures (Papay, 2011). This raises important questions for teacher accountability policies. Other choices about model specification depend on how districts plan to use these estimates. For example, some models automatically take a teacher's estimated effect, which is

essentially using the teacher's estimated stability of teacher value-added estimates as a "fixed effect" (Papay, 2011). Other models "shrink" teacher estimates back toward the mean, with the amount of reduction depending on the available information concerning teacher performance including the sample size and variability of student test scores (Papay, 2011). Within these models used to examine accumulated teacher effectiveness, extreme results can occur both from true differences in teacher performance and from sampling variation (Papay, 2011).

In addition to these issues, research suggests that three years of teacher-level data are needed to make the most accurate value-added estimates (Amrein-Beardsley et al., 2016). Further, some states' tenure and due process laws allow districts to terminate or untenure teachers with only one or two consecutively unsatisfactory value-added scores (Amrein-Beardsley et al., 2016).

In addition, teachers of gifted students often have lower ratings because high-achieving kids test high on standardized tests and often make smaller gains (Duffrin, 2011). Eric Hanushek, of Stanford University, supports the use of value-added measures for formal evaluation, although he concluded that the scores do not show enough differences among most teachers who fall in the middle range (Duffrin, 2011).

This creates a system where teachers are concerned for their value-added scores and could choose careers at schools that would provide the best advantage for achieving the needed scores. Because of this, within our current education system, there is little dispute that disadvantaged students are more likely to have the least qualified teachers (Clotfelter, Ladd, and Vigdor, 2006). Schools with higher socio-economic status tend to

attract a higher caliber of teachers since they routinely turn out students who achieve higher scores on high stakes tests. This creates a cyclical system where teachers who want to achieve highly effective status on their value-added models and are rewarded with tenure, will work at schools where this goal is easier to achieve and does not put their livelihood at risk. This results in schools in lower socio-economic areas having lower ranking teachers according to our current evaluation system.

The performance measure used in value-added models needs to give teachers accurate information about their productivity (Brehm et al., 2017). Given that schools have small class sizes and there is a need to base performance pay on one year of testing outcomes, it may not be feasible to obtain precise enough value-added estimates in combination with individual incentives to impact teacher effort, including student achievement (Brehm et al., 2017). Further, it was found that teachers are unresponsive to many incentives since in many cases, they may already be putting forth a maximum amount of effort, or they may not know how to increase student achievement (Brehm et al., 2017). In studies conducted on teacher responses and motivation within value-added models, there seemed to be a lot of noise within the data. Noise is defined as the random irregularity in any real-life data. Additionally, an increase in noise reduces the odds of winning an award given there is likely to be no change in effort for those above the threshold.

There are clear disadvantages to working in lower socio-economic schools, some teachers and unions have turned to the courts to combat the perceived systemic failures. Several court cases that challenged the authority or jurisdiction of the state Department of

Education to enact evaluation regulations or laws that use value-added models have appeared (Paige et al., 2019).

Teachers and their unions argue that the use of value-added models leads to teacher isolation where teachers would “close their classroom door” and revert to working alone instead of working with their team (Paige et al., 2019). Additionally, they argued that using of high-stakes testing causes teachers to leave the profession and avoid high-needs schools and seek safer assignments where they can avoid the risk of low scores (Paige et al., 2019). In the end, courts have sided with state governments and found that there was a legitimate interest to increase student academic performance by improving the quality of instructional, administrative, and supervisory services in the public schools (Paige et al., 2019). This was considered the “rational basis test” (Paige et al., 2019). While courts may uphold the use of value-added models under a rational basis test, courts were wary regarding use of value-added models when making significant decisions with respect to teacher employment status (Paige et al., 2019).

Given the above issues, in her essay on value-added models, Amour-Garb (2009) suggested more plausible policy options for ensuring teacher quality which would benefit both the teachers and students. These include the use of teacher value-added models as one part of the teacher tenure decision and the use of overall evaluations conducted at the school level as a complimentary piece of evidence. The use of these evaluations is one part of the compensation decision and would give teacher value-added information to school principals without directly attaching high stakes (Amour-Garb, 2009). School-level accountability would also focus on school value-added models, but would put

pressure on principals to improve, giving them an incentive to use teacher value-added information in any meaningful way (Amour-Garb, 2009).

Adding to earlier information, individual value-added ratings, according to Duffrin (2011), can be skewed by random events including student illness or a neighborhood crisis, all of which are outside a teacher's control. Therefore, value-added information should be combined with other information and the understanding of people who know what's going on in a classroom (Duffrin, 2011).

In theory, teachers should only be assessed for student learning that occurs when the student is within their class (Papay, 2011). Because of this, a value-added measure would compare a student's performance on the first day of the school year to their performance on the very last day (Papay, 2011). However, in practice, tests rarely occur at the very beginning or end of the school year and thus leaving a portion of one teacher's instruction to be attributed to another teacher, from the previous year (Papay, 2011).

Further, it is often difficult to attribute students' learning only to one teacher. This is due to both the mobility of the student within and across schools, and because students often learn skills from several teachers simultaneously (Croninger and Valli, 2009). This can be seen particularly in middle and high schools where students often have different teachers for many of their subjects.

Finally, most value-added models use a previous year's test score as a measure of past performance and any learning loss or gain that occurs over the summer becomes intertwined with value-added estimates (Papay, 2011). If a teacher's class has a disproportionate amount of students whose skills were lost over the summer, those

teachers may actually demonstrate greater gains in impacts on student learning than teachers with similar estimated effects whose class did not demonstrate learning loss during vacation (Papay, 2011).

If value-added measures continue to be used as part of the evaluation about teacher performance, more attention should be paid to the tests themselves (Papay, 2011). Currently, tests are designed to measure student, not teacher, performance (Papay, 2011). The tests are designed to identify a district's best teachers and differ from those designed to assess student proficiency. Further, when a test is administered must be considered very carefully (Papay, 2011). An accountability system based on several different assessments would provide data that can be used for both coaching and instructional responses on all tests in that subject area (Papay, 2011). Further, while the ability to raise student test scores is important when considering teacher quality, it represents only one of many key factors of what makes a great teacher (Papay, 2011). Combining this test-based accountability with other performance measures could help to address these issues (Papay, 2011). Given the amount variation in any single assessment of value-added models, combining multiple sources of information could provide both schools and teachers with a better understanding of their performance (Papay, 2011).

Summary

This chapter has looked at how value-added models came into existence within our current public education system by building on previous policies such as No Child Left Behind, Common Core, and Race to the Top initiatives. Value-added models directly tie teacher evaluation to student performance on end of year high-stakes tests.

The policy, meant to ensure accountability on the part of the classroom teacher for both content and student retention of information, is based on an understanding that all students begin public school on an equal footing. This chapter has discussed how that understanding may be flawed. Value-added models were a result of federal policy and then implemented in a rapid fashion by states. This current study will build on the understanding of value-added models established in this chapter and look further as to why states chose to implement the policy.

CHAPTER 5

Innovation of Diffusion

A person cannot claim to understand policymaking without being able to explain the process through which governments adopt policies (Berry and Berry, 2014). With this understanding, public policy scholars have conducted extensive research into the innovation and diffusion of policy (Berry and Berry, 2014). Most policies spread gradually. The spread is incremental and normally found to be a product of policy learning and emulation (Boushey, 2012). Recently, however, it has been found that a substantial portion of innovative policies trigger a rush by states to copy a new popular idea (Boushey, 2012), one of these being value-added models.

It is the purpose of this chapter to discuss what policy diffusion is and why it is important for scholars to study. In the chapter I will explain the mechanisms for diffusion such as learning, imitation, normative pressure, competition, and coercion. I will also explain the types of regional diffusion. I will next look at how federalism plays a part in the diffusion process by looking at federal influence. I will go on to explain how both internal and external determinants such as problem severity, electorate attitudes, political culture and policy coalitions influence the diffusion of a given policy. Lastly, I will explain the general model I am testing, the federal influence on the model and the internal determinants that I will use to help explain value-added policy adoption.

What is Diffusion and Why it is Important

Policy diffusion is the process by which innovation of new policy spreads (Gray, 1973). The diffusion process requires contact among states or districts, a transmission process, and acceptance of the given policy by additional states (Lave and March, 1993). Further, policy innovation is a program that is new to the state government that is adopting it; a specific policy can be innovative in a one state even if it is not new to the other surrounding states (Berry and Berry, 2007). This policy diffusion process explains the spread of an idea, product, or technique over a given period of time. Further, diffusion contains three interrelated components: the innovation or ‘thing’ that diffuses, the adopter (the state), and the adoption environment (the diffusion milieu) (Feller and Menzel, 1977). For this study, innovation will be defined as the first use by a given state of a new idea, where new means only new to the adopting agent (Rogers and Shoemaker, 1971).

Much attention has been focused on how variation in state political, institutional, cultural, economic, and social characteristics allow for governments to innovate over time (Boushey, 2012; Berry and Berry, 1990; Canon and Baum, 1981). Through this body of research, it has been demonstrated that both political and institutional attributes including ideology (Berry and Berry, 1990), legislative professionalism (Volden, 2006), political competition (Karch, 2006), and in some cases, public opinion (Pacheco, 2012) shape a government’s propensity to innovate.

Yet, further research has sought to identify the role of “carriers” in the diffusion of innovations and demonstrate how professional organizations (Boushey, 2012; Balla; 2001; Gray, 1973), interest groups (Mintrom and Vergari, 1998), and policy

entrepreneurs (Mintrom, 1997) all work to collaborate amongst state governments to target policy change and effectively show policy lessons through their own successes and failures which diffuse across state lines.

When thinking of diffusion models, we usually think of a set of governmental jurisdictions (e.g., the fifty states in the United States) as a social system and maintain that the pattern of policy adoptions by member governments results from one member emulating the behavior of another member (Berry and Berry, 2014). In the most general sense, the probability of adoption of a policy by one governmental jurisdiction is influenced by the policy choices of other governments in the system (Berry and Berry, 2014).

As a property of the adopting organization, the behavior indicates innovativeness as a part of the diffusing idea and further indicates adoptability. This demonstrates that if a state adopts a new policy relatively quickly and extensively, the action alone provides evidence both that the state is innovative, and that the idea is adoptable (Down and Mohr, 1979). Diffusion represents a process of quick adoption, which imitates and mimics policy learning (Foucault and Montpetit, 2011). The demand to adopt new policies by policy makers creates an environment that encourages the replication of peer experiences with new ideas, based on anticipated payoffs (Boushey, 2012).

Innovation is a process that occurs in two stages (Zaltman et al., 1973). The first is the diffusion stage, this stops when a prospective adopter may have heard about the innovation and how it might be of benefit to the organization (Zaltman et al., 1973). The

second stage is the adoption stage. This starts at the point of awareness and continues until a decision about adoption has been reached (Zaltman et al., 1973).

Once the potential innovator becomes aware of the “new” policy and chooses to adopt, Feller and Menzel (1977) describe the next steps which are the pattern of organizational innovativeness. They explain diffusion as a source of discrete and continuous alterations in relation to a specific set of adopters (Feller and Menzel, 1977). Further, the diffusion contains external determinants to the organization which pressure the organization to alter its existing policy practice. These define the range of feasible alternatives to the current policies. It also looks at which conditions for which information flows and evaluates the performance characteristics of the proposed innovation. These social environmental components are then filtered through the internal decision-making process of the organization to produce a decision to accept or reject the innovation (Feller and Menzel, 1977).

Feller and Menzel (1977) suggest that these differences affect the rate of adoption of specific techniques, the significance of other factors that influence diffusion patterns across a set of potential adopters, and the likelihood that there will be any systematic pattern of organizational innovativeness. Further, this pattern of organization creates a chain of events that can cause disruption throughout the adoption cycle. This is due to the chain having broken or missing links that can completely block or lengthen the policy diffusion process (Feller and Menzel, 1977).

The innovation-decision design combats the effects of interaction by bringing interacting variables together in the same equation (Winter, 1968). In terms of decision-

making elements or choice as opposed to state organizational variables, the determinants for adoption are contingent upon the effect that decentralization has on innovation (Down and Mohr, 1979). The decision to innovate also depends on the amount of agreement on values and information that exists among state decision makers. If decision makers share the same ideals and are all well-informed, then there should be innovation of policies between states (Kaufman, 1960). Additionally, state wealth or slack will play an outsized role in a state's calculations surrounding the potential adoption of an innovation (Down and Mohr, 1979).

Further, the properties of innovations, such as, compatibility and relative advantage, cannot be measured or their impact assessed without first referencing the state that is potentially implementing the policy. It is because of this that any model that tries to anticipate the extent and time of adoption must be sure to include both characteristics of the state and the innovation in the context surrounding a particular decision (Down and Mohr, 1979).

When considering the variables for policy diffusion and adoption, we need to consider why a particular variable is expected to be related to innovativeness. This can help to reveal a greater level of generality and abstraction (Down and Mohr, 1979). Predictors need to be framed to be broadly applicable, which will cut down on instability within the models (Down and Mohr, 1979).

Mechanisms for Diffusion

Recent literature has identified a variety of alternative mechanisms by which the policy choices from one jurisdiction can influence the choices of other governments. At

least five mechanisms have been discerned: learning, imitation, normative pressure, competition, and coercion (Berry and Berry, 2014).

Learning occurs when policymakers in one jurisdiction (or state) derive information about the effectiveness (or success) of a policy from previously adopting governments (Levy, 1994). Success can be defined not only in meeting policy objectives set forth, but also in achieving personal political goals of those who lead such as winning reelection (Shipan and Volden, 2008).

Imitation occurs when one government seeks to have the same characteristics as another (Berry and Berry, 2014). Government A adopts a policy adopted by B simply to look like B (Shipan and Volden, 2008). Imitation occurs because policymakers within government A see B as worthy of mirroring. This mirroring effect prompts government A to adopt any policy that B adopts independent of any evaluation of the policy or its effectiveness to achieve its set goals (Simmons et al., 2006). In many cases, policymakers within the governments looked to “leader” or “signal” governments. The governments may be characterized as leaders due to their size or wealth (Karch, 2006). They may have also earned strong reputations or high levels of credibility through previous actions (Grupp and Richards, 1975). In other cases, the governments imitated share characteristics that make them role models or valuable peers (Weyland, 2004). For example, many scholars have found that policymakers have a tendency to emulate policy adoptions of governments that share similar partisan and ideological orientations (Volden, 2006). It has been found that the major difference between learning and imitation is that learning focuses on the action (the policy being adopted by another

government) and imitation focuses on the actor (the other government that is adopting the policy) (Shipan and Volden, 2008).

These previously mentioned factors are often coupled with normative pressure, which is when a government submits to normative pressure of fellow adopters (Berry and Berry, 2014); when there is adoption of a policy not because there is imitation of any particular government or learning from the experience of other adopters, but because a government observes that the policy is being widely adopted by other governments, and because of the adopters shared norms (DiMaggio and Powell, 1983). Shared norms often occur across societies (and states), allowing some policies to be adopted and later taken for granted (DiMaggio and Powell, 1983). Some scholars have pointed to the role of experts within (state) governments in forging a consensus on norms (Haas, 1992).

For policies made by administrative agencies instead of legislatures, the shared norms that accompany diffusion have been argued to be the result of professionalization of agency personnel regardless of the area in which they serve and develop a shared understanding of “good” policy by virtue of their common professional training (Teodoro, 2009) and the new policy is further thought to be within “best practices” for the same policies to be adopted elsewhere (Walker et al., 2011). Shared norms are often further reinforced by the participation of those within intergovernmental professional associations that constitute networks across which policies can diffuse (Berry and Berry, 2014).

Further, a policy diffuses through competition when a government’s decision on whether to adopt the policy is motivated by the desire of state officials to achieve an

economic advantage over other jurisdictions or to prevent other jurisdictions from securing an advantage over it. In diffusion via competition, a government's adoptions can change the payoffs of adopting (Garrett, 2006). In location-choice competition, one type of competition that many scholars describe, governments seek to influence the location choices of individuals who are in a position to bring effective change in more than one place, which usually is their own location (Meseguer and Gilardi, 2009). This would bring potential monetary advantages to the area for which the policy is being implemented, making the area more desirable.

Diffusion due to coercion is when entity A is more powerful than entity B, then A is enabled to take actions that increase B's reasons for wanting to adopt a certain policy (Berry and Berry, 2014). This is sometimes accomplished by threatening action if the weaker entity, for our purposes a state, does not acquiesce (Simmons et al., 2006). Researchers studying diffusion in the American federal system focus on vertical coercion (diffusion across levels of government). In some cases, the national government can simply mandate certain policies be enacted by states (Zamani-Gallaher, 2014) such as emissions standards (Bhat et al., 2016). Another example would be that the US Supreme Court can make rulings that limit state policy choices (Hoekstra, 2009). Other cases where I find federal influence is a grant-in-aid creating a financial motivation for a state to adopt where the government uses a carrot rather than a stick by creating an incentive for a state to adopt a policy (Berry and Berry, 2014).

Each of these mechanisms focuses on a different channel of communication such as federal directives or court rulings and differences across government means. What is

found is that each state relies on one, or at times the combining of multiple mechanisms, to justify why one government would emulate another government when making policy decisions (Berry and Berry, 2014).

Situation Framework

Building upon the mechanisms already discussed, we find that when a state finally chooses to adopt a policy, this takes place within a situation framework (Hartgen, 1973). The situation depicts a variety of constraints that serve to limit the effectiveness of a given innovation. Further, these constraints include societal customs and norms, as well as physical characteristics and economic conditions that make up a state (Hartgen, 1974).

A state's resources are also either constraints or advantages that need to be considered since they are needed to offset costs (Downs and Mohr, 1979). Thus, various resources would logically become even "better, more stable predictors of innovation when the associated costs are also considered" (Downs and Mohr, page 391, 1979). Building on this, innovation on the basis of costs and benefits might seem naive considering any behavior can obviously be reduced to "pleasure and pain" (Downs and Mohr, page 392, 1979). Despite its universal applicability, benefit-cost analysis would be not be used to explain all behavior. For the normal, routine types of state decision-making, a plethora of concepts including sequential attention to goals, resolution of conflict, reaction to feedback, and problematic search can be employed (Cyert and March, 1963). Further, for adaptive behavior, a technique including stimulus, reinforcement, and learning comprise demonstrates a better response. A benefit-cost configuration is a good descriptive model for innovation because innovation is

instrumental as it is supposed to achieve a better result than the current policy. Therefore, the consequences, both positive and negative, take up decision-making attention and create a “comparative rational process” though the results may not be exact (Downs and Mohr, 1979).

When utilizing the benefit-cost model, four key factors are considered when undertaking a new policy; they include risk, average cost, uncertainty, and instability (Downs and Mohr, 1979). Risk considers the degree of concern over possible disaster. Average cost includes the cost of overall implementation and discontinuance. Uncertainty considers the lack of confidence that the state has in its benefit-cost calculations. Lastly, instability considers the future of benefits including fear that the benefit/cost ratio will without warning decrease at some point (Downs and Mohr, 1979).

The ability for a government to innovate often relies on the extent of the resources (including financial) that they have at their disposal (Berry and Berry, 2014). A state would need sufficient resources in order to adopt a potential new policy. Some new programs require major monetary investments, and therefore there must be the financial resources available for adoption (Alland, 2004). This leads to the understanding that the fiscal health of a government will have a positive or negative impact on its propensity for policy adoption (Lowery, 2005). It can be argued that the capacity of a government to be able to finance extensive public services is the ultimate determinant of a government’s ability to innovate (Berry and Berry, 2014).

In general, the focus on resource availability has either been the total of financial resources, or within the decision-making unit such as slack resources (Feller and Menzel,

1977). According to Downs and Mohr (1979) five types of resources appear to have the greatest impact, including wealth, manpower-expertise and time, equipment, information, and staff tolerance for change. In other words, a state must have the financial ability, as well as the human capital (both in numbers and expertise), in order to implement. There also needs to be a learning component which is often gained through policy emulation taken from 'cue' states, as well as the willingness of the participants such as states and school districts in the case of this study, to see the policy to fruition.

Further, states may be required to meet additional performance standards as the consequence of new legislation (Feller and Menzel, 1977). Education policies would be a good example of this, with their attendant requirements that states develop implementation plans to meet federally established educational standards. The need to meet these requirements leads state and local agencies to look for the techniques that are parts of enforcement programs and therefore makes them more open to the initiatives (Feller and Menzel, 1977).

State policy innovations, such as the previously referenced Common Core and value-added models, are heavily underwritten by federal funds, and are scrutinized by federal officials in order to demonstrate compliance. Changes in objectives at the federal level make it easier for state agencies who are in on-going negotiations with federal agencies to receive certain types of funding which will keep states in tow with federal policy objectives (Feller and Menzel, 1977). Additionally, federal pressures on state and local governments to innovate are usually coupled with new financial resources which serve to coerce states into complying with federal mandates. Federal funds appear to be

the most important resources in state and local government, and therefore, are a major influence on the obtaining of new technologies (Feller and Menzel, 1977).

In the above examples, states become funnels for which federal funds flow through to the cities/counties. State influence may extend to the approval of specific technologies which are later implemented at the county level. States may also advise cities on best-practices and sway them to adopt state led initiatives. Additional actors such as vendors may focus their efforts at convincing state agency officials of the superiority of their product before going to local districts (Feller and Menzel, 1977).

Policy scholars have also hypothesized that “problem severity” is an important determinant of the motivation to innovate (Berry and Berry, 2014). Problem severity can influence the motivation of government officials to adopt a policy by defining the need for the policy or indirectly by stimulating demand for the policy through the incitement of societal groups (Berry and Berry, 2014). For example, poor economic conditions during the COVID 19 pandemic contributed to the adoption of various aide programs which saw an increase in the demand and need for assistance for basic needs such as housing, food, and utilities (Tracking the COVID-19 economy’s effects on food, housing, and employment hardships, 2022). Additionally, in his study, Stream (1999) looks at how the rate of those who are uninsured within a state influences the likelihood that the state will adopt various insurance reforms. When looking specifically at education reforms, Mintrom and Vergari (1998) argue that the greater the ratio of state education funding to local funding, the more likely that a state legislature will consider a reform such as school choice.

Alternatively, costs must also be calculated when considering policy innovation. These can be broken down into decision costs and implementation costs (Downs and Mohr, 1979). Decision costs are characterized as the costs of arriving at a decision to implement an innovation or not, including to what extent and at what rate. Implementation costs include the costs of actual implementation of the innovation (Downs and Mohr, 1979).

Model Building of Value-Added Policies

Building off of previous literature, I construct a diffusion model that serves to show how value-added policies have spread from one state to another. I specifically further build on research regarding policy diffusion and state adoption of federal policy which have largely been pioneered by Walker (1969), Gray (1973), and more recently, Berry and Berry (1990). Walker (1969) concentrated on state diffusion and the rate at which it occurred. Walker (1969) specifically attempted to answer why a state would be more apt to adopt a policy program by looking at what informed the decision process of the states.

I focus specifically on the diffusion of education policy, looking at why states would adopt the policy of tying teacher pay and evaluations to student test performance. I theorize that in the context of educational policy, states are more likely to adopt policy that has been enacted by “early adopters.” Once a policy is implemented by neighboring states, states will then be more likely to adopt the proposed policy due to increased regional pressure. In conjunction with state pressures, internal demands of state governments such as economic constraints. Additionally, the federal government also

influences the adoption of state policy through incentives such as grants. The influence of the federal government, internal components, and the adoption of policy in neighboring states will reinforce one another when considering educational policy adoption; all working together and creating a higher propensity for policy adoption.

Regional Diffusion

Further, Walker (1969) assumes that “pioneering states” gain their reputation because of the speed at which they readily accept new programs. These pioneer states would have what Rogers and Shoemaker (1971) would call 100% psychological trialability; these are instances in which the potential adopter does not need experience to decide whether or not to adopt, but instead, simply tries out the innovation. Walker (1969) analyzes this with the use of a regression model and developing an innovation score based on at least twenty states having enacted policies in the areas of welfare, health, education, conservation, planning, administrative organization, highways, civil rights, corrections and police, labor, taxes, and professional organization. What Walker (1969) found was that certain states were seen as “pace setters” within their region. Neighboring states look to these states to take their cue on policy adoption. These states would serve as what Downs and Mohr (1979) call an advocate; for each new idea studied, these states would serve as a source to other potential adopter states. These “pace setters” (Walker, 1969) have the important conceptual role of specifying the benefits and costs that are connected with a policy and the point at which adopting organizations can assess these benefits and costs accurately (Downs and Mohr, 1979).

Walker (1969) concluded that there was a continuum of how quick states are to accept policy innovations. Walker divided the country into five regions and conducted a regression of the rate at which a state would adopt a given policy. This created a hierarchy of adoption that demonstrated either a likelihood for innovation or a reluctance to accept innovation. Walker also noted that improved communications between states, especially state officials speaking to their regional counterparts, increased the acceleration of diffusion. He concluded that there is a national system of emulation based on regional groupings. These systems set the norms and national standards for policy adoption. Walker's work was the first of its kind to examine the diffusion of policy when taking into consideration regional proximity.

Adding to the literature of policy diffusion among states, Virginia Gray (1973) simplified her focus on policies, looking specifically at education, welfare, and civil rights. Gray found similar results to that of Walker: an interactive effect where states who have not yet adopted are influenced by adopters in their region (Gray, 1973).

Significantly, Gray found that states vary in their level of support for policy adoption based upon specific stimuli such as federal grants. This differs from what Walker anticipated, which was that states would show uniform adoption simply on the criteria of the adoption influence of neighboring states. Gray's study shows that though neighboring state influence is one component of adoption, it is only part of the reason why a state would adopt a particular policy. The policy in conjunction with state dynamics and incentives were highly correlated with adoption; in other words, whether a state had a Republican or Democrat governor and if the state was offered a large grant

from the federal government played a large role in the adoption of policy. Further, adoption of one policy under a particular area, education for example, did not guarantee the adoption of further education policies (Gray, 1973). This adds to the understanding of diffusion that was fostered under the Walker model and demonstrated, unlike Walker's prediction, that policy diffusion in states does not just focus on "innovative states" for their cue on whether or not to adopt a specific policy. The states also look inwardly to ensure the policy aligns with the needs and political ideologies of the state itself.

Building on this, Feller and Menzel (1977) found the primary determinant of decision making when concerning the adoption of new innovations within states is the mission agency nor the central administration. Additionally, agencies adopt new technological innovations in response to pressures that come from outside of the organization (Feller and Menzel, 1977). In many of the circumstances, the pressure is related to the technology and can be likened to being approached by a salesman offering a new product (Feller and Menzel, 1977). While in other cases, agencies are confronted with new problems and then will seek a solution to the problem (Feller and Menzel, 1977). This may involve the adoption of a completely different technique (Feller and Menzel, 1977).

Federalism

What many of these studies have hinted at, but not fully vetted, is the role of federalism in policy diffusion. Andrew Karch (2006) looked at national intervention regarding diffusion in which he claims the motivation to implement a policy stem from a "problematic social condition" such as high unemployment rates. He claims that it is a

trifecta of influence from the state, regional, and national forces working in conjunction that determine the adoption of policy. The study only focused on what may prevent a state from adopting a policy, such as funding. This allows us to understand why a state that may want to adopt a policy did not do so but does not examine the internal determinates of a state that would specifically allow for adoption to occur.

Political responses to new policy ideas change based on how politicians and the public define a problem (Boushey, 2012). When considering how problem definitions can change political responses to public issues, focus is needed on agenda setting. By looking at the policy problem from a different angle that policy reform can begin to take shape and generate interest in the policy problem at hand (Baumgartner and Jones, 2009).

Scholars have found that changing policy environments lead to the acceptance of a different policy image (Kingdon, 1984). Additionally, abrupt changes to policy can result from a focusing event that provides a shock to the system and redirects attention to the central policy issue such as our nation's failing schools highlighted in a Nation at Risk (Baumgartner and Jones, 2009).

Federalism can make policy coordination difficult (Pralle, 2003). However, federalism can also create opportunities for policy innovation to take place, as municipal, county, and state governments create a new way of responding to local concerns (Boushey, 2012). This takes place through venue shopping where activists and interest groups take their concerns to multiple government venues in an attempt to obtain increased support for their ideas (Baumgartner and Jones, 2009). This can create favorable conditions for the rapid spread of innovative programs or new policy images

(Holyoke, 2003). However, federalism can create both opportunities and constraints for the spread of new policy. Baumgartner and Jones (2009) found that expanded attention to an issue can result in the nationalization of the issue and can lead to the nationalization of policy attention and what was once only a localized concern has now garnered national merit. This allows for the federal government to act as a powerful engine of change for state policy adoption.

Further, positive feedback cycles can also occur even without the federal government being at the helm (Boushey, 2012). Issue attention can also be nationalized through the collaborative efforts of interest groups that advocate for policy reform, bringing their campaign across states (Boushey, 2010).

As alluded to previously, it is the combination of policy image and institutional venue that is critical for understanding why policy diffusion may occur (Boushey, 2012). When issues gain national attention, policy can rapidly diffuse when its adoption across the country is mandated by the federal government (Boushey, 2012). When an event occurs that creates a large amount of attention within a state, diffusion may occur and create a “policy bandwagon” as states attempt to quickly emulate the innovations of their peer states (Boushey, 2012).

Internal and External Determinants

Diffusion of policy needs to take in consideration both internal and external characteristics. Examples of internal characteristics include political factors such as state party control, urbanization, and relative state poverty levels. External determinants look at how policies spread from one state to another and the factors surrounding the

occurrence (Li, 2017). Attributes of innovation include cost, relative advantage, communicability, complexity, trialability, and susceptibility to successive modification. While attributes of the adopter include motivation, professionalism, informedness, resources, pride, risk, and propensity (Feller and Menzel, 1977). However, it is important for the potential adopter to understand both what policy innovation can offer in terms of benefits, as well as the position of the adopter to innovate (whether they have the resources and motivation to adopt).

In order to understand innovation of policy adoption, I need to distinguish between the diffusion and adoption stages. This eliminates a “bi-dimensionality” in the thinking of time of adoption (Downs and Mohr, 1979). By restricting attention to the adoption stage, it becomes possible to research both the time of adoption and the extent of use within the same theoretical framework (Downs and Mohr, 1979). However, when considering the earliness of adoption after the point of awareness this becomes a matter of state strategic decision making (Downs and Mohr, 1979).

Expanding on this, Berry and Berry (1990) added an additional element by utilizing event history analysis to look at the influences of nearby states. Berry and Berry (1990) used similar models and looked at state lottery adoptions which reflected both the internal and regional influences. This built upon research done by Walker (1969) and Gray (1973) by combining what both had done into one comprehensive study of policy diffusion. Berry and Berry (1990) acknowledged, as did Gray (1973) previously, that a state will not blindly emulate its neighbor states’ policies without its public officials being influenced by the political and economic environment of its own state (Berry,

Stokes, and Berry, 1990). Previous models look at a state's motivation to innovate and the availability of resources to overcome any obstacles. They further theorize that as motivation to innovate increases so do the obstacles. The increase in obstacles is due to unexpected increases in monetary demands and resources needed to implement. Also, unpopular policy should be easier to overcome and justify to voters if the policy has first been adopted by nearby states (Berry, Stokes, and Berry, 1990). The study found both elements to be true; innovation is positively associated with the desire to take new approaches, and negatively associated with the strength of the obstacles (Berry, Stokes, and Berry, 1990). This is directly related to the availability of resources to overcome the obstacles. They also found that neighboring state influence was highest when internal characteristics of a state were favorable for adoption; in the case of a state lottery this would include the poor fiscal health of a state which would benefit from the influx of currency generated by a state lottery (Berry, Stokes, and Berry, 1990). The study furthered the use and understanding of event history analysis and its use on rarely occurring events such as looking at specific policy and route to adoption.

All of these groundbreaking pieces have expanded the scope of policy diffusion which stretches across policy domains such as education (Key, 1949), technology (Menzel and Feller, 1977), energy (Regens, 1980), and even tort law (Canon and Baum, 1981). Normally, what these studies have focused on is the timing of when a state adopts a given policy.

Additional Explanations for Adoption

When considering the proliferation of value-added models, what we find is that traditionally, the two most dominant factors in why a state would adopt a specific education policy were proximity to fellow adopters and partisanship within states. Building off earlier concepts, Republican led administrations tend favor personal accountability. For example, when considering healthcare, Republicans would say that people have a responsibility to take care of their own needs; if healthy people fail to take out health insurance, and then fall sick, that is therefore of their own doing. Democrats, on the other hand, would retort that we owe people health care irrespective of the choices they've made (Illing, 2017).

With education policy, Republican led administrations have been some of the first adopters of having funding tied to results (Jochim and Lavery, 2015). Policies that favor personal responsibility tend to be more conservative in nature. Education policies are no exception. Thus, a policy that ties classroom teacher accountability to set student achievement, along with merit incentives should tend to be more readily accepted and adopted in more conservative states and propagated by a conservative executive.

When such measures are on the state's agenda, states with lower per capita income, and thus lower per pupil spending, are more apt to hold accountability for funds disbursed since the resource is at a minimum and needs to be closely monitored. These same states tend to also be conservative and with lower incomes and thus states would want to ensure a return on investment of funds distributed.

Education policies that favor a direct return on funds disbursed are meritocratic in nature. As such according to previous literature, merit policy can be explained in part by partisanship and urbanization levels within the state (McLendon et al., 2006; Li, 2017; Howell and Magazinnick, 2017).

In this study, I address this question by combining the methods put forth by Berry and Berry with the utilization of event history analysis, as well build upon the work of Gray and Walker by noting neighboring state influence and internal determinants. This will give an understanding of which states are more likely to adopt policies based on the pressure from “early adopters”, as well as look at the motivating factors such as pressure groups, in determining why states are more apt to adopt certain policies.

Summary

This chapter has looked at the concept of the diffusion of innovation. There are several reasons a given policy may diffuse including ideology, mimicking effects, emulation, regional pressure, and cues from known policy innovators. Previous literature has highlighted the importance in understanding the roles that each of these factors plays in how a policy can spread. This study will build on this concept of how policies diffuse. I will look at various reasons for policy diffusion throughout the country, taking into consideration the effects of previous similar policy adoption along with the role both internal and external determinants may play.

CHAPTER 6

Data and Methods- Hypotheses for State Adoption of Value-Added Models

The focus of this research looks at why value-added models diffused across states. Using lessons learned by scholars of policy diffusion and innovation, I will present a series of hypotheses that may help to explain the adoption of the policy of interest. I will then go on to explain the variables that were controlled for within the models and their significance in previous literature. Next, I will identify the data that will be used to operationalize the concepts contained in the hypotheses. Lastly, I will explain the statistical methodology that will be used to test each of the hypotheses.

When considering the adoption of new educational policies, implementation often depends on adequate funds to support professional development of teachers and other staff, additional investments in curriculum, tests, technology and other instructional supports. In many cases, changes to state accountability systems, such as linking student test scores to teachers, and using student test scores to evaluate teachers, fall under the purview of elected state policymakers (Jochim and Lavery, 2015). What this demonstrates is a need for funding and a favorable state legislation for policy implementation to take place. There is also a learning aspect to policy adoption, as previously discussed, with states taking their cue from “early adopters” of policies and then reconstructing the policy to fit their particular state needs. With the vast majority of states, 32 out of 50, choosing to implement value-added models of merit pay, I seek to explain why the policy diffused across states by proposing seven hypotheses that seek to explain diffusion rates.

Hypotheses

States will take their cues for adoption from early adopter states. Among the early explanations for policy adoption, scholars identified the importance of states regarded as innovators and, also, regional influences. Hence, this leads to the first three hypotheses. The innovative state (first state to adopt the policy) within their region will demonstrate the effectiveness and/or challenges of the policy to the other states. As additional states within the region adopt similar policies (mitigating for challenges previously presented), there will then be pressure for any states that have not adopted the policy to follow suit and conform with the region.

Hypothesis 1: *The probability that a state will adopt the educational policy of value-added models for teacher evaluations increases when it is “signaled” by a regional cue state. Diffusion will begin at the cue state and move outward from this origin point.* Each region of the U.S. has a “cue” state for policy adoption. These four cue states (New York, Nevada, Kentucky and Oklahoma) are normally the first within the region to pilot a new educational policy within their schools. The states reflected as cue states are recognized for first adopting Common Core within their region. They are as follows: New York for the Northeast, Nevada for the West, Kentucky for the Southeast, and Oklahoma for the Midwest (Achieve, 2013). Walker (1969) noted that certain states were seen as “pace setters” within their region and neighboring states look to these states to take their cue on policy adoption. Once the policy has been adopted by the cue state, it signals to surrounding states to adopt the same or similar educational policies.

This differs from regions since regions are broader in nature whereas, with cue states, fellow state leaders look specifically at one state for innovation with diffusion outward from the one cue state. While cue states provide a signal regarding policy adoption to an entire region, the diffusion of policy also is influenced by the actions of neighboring states.

Hypothesis 2: *The probability that a state will adopt the educational policy of value-added models for teacher evaluations increases with the extent to which bordering states adopt the policy.*

Similar to the case with cue states, many states are hesitant to be the first to adopt a new and innovative policy. However, if the policy is first tested by a state that borders and the policy is viewed as worthwhile and continued, then a state is more likely themselves to adopt the same or similar policy.

Hypothesis 3: *The probability that a state will adopt the educational policy of value-added models for teacher evaluations increases with the extent to which states within their circuit court regions have adopted the same policy.*

The concept of regional influence does not have to mean only being next door to another state. It can also mean that states within a region can influence one another through shared regional issues. For instance, state compacts are tools that states use to address problems that are not contained within formal political boundaries. For this reason, I examine regional influence by looking at geographic clusters of states.

States within the same region share similar ideologies (Walker, 1969). This hypothesis tests the notion that states are likely to mimic one another when they share

regional commonality. In other words, I am looking not just at the neighbor that a state directly borders, but any neighbor within the same subdivision, in this case circuit. Circuits will be defined by which federal circuit court each state falls within. Regional political ideology will be represented by circuit court regions.

It is therefore assumed that states within the same circuit would be apt to adopt similar education policies with diffusion occurring around the same time. Additionally, circuit regions serve as a proxy for the diffusion of regional ideology. There are eleven total throughout the US. Previous diffusion literature has not looked at circuits court regions and with partisanship at time high levels (Taylor and Watts, 2020), I should see diffusion of policy between states within the same circuit court region. Additionally, regional circuit courts are a reasonable proxy for general regional influence since rulings handed down by the courts should be adopted in a similar fashion by all states within the circuit.

The diffusion of innovation literature also identified the impact of policies that lay the groundwork for other policies. It is by understanding how a policy diffuses that we are able to foresee favorable conditions for future policy endeavors. It is that general explanation that I am drawing on to develop hypotheses #4-6.

Hypothesis 4: States that adopted the policy of performance-based funding for higher education are more likely to adopt meritocratic measures proposed by the federal government when concerning K-12 and value-added models.

Performance-based funding ties state education funds to set outcomes such as graduation rates. The practice is both coercive and a form of accountability for institutions in higher education.

States having piloted the accountability program of performance-based funding and shown success in the increase graduation will signal for the additional piloting of these type programs in the K-12 system. Value-added models are similar in how they were rolled out and hold accountability. States that have already had success with performance-based funding in higher should be better primed and therefore more likely to adopt future accountability policies such as value-added models.

Hypothesis 5: States that adopted the policy of incentivizing teachers to work in targeted, low incomes areas, are more likely to adopt meritocratic measures proposed by the federal government when concerning K-12 and value-added models.

Target schools are defined as schools that are low performing (not meeting minimum standardized testing requirements and/or making required student testing gains) that are targeted for reform (Torre et al., 2012). In many cases these schools have been “taken over” (under the direct purview) of the state (Torre et al., 2012). Hard-to staff schools are defined as urban and rural, high-poverty, high-minority, and low-achieving schools (Opfer, 2011). Hard-to-staff schools are often found in inner cities with a high percentage of the student population living in poverty (Allen, 2000). Additionally, hard-to-staff schools are also found in rural areas that are in economically depressed or isolated districts and tend to offer low salaries or lack the amenities that more urban areas have that would attract better qualified teachers (Allen, 2000). Incentivizing teachers to work

in these targeted schools allows for states to staff “hard to employ” positions within their district and allows for teachers to receive monetary compensation for additional effort being put forth in these turn-around situations. With increased effort equating to an increase in pay, this scenario becomes meritocratic in nature. It is reasonable to presume that a state that is willing to engage in one type of meritocratic practice would also employ another similar measure.

Hypothesis 6: States that adopted common core policies, are more likely to adopt meritocratic measures proposed by the federal government when concerning K-12 and value-added models.

States who had previously adopted Common Core policies to stay in compliance with federal curriculum mandates would therefore feel the same compulsion to stay in compliance with federal policies and therefore more apt to adopt value-added models.

Internal Factors

In all of my models I include controls for Republican government control of the state, urbanization, per pupil spending, poverty, and strength of teacher unions. Previous research has shown these variables to be significant and can add to why state may adopt a meritocratic policy.

The probability that a state will adopt the educational policy of value-added models for teacher evaluations is greater when the governorship and both houses of legislature are under Republican Party control.

The Republican Party is conservative in fiscal matters. An educational policy that would only fiscally reward a public university or college would fall under the umbrella of a conservative policy and potentially punitive in nature. Conversely, Democrats tend to support less punitive policies, therefore, states where Democrats control the state legislature and governorship should be less likely to adopt the educational policy (Allen et al., 2004).

Previous studies such as Berry and Berry (1990), McLendon et al. (2006), Yi (2017), and Walker (1969) all include Republican state control in their models. All mentioned scholars have found the presence of Republican control to be significant in their understanding of why diffusion was occurring. This is an important finding since states with Republican leadership were more likely to adopt meritocratic policies since, as stated in previous chapters, accountability measures tend to be conservative in nature and follow closely with Republican ideals. This would demonstrate that ideology is a key factor when considering policy adoption.

The probability that a state will adopt the educational policy of value-added models for teacher evaluations is greater when the state has a lower percentage of urban areas.

An area is considered urban if it has a population of 50,000 (US Census Bureau, 2016) residing within the geographically defined city limits. Large urban areas like New York City have the capacity to identify and elect to be the first to implement policies that many small rural jurisdictions cannot support. Rural departments are more likely to have less experienced staff, tend to have districts with limited resources, and are often isolated

from other districts which may provide support or guidance. They also may have a lower tax base from which to pull additional education funding. In addition to limitations on staffing and funding, small rural jurisdictions tend to have fewer local partner organizations (Harris and Mueller, 2013). In larger districts, schools may have more opportunities to work with organizations, citizen groups, businesses, and others who aid in policy development and other activities (Harris and Mueller, 2013). It is because of these partnerships within the urban communities and better trained educators that makes states more likely to enact a new policy.

Steinhubel et al. (2020) in their work looked at rural versus urban adoption of technology. They found that the more urban an area, the more likely it is for adoption to occur due to things like social interaction which effects the overall number of possible adopters and supports further diffusion. This is key since the more urban areas allow for additional access points in the understanding and perpetuation of the novel policy and leads to higher adoption rates.

The lower the per capita per pupil spending is, the higher probability that the state will adopt the educational policy of value-added models for teacher evaluations.

Grants were designed to help states accomplish goals (Welch and Thompson, 1980). States that invest in students expect to see a return on their investment. However, states that spend less per pupil in higher education would be more likely to become vested in the program of performance-based funding since their public colleges and universities are more reliant on state funds. To this extent, the funds being offered and tied to higher education graduation rates, would be most attractive to those states that also

have some of the lowest per pupil spending in the country: \$2,419 (AZ), \$3,542 (VT), \$3,185 (NH), \$4,302 (PA), and \$4,494 (KS) (Science and Engineering State Indicators, 2020). Regarding per pupil spending, I use data derived from the USA Facts study on geographic disparity which looks specifically at per pupil spending along with state educational rankings over time (USA Facts, 2021).

Further, Jackson et al. (2016) found that per pupil spending each year for all 12 years of public-school leads to 0.31 more completed years of education, an overall return of 7 percent higher wages, and a 3.2 percentage point reduction in the annual incidence of adult poverty. This would lead to the understanding that per pupil spending is vital to a student's education, especially when I consider those in the lower income brackets. This finding demonstrates that it is likely that states who are dependent upon federal education funds due to lower per pupil allocations would be more apt to be in compliance with any federal mandates.

The poverty of a state will make the state more likely to adopt the policy of value-added models for teacher evaluations.

Based on previous research, I should see states with high poverty rates more likely to adopt the policy of tying teacher pay/evaluations to student test scores. The influx of state funds attached to policy adoption should serve as an incentive for lower social economic states and help in an attempt to elevate student achievement.

In their study, Welch and Thompson (1980) looked at federal incentives that were used to ensure compliance by states and specifically at diffusion rates. They found that the federal government, through its fiscal incentives, could speed the rate at which policy

innovation occurred in the state. Further, incentives of direct fiscal aid are found to be more effective than indirect incentives in encouraging rapid early diffusion, but both types of federal incentives are found to be about equal in promoting diffusion through all states (Welch and Thompson, 1980). What this amounts to is that states that have higher overall poverty rates, are among the first to enact federal policies when they are tied to funding as doing so provides aid to an already burdened system.

The strength of the teachers' union in a state is negatively correlated with the likelihood that a state will adopt the educational policy.

Teacher unions have a vested interest in protecting their teachers. This may come in the form of pay negotiations, maximum hours worked, requirements levied by the state, etc. To effectively negotiate, the union needs to be seen as strong. This can be accomplished through a vast member base, strong leadership, and fiscal health. When a union is seen as weak, the state can impose more demands upon teachers without additional pay increases. The state would also be able to enact policy that may be seen as detrimental to the teacher's job security such as tying their evaluation to the test scores of their students. I propose that a strong union would be able to block such a policy from taking effect and ensure teacher job security.

Additionally, Hartley and Flavin (2011), Marianno (2020), and Jha et al. (2019) all found that teachers' unions, depending on their strength, were able to lobby for either the adoption or rejection of proposed education policies within the state legislature. This is a significant finding since this would help to explain why education policies within a

state may be adopted or fail to get traction on the floor of the state legislature which could further or halt policy adoption.

Data

State adoption of value-added models is visualized using data from the National Council for Teach Quality study which breaks down the states which adopted value-added model and the year in which adoption first took place (Teacher and principal evaluation policy, 2019).

The dependent variable or hazard rate $ADOPT_{i,t}$ is the conditional probability that state i will adopt the value-added model given that the state has not adopted a value-added model prior to year t (Berry and Berry, 1990). The hazard rate measures the likelihood of an event to fail (die) depending on the age it has reached. It predicts the amount of time until a certain event occurs, such as the death or failure of a value-added policies or component (Liberto, 2021). The predicted values of the dependent variable are in essence the probability that the policy is adopted. $ADOPT_{i,t}$ is measured as a dummy variable equaling one if state i adopts a value-added model in year t , zero otherwise (Berry and Berry, 1990). It is because of this that the model takes the form of a logistic regression. Data on state adoption dates of value-added models will be derived from the National Council on Teacher Quality (State-by-State Evaluation Timeline Briefs, 2018).

$BORDER_{i,t-1}$ is the term reflecting influence of bordering states and denotes the proportion of previously adopting neighboring states. In other words, the number of states sharing a border with state i that had adopted a value-added model prior to year t . The values were estimated by calculating the ratio of states sharing a border that had

previously adopted a value-added model. I examined the influence of bordering state adoption at zero-, one-, and two-year lags. However, it was found that the one and two-year lags did not produce substantively different results from a 0-year lag, so only the 0-year lag is retained for the reported model. Further, I followed the other education policy literature (McLendon et al., 2006; Berry and Berry, 1990) in dropping the first adopter (in this case, Florida) from the analysis to prevent bias. $BORDER_{i,t-1}$ is a continuous variable ranging from 0 to 1. A separate model was estimated that includes the border state variable so that the presence of this variable would not interfere with the presence of the variables for cue states and/or circuit court regions.

When considering “cue” states, I used the idea of states labeled as “pace-setting” states in previous studies conducted by Walker (1969) and Downs and Mohr (1979). The states reflected as cue states are recognized for first adopting Common Core within their region.

$CUE_{i,t-1}$ is the term reflecting whether a state qualifies as a signal state within the totality of the United States. Once the “cue” state adopts the value-added model, I should see this policy diffuse to other states in the following year. The four regions are the Northeast, Southeast, Midwest, and West. The cue states for each of these regions are New York, Kentucky, Oklahoma, and Nevada, respectively. I identified the $CUE_{i,t-1}$ states based on their early adoption of other educational policies and their reputation as innovative in educational policy (Walker, 1969). Cue is a continuous variable ranging from 0 to 1.

US courts of appeals, or circuit courts, were taken from USCourts.gov which delineates the twelve circuits and which states fall into each specific circuit (Court role and Structure, 2022). Only eleven circuits will be compared since the twelfth circuit of Washington D.C. will not be used and dropped from the data. States within the same circuit should share similar political ideologies and therefore adopt similar policies, so using the circuits as a regional measure should demonstrate diffusion among states within.

As with the border state model, a separate model was created for testing the hypothesis related to US courts of appeals so that there would not any interference with border or cue states. As with the previous model, ADOPT looks at the probability that a state will adopt the value-added model policy.

Ten binary variables were created to represent 10 of the 11 U.S. courts of appeals circuits included in the data. The base category not explicitly represented with a binary variable in the model is the 11th Circuit which contains Florida. This decision was made because Florida was the first state to adopt value-added models and serves as the region against which others are compared.

In order to determine the extent that states who first adopted performance-based funding would then be apt to adopt a further accountability policy of tying teacher pay to student test scores, or value-added models, data on state adoption dates will be derived from the National Council on Teacher Quality (State-by-State Evaluation Timeline Briefs, 2018).

$PBF_{i,t}$ is a binary variable with the value one representing adoption of performance-based funding and zero for no adoption. I also examine whether early adoption of performance-based funding serves as a predictor for value-added policy adoption by dividing the states into three categories of early, mid, and late adoption in the PBF wave variable.

In determining incentivization for targeted schools, I use data derived from the National Center for Education Statistics. The data was collected over the 2011-2012 school year and displays which states (a total of 20) enacted an incentivization program within the state to attract K-12 teachers to hard-to-staff schools/areas (State Education Practices SEP, 2012). $INCENTIVE_{i,t}$ is an ordinal variable ranging from zero to two that represents the degree of incentivized teacher pay at target schools. This is estimated by a combined value of targeted schools and hard to staff teaching assignment variables. A value of zero indicates that a state is neither hard to staff nor a target school, whereas a value of one represents one or the other, and a value of two indicates that a state is both targeted and hard to staff.

Data for adoption of Common Core policies is derived from the official Common Core standards website (Standards in your State, 2022). The site breaks down the policy by state and gives exact adopt data, and in some cases, repeal dates. $COMMON_{i,t}$ is a binary variable with a value of one indicating that a state adopted Common Core and a value of zero if a state did not.

When looking at Republican control, it will be broken down into two variables, one for governor and one for the legislature. The governorship variable is a binary where 1 indicates a Republican governor in office and zero indicates a Democratic governor.

The legislature variable represents party control of the state legislature by the Republican party. A coding of 1 indicates Republican control of both houses within the legislature and a coding of 0 indicates that there is not Republican control of both houses within the state. Nebraska, which is unicameral, will be coded as 1 for Republican control and 0 for not having Republican control of the legislature. The data for both variables will be taken from Gubernatorial and Legislative Party Control of State Government in Ballotopia (2019).

To understand which states have higher relative poverty rates, I will use census data from 2002-2019. The variable is the percentage of a state's population that is in poverty, ranging between 0 and 1.

The student spending variable takes into consideration how much each state spends on average per student. To accomplish this, I will use data reported by states regarding per pupil education spending to the USA Facts study on geographic disparity (Spending per student in k-12 public schools, 2021). Data collected from years 2011-2016 will be used. The data is reported from both state and local contributions. The variable represents state spending in thousands of dollars per pupil.

Depicting which states have a higher urban population, I will use census data from 2010. The variable population density is the proportion of a state's population living in an urban area, ranging from 0 to 1.

Teacher unions, if strong, can have substantial influence over the ability for a state to implement a policy that will affect teacher pay. The variable represents the relative power teacher unions have in a state. For this study, union strength, is an ordinal variable that ranges from 1 to 5, with 1 being a strong, robust union that has previously shown influence on policy and 5 being either an extremely weak, or the lack of a union, and therefore wields no influence on policy. Data for the variable was derived from a study conducted by Northern et al. (2012) which gave an overall ranking of state teachers' unions by strength and tier based up the following five criteria: 1) Resources and Membership, 2) Involvement in Politics, 3) Scope of Bargaining, 4) State Policies, and 5) Perceived Influence.

The sample used in the analysis includes the fifty states in the United States, but does not include the US territories, Puerto Rico, or Washington, DC.

Model Estimation Method

In this study, I will be using event history analysis which is the application of logistic regression in a time series model. The research centers around the reasoning behind why one state would adopt the value-added educational policy and considers potential adoption motivators. The probability of the policy adoption will be determined based on the events described in the hypotheses.

The dependent variable is the policy adoption of value-added models. The unit of analysis will be state-years. For event history analysis, the first step is to establish the first year any state adopted the educational policy of the value-added model for teacher evaluations (advent of the policy), which was Florida in 2011. With not all states having

currently adopted the policy, the end date will be 2016 since no additional states have adopted value-added models past this time. This gives a total time span of 5 years.

The methodology being employed is event history analysis. As explained by Berry and Berry (1990), Event History Analysis goal is to explain qualitative change that occurs in the behavior of an individual (or in this case a state). It is a longitudinal record showing if and when the event was experienced. In a discrete model, used in this study, the period of analysis is divided into distinct units (years) (Berry and Berry, 1990).

Event history analysis is frequently used by scholars to model policy adoption. I follow previous studies on education policy (Li, 2017) and policy diffusion (Walker, 1969; Gray, 1973; Berry and Berry, 1990, 1992, 2007), and use a Cox proportional-hazards model. I chose this nonparametric model because I had no expectations as to the probability of states adopting value added policies being time dependent. The dependent variable, then, is the hazard ratio of value-added policy adoption. For each model, I examined the Schoenfeld residuals and tested each model for violations of the Cox proportional hazards assumption (Fox, 2002; Liu, 2017).

CHAPTER 7

Results

In this chapter, I first describe the division of my three models and report the results of my regressions, predictive margins tests, and average marginal effects. I note areas of significance with my independent variables that test for the adoption of value-added model policies. Second, I follow this with further discussion that investigates the findings within the independent variables. This will add to the findings of previous literature and speak to potential reasons for policy adoption. Lastly, I discuss my conclusions regarding the findings of the quantitative analyses conducted in this chapter.

The time span of my data ranges from 2011 when value-added policies were first implemented until 2016 the year that the last state adopted the policy. To evaluate the hypotheses, I estimated eleven different models. Figure 8.1 below depicts the hazard ratio for the dependent variable (adoption of value-added models). The “steps” show the proportions of states that chose to adopt policy over the six-year time span, with many choosing to adopt within the first two years of policy roll-out. There is an even pattern of adoption for the subsequent four years. The first state to adopt a value-added model did not do so until two years after the Race to the Top grant program was created. By the end of the fourth year, more than half had done so, and by the end of the period under study only a handful of states had yet to adopt such a model.

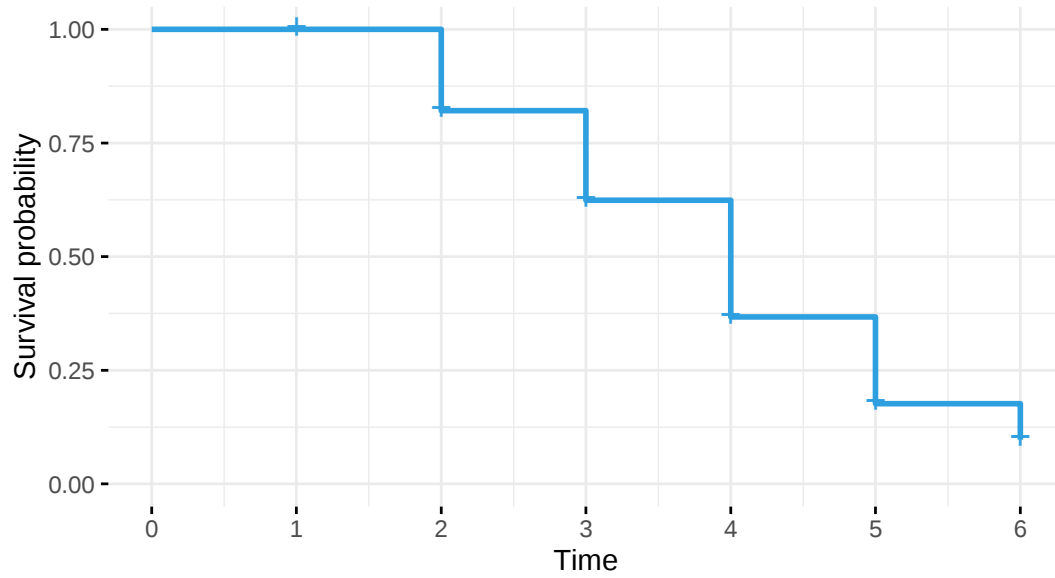


Figure 7.1
Kaplan Meier Estimate of
Diffusion

Results and Discussion

Model 1, reported in Table 7.1, tests the *cue state* variable that looks at which states tend to be first adopters of education policy innovations in their region. I use the Common Core policy adoption timeline to designate signal states in each region (New York, Kentucky, Oklahoma, and Nevada) and look at if there are influences in the behavior of others within their region. I also look at whether the timing of performance-based funding adoption influences state adoption.

Teacher incentives tests whether states that provide incentives to teachers such as additional pay would therefore be more likely to further adopt the incentivized program of value-added models. Lastly, model 1 tests whether having priorly adopted the policy of Common Core would influence states to adopt value-added models.

In the rest of the models, I chose to include states typically excluded from the policy literature; Hawaii, Alaska, and Nebraska, as I had no reason to assume that they would be excluded from the effects the model. Hawaii and Alaska, for instance, are often excluded because they are not contiguous, hence their exclusion in our estimation of geographic diffusion. Nonetheless, both states still have all the other characteristics I posit explain adoption. I argue it would be unreasonable in the digital age to discount states solely based on their geography. I also include Nebraska despite its unique unicameral legislature, as I argue many of those factors such as poverty and urbanization are exogenous. Moreover, even if policy is easy to pass, I still expect to see those same dynamics I expect in other states including the significance of party control and spending.

Table 7.1
Policy Indicators

	Model 1			
	(1)	(2)	(3)	(4)
Cue States	3.473 ^{***} (0.490)			
Performance Based Funding Adoption		1.345 ^{***} (0.354)		
Teacher Incentives			0.679 ^{***} (0.225)	
Common Core Adoption				2.613 ^{***} (0.680)
Teacher Union Strength	1.157 ^{***} (0.137)	1.171 ^{***} (0.136)	1.278 ^{***} (0.156)	1.278 ^{***} (0.157)
Republican State Legislature	0.546 (0.371)	0.595 (0.369)	0.603 (0.384)	0.610 (0.375)
Republican Governor	1.046 ^{***} (0.366)	0.926 ^{**} (0.361)	0.803 ^{**} (0.375)	1.005 ^{***} (0.372)
Population Density	0.0004 (6.810)	0.001 (7.174)	0.028 (7.244)	0.162 (7.601)
Percent Poverty	1.098 ^{***} (0.042)	1.080 ^{***} (0.040)	1.074 ^{***} (0.041)	1.070 ^{***} (0.042)
State Spending Per a Pupil	1.001 ^{***} (0.007)	1.007 ^{***} (0.007)	1.004 ^{***} (0.007)	1.009 ^{***} (0.007)
N	203	203	203	203
R ²	0.063	0.035	0.046	0.042
Max. Possible R ²	0.735	0.735	0.735	0.735
Log Likelihood	-128.338	-131.371	-130.188	-130.551
Wald Test (df = 7)	12.460 [*]	7.300	9.200	8.440
LR Test (df = 7)	13.208 [*]	7.142	9.507	8.782
Score (Logrank) Test (df = 7)	13.193 [*]	7.487	9.590	8.717
* p < .1; ** p < .05; *** p < .01				

¹ Model 1 reports hazard-ratios with standard errors.

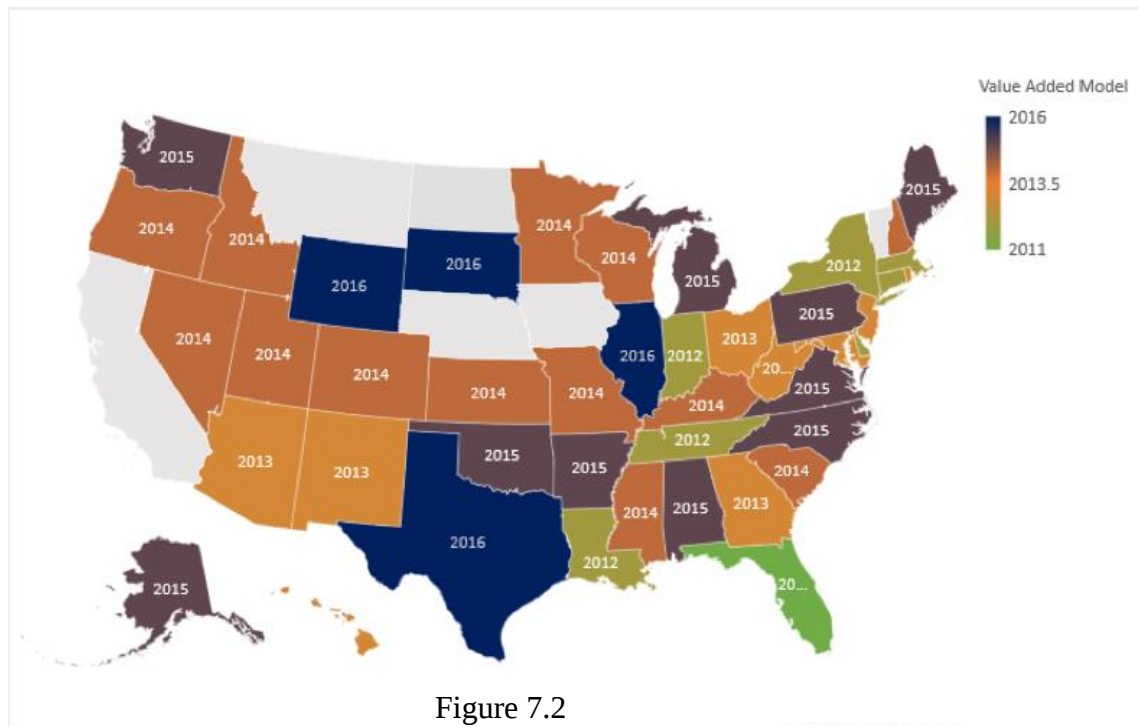


Figure 7.2
Value-Added Model
Adoption by State and Year

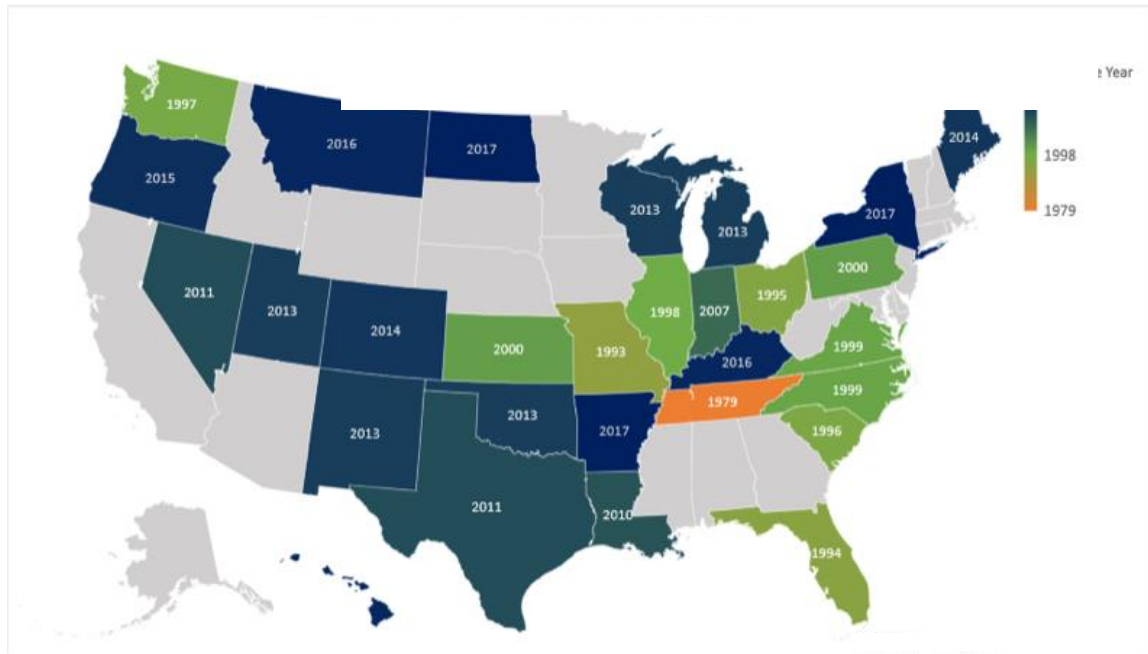


Figure 7.3
Performance-Based Policy
Adoption by State and Year

Region	States
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New York, New Jersey, Pennsylvania, Delaware
Southeast	Georgia, Florida, South Carolina, North Carolina, Tennessee, Alabama, Mississippi, Arkansas, Oklahoma, Louisiana, Texas, Maryland, Virginia, West Virginia, Kentucky
Midwest	Illinois, Indiana, Michigan, Ohio, Wisconsin, Iowa, Minnesota, Kansas, Missouri, Nebraska, North Dakota, South Dakota
West	California, Oregon, Washington, Nevada, Arizona, Colorado, Utah, New Mexico, Idaho, Montana, Wyoming

Figure 7.4
Region and State Break-up for Cue States

When testing the key hypotheses of my models with the just mentioned variables, I include in all of the models the following control variables: Republican control the legislature, Republican governorship, state education spending per pupil, percent of state population living in poverty, strength of teacher unions, and population density.

Internal variables for *teachers union*, *governor party*, *percentage poverty*, and *per pupil spending* all demonstrated statistical significance in the expected direction in the model. This is in line with previous policy diffusion literature.

One control variable, *spending per a pupil*, was slightly significant when testing for robustness. Though, as a whole the models were not. I followed Thernau et al. (2021) in re-estimating the models by transforming the variables and using a stepwise function. For all the models, I tested the assumption of time invariance and found that none of our models violated the assumption globally. However, both spending and cue states individually were almost significant. To ensure the robustness of the results, I re-estimated the models first with a step function and second with a continuous function to describe the time-varying coefficient. The purpose of the step function model is to split the analysis time into several intervals and stratify the Cox proportional model for these time intervals. This helps to account for any variation in the effects over time. The steps were chosen based on a graph of the Schoenfeld test for each variable. I also estimated the models by transforming the time-varying coefficients with a parametric continuous function. Both re-estimations had similar results to the primary models. The results from these were functionally the same. In addition, I followed Li (2017) in generating the Cox-

Snell residuals from the martingale residuals, which I then plotted against the Nelson-Aalen estimates to assess the fit of each model.

Table 7.1 shows the effects of policy diffusion by cue states, performance-based funding, teacher incentives and Common Core adoption. All of the variables tested were significant at the .01 level. *Cue states* is a binary variable with values of 0 and 1.

Performance-based funding is a binary variable with zero indicating a state did not adopt performance-based funding and a value of one indicating that a state did adopt. *Teacher incentives* is an ordinal variable ranging from zero to two. A value of zero indicates that a state is neither hard to staff nor a targeted school, whereas a value of one represents one or the other, and a value of two indicates that a state is both targeted and hard to staff. Lastly, *Common Core* is again binary variable with a value of one for adoption and zero for no adoption.

Cue states had the largest impact and increased the likelihood of adoption, making states surrounding cue states three times more likely to adopt the value-added model policy. States were grouped by larger regions, as seen in Figure 7.4, of northeast, southeast, midwest, and west. This demonstrates that states look to policy innovators within their region when considering which policies should be adopted and implemented within their own states. This is in line with expectations for hypothesis 1.

With states having previously adopted performance-based funding, the regression shows that they are 1.3 times more likely to adopt the value-added models. If I look at the heat maps in Figure 7.2 and 7.3. There is clustering of state policy adoption based upon region which further demonstrates the significance of the variable. Both the performance-

based funding and value-added models hold a form of accountability both institutionally and on the part of individuals. Having piloted this similar policy, it makes sense that states would be more likely to adopt. This is again in line with expectations for outcomes as described in hypothesis 4.

The finding for teacher incentive programs is interesting. According to the data, states are less likely to adopt value-added models if they have incentivized programs already in place. Both the incentivized programs and value-added models are a form of “reward” for teachers, with incentivized programs offering monetary benefits for agreeing to teach in either hard-to-staff or target schools and value-added models monetarily rewarding teachers for meeting testing objectives outlined by the state. It is therefore surprising that if one is adopted, the other is less likely to be adopted. This may be for a variety of reasons including policy overlap, lack of additional funds to implement both policies, or a re-focusing to have the “best” teachers in place first prior to implementing policies focused on student test scores. Though this finding is significant, as expected in hypothesis 5, there was not the expectation of a negative association.

Common Core adoption was a precursor to the Race to the Top initiatives. The policy adoption signalled an effort to comply with newly emerging federal education policies. What is found in the model is that federal initiatives did not coerce states to implementation of Common Core standards, however, a component of Race to the Top was the implementation of value-added models. It, therefore, would follow that by signaling a desire to be in compliance by adopting Common Core, the same states would also want to be in compliance with value-added policies. The data supports this

assumption and found that states who adopted Common Core were more than 2.5 times more likely to adopt value-added models. This finding does fall in line with the expectations of hypothesis 6.

When looking at the control variables, teacher union strength, governor party, percentage in poverty, and per pupil spending all demonstrated statistical significance across all models. The results are in line with findings from previous research (McLendon et al., 2006; Li, 2017; Miller and Morpew, 2017). The only difference between this study and previous literature, is that I did not find party control of a state legislature to play a role in policy adoption whereas previous research has reported such a relationship (McLendon et al., 2006; Li, 2017; Miller and Morpew, 2017). This may be accounted for the Republican governor variable, which again was significant, having an outsized role in what education policies were adopted, especially concerning policies such as value-added models that are tied to federal funds (McLendon et al., 2006).

In model 2, I examined the role of federal circuit court region on policy adoption. The expectation is that states within a circuit would have similar ideologies and therefore would mimic each other with the adoption of value-added models. The results are presented in table 7.2. There are 12 circuit courts and omitted the DC circuit, leaving 11 circuits for comparison. There are only 10 circuits listed as variables in the model, since one circuit court region was omitted to avoid falling into the dummy variable trap. With that being the case, all the coefficients indicate whether each of these regions differ in comparison with the base category which was circuit 11.

Table 7.2
Circuit Court
Regions

	Model 2
Circuit 1	0.767 (1.020)
Circuit 2	0.200 (1.339)
Circuit 3	3.250*** (1.068)
Circuit 4	0.451 (0.972)
Circuit 5	0.438 (1.050)
Circuit 6	0.786 (0.958)
Circuit 7	0.921 (1.054)
Circuit 8	0.114 (0.962)
Circuit 9	0.453 (0.881)
Circuit 10	0.457 (0.916)
Teacher Union Strength	1.502*** (0.217)
Republican State Legislature	0.637 (0.398)
Republican Governor	0.576 (0.403)
Population Density	0.853*** (0.095)
Percent Poverty	1.071*** (0.059)
State Spending Per a Pupil	1.004*** (0.009)
N	203
R ²	0.102
Max. Possible R ²	0.735
Log Likelihood	-124.038
Wald Test	19.230 (df = 16)
LR Test	21.808 (df = 16)
Score (Logrank) Test	21.020 (df = 16)
*p < .1; **p < .05; ***p < .01	

2

² Model 2 reports hazard-ratios with standard errors.

I find that, in general, circuit courts regions are not significant predictors of policy adoption. The only region to come up significant is that of the third circuit court region which contains the states of Pennsylvania, Delaware, and New Jersey. This coefficient indicates that states in the third circuit court region were on average more likely to adopt value-added models than were states in the 11th circuit court region (the latter includes Florida, the first state to adopt this policy). In contrast, being in other circuit court regions does not increase nor decrease the likelihood of adoption in comparison to states in the 11th circuit court region. This suggests that what was proposed in hypothesis 3 was incorrect. About 125,000 Pennsylvania residents work in New Jersey, and almost the same number of New Jersey residents work in Pennsylvania, according to U.S. Census data (McCrystal, 2018). This phenomenon creates two states that share a large number of residents and could influence the speed of policy diffusion for the region.

The results for the diffusion of value-added policies within federal circuit courts suggests that general regional diffusion does not occur between states in the broader sense. Instead, cue states and direct borders seem to better explain regional diffusion of the policy. Also, the variables representing the internal state characteristics of teachers' union strength, population density, proportion of population in poverty, and state education spending per pupil are all statistically significant in the expected directions in the model.

Model 3, reported in Table 7.3, estimated the relationship of geographic diffusion and value-added policy adoption. I measure geographic diffusion as the proportion of neighboring states that adopted value-added policies lagged by one year (Berry and Berry, 1990). I chose to lag this measure as I expected neighboring states would wait a

year to see how their neighbors fared after implementation. I exclude Alaska and Hawaii from this model as they are not contiguous with the other 48 states.

For hypothesis 2 regarding border states, I find the results to be in step with existing theories on policy diffusion and adoption. The values of the variable were estimated by calculating the ratio of states sharing a border that had previously adopted a value-added model. Additionally, 1-year and 2-year lags were tested and were found to not be significant. The results for the variable not lagged are reported in this table and is statistically significant, indicating that “border” states do influence the adoption of value-added models. However, this finding is in contrast with previous research which found a negative association with borders and performance-based funding (Li, 2017). Referring back to the results from the heat maps (Figure 7.2 and 7.3), at first, I do not see much in the way of diffusion in the east, but the further west I go on the map, the more diffusion is present, suggesting regional diffusion. There is evidence that regional influence on the probability of value-added policy adoption varies depending on the level of motivation to innovate. Overall, when a state is bordered by only states that have already added a value-added model, a state is 1.3 times more likely to adopt that policy as well. This is because it has often been found that states like to “wait and see” if the policy is a total failure before following their neighbors but do not like to wait too long for fear of missing out on the benefits (Berry and Berry, 1990).

When looking at the variables, what is interesting is that unlike previous literature that found statistical significance concerning gubernatorial and legislative state control influential in policy diffusion (McLendon et al., 2006; Berry and Berry, 1990), the results

for this study do not corroborate that finding. For this model and the value-added policy, I did not find that either political party control variables to be statistically significant in explaining the diffusion of the policy. This may be because the federal government is pulling the purse strings through Race to the Top initiatives and states are following suit based upon what their “neighbor” is adopting in the way of policy which other studies have shown (Howell and Magazinnik, 2017).

Conclusions

The overall results of the quantitative analysis of this research strongly support the argument that cue states, prior adoption of performance-based funding, prior adoption of Common Core standards, and the higher the proportion of bordering states that have adopted the policy all increase the likelihood that a state will adopt a value-added policy. Also, it is worth noting that though the prior adoption of teacher incentive policies is statistically significant in the reported models, it is that the previous adoption of incentives reduces the likelihood that a state will adopt a value-added policy. The explanation behind this finding is that states only have enough resources to focus on the implementation of a few policies within a policy domain, in this case education (Madsen, 1994; Paciotto and Delany-Barmann, 1994; Crampton, 2004). Thus, a state would lack the funds and potentially the manpower needed in order to implement both value-added models and teacher incentives.

Diffusion through cue states and border states allows for a greater understanding of a policy implementation, thus increasing one’s know-how competencies since,

Table 7.3
Border Diffusion

	Model 3
Borders	1.310** (0.588)
Teacher Union Strength	0.095 (0.144)
Republican State Legislature	-0.435 (0.394)
Republican Governor	-0.071 (0.390)
Population Density	-9.778 (7.243)
Percent Poverty	0.057 (0.045)
State Spending Per a Pupil	0.001 (0.008)
N	195
R ²	0.057
Max. Possible R ²	0.729
Log Likelihood	-121.513
Wald Test	12.310* (df = 7)
LR Test	11.429 (df = 7)
Score (Logrank) Test	12.624* (df = 7)
*p < .1; **p < .05; ***p < .01	

3

³ Model 3 reports hazard-ratios with standard errors.

essentially, it is “learning by doing” or experience (Rowland, 2020). Further reinforcing this argument is the finding that all cue states essentially serve as role models for their fellow states. Cue states serving as representatives should reasonably have more time and possibly more resources in respect to other states to become experts in the particular policy (value-added models) as they have had an extended amount of time to pilot and implement the policy.

Quantitative analysis of the pre-existing data on mechanisms for policy diffusion, though, offers mixed results. In this study, control variables for governorship and state legislative party control were found to be statistically significant in one model, but not the other. Results in Table 1, which depicted the role of the governor as statistically significant, still did not find the state legislature to effect policy adoption. When considering Table 3, which looked at the significance of border states, neither the role of the governor nor the state legislature was found to impact policy adoption. This may be due to the out the outsized role of either the governor for Table 1 or the federal government initiatives for Table 3; both of which hold the purse strings for implementing states and have coercive tactics at their disposal. Moreover, experience by the governor with prior policy may also contribute to their role in driving policy (McLendon et al., 2006).

Concerning the models themselves, it is worthy to note the analyses of this study further reinforce the greater predictive power of separating explanatory variables into various statistical models to demonstrate reasons for policy adoption. Border states may overlap and conflict with cue states and not demonstrate the full picture for the diffusion

effects. Thus, any statistically significant factors that increase the likelihood of adopting value-added models can be seen in conjunction with factors that showed significance in previous literature. These outcomes also supported many previous arguments that internal state factors also contribute to a state's ability to innovate (McLendon et al., 2006; Berry and Berry, 1990).

CHAPTER 8

Conclusion, Implications, and Further Work

This study began as an endeavor to understand the conditions under which state governments adopt new meritocratic policies such as value-added models for K-12 public education. I conceptualized the adoption of the policies to be a function of diffusion with regional influences playing key roles in the rates of policy adoption. Further, this study differs from previous research with a focus on demographic, economic, and political characteristics of the states (McLendon et al., 2006). Previous programs similar to value-added models have budget allocations directly tied to campus performance and held that campuses may lose funding if they “fail to perform adequately” (McLendon et al., 2006).

Additionally, previous research found that Republican-controlled states may favor meritocratic policies because these initiatives offer elected officials the strongest advantage for building up accountability pressure (McLendon et al., 2006; Beyle, 1999). In contrast, Democratic-controlled states, who tend to be less ideologically inclined to enact a government-accountability agenda (Baumgartner and Jones, 1993). Further, it has been found that a congressional minority will, after they themselves become the majority, seek to oversee more actively the programs of the opposition of the former minority (Hofferbert, 1966).

Building on these previous findings and using event history models on a panel dataset of 50 states in the US over 5 years, this paper seeks to investigate what influences the adoption of meritocratic policies such as value-added models by states in the context

of the Race to the Top legislation. Based on the review of the extant literature, I developed a conceptual model to empirically investigate this research question.

The study looks at six key variables: border influences, regional influences (including federal circuit court region and cue states), prior merit-based initiatives (such performance-based funding and teacher credential incentivization), and Common Core standards laying the groundwork for compliance with policy using a proverbial monetary carrot and stick.

This research emphasizes the need to examine education policy diffusion as not just a temporal process, but one made up of various stages (Ampaw and Jaeger, 2011). This implies that policy makers working with states to adopt new education policies should consider implementation issues and provide distinct support mechanisms at each stage. Additionally, the impact of federal incentives on the rate at which states adopt policy innovations had not previously examined. To combat this, one factor examined in this work is the time of policy invention. There is also a growing trend toward the nationalization (Welch and Thompson, 1980) of social, economic, and political policies.

This study found a positive correlation between states that have adopted prior performance-based funding initiatives that trickle-down from policies found in higher education to K-12 public schools. This is significant because once the door to diffusion is opened in one direction, we may find that policies trickle up from K-12 to higher education as well. This would mean the public school system is looked at more as a cohesive whole rather than parceled out pieces (individual school districts) as I currently characterize the system.

I also found that both cue and border states positively affected whether a state was more apt to adopt value-added models. This demonstrates a learning component that diffuses along with policy. Once a cue state has piloted a new education policy, border states can then learn from prior implementation and course correct any challenges in their own state at which point I should see further policy diffusion.

It was discovered that if a state adopted Common Core standards, they were more likely to adopt value-added models. The two policies build off each other with Common Core setting the groundwork for teacher standards and goals for each grade level benchmark and value-added models holding accountability on the part of the classroom teacher to ensure the standards and goals are adhered to and learning gains are met. It would, therefore, make sense that if state adopted one policy, they would become more apt to adopt policies that fall within the same domain.

States that adopted teacher incentive programs were shown to be less likely to adopt value-added models. This may be due to the five core issues discussed in previous chapters of reliability, validity, bias, transparency, and fairness (Amrein-Beardsley et al., 2020). This also may be due, in part, to budgetary restrictions within the state. Prior research has shown that as education becomes a lower priority within a state, there is both a reliance on federal dollars and less likelihood for innovation (Willner and Grönblom, 2009). It would therefore stand to reason that if a state has already implemented one policy that rewards teachers, such as incentives, then they would be constrained to be able to implement yet another through value-added models.

Though states within the same federal court circuit region may indeed at times share similar ideologies and goals (Broscheid, 2011), this does seem a common push to foster education policy innovation between the states. Instead, I found no correlation between states that were within the same federal circuit court region and policy diffusion. However, in line with prior research, I did find the control variable for party control of the governorship to be significant (McLendon et al., 2006), which gives credence to the robustness of my findings and builds upon what is previously understood in education policy diffusion.

This study has furthered the understanding of policy diffusion among states by empirically demonstrating that external determinants are important to understanding diffusion. Previous studies have focused on the internal makeup within states when explaining why a policy would diffuse. As discussed, previous findings demonstrated that factors such as partisanship within the state were critical to the understanding of why policies were adopted. However, this study showed that factors occurring outside the scope of internal determinants were just as critical. The results derived from the border states model illustrated this point, as no internal determinant controls were found to be significant. However, the imitation and normative pressure of bordering a state was significant and demonstrates the importance of looking at both factors internal and external to a state to have a full understanding of why policies diffuse.

Race to Top Coercion

A component not looked at within this study was coercion. This, however, is an important element to discuss since coercion could help to explain why states chose to first

adopt Common Core standards and later advance those standards with Race to the Top and with-it value-added models. A study conducted in 2017 looked specifically at the increase in policy adoption rates under Race to the Top (Howell and Magazinnik, 2017). They found that part of the explanation for this general increase in policymaking activity has to do with the structure of the competition itself, being an uncertainty about what dynamics would need to be put in place in order to win federal funds. States that chose to participate in the competition portion of Race to the Top devoted state resources to adopt policy reforms that the state thought would increase their chances in each round of the competition (Howell and Magazinnik, 2017). They do note that policy diffusion, which is looked at specifically in this study, also seems to be at least partly responsible for the national influence of Race to the Top (Howell and Magazinnik, 2017). They show that Race to the Top leveraged the forces of both “vertical and horizontal policy diffusion” in order to encourage legislative activity across states (Howell and Magazinnik, 2017).

Though the findings did not expressly point towards coercion of policy adoption through Race to the Top initiatives, in a study conducted by Howell and Magazinnik (2017) found that coercion through the use of monetary incentives did encourage many states to adopt value-added model policies. Howell and Magazinnik (2017) surveyed state legislators in the spring of 2014 and asked state legislators to reflect on the importance of Race to the Top for the education policy deliberations within their states. They also found that roughly one-third of legislators reported that Race to the Top had either a “massive” or “big” impact on education policymaking in their state. Additionally, 49 percent

reported that it had a “minor” impact, and only 19 percent claimed that it had no impact at all.

When considering the enactment of policies such as value-added models after the rollout of Race to the Top, it was found that states around the country enacted these policies at a much higher rate in the aftermath of Race to the Top than prior to its announcement (Howell and Magazinnik, 2017). It was found that between 2001 and 2008, states on average only enacted about 10 percent of these policies. In contrast, in the years between 2009 and 2014, 68 percent of the same types of policies were enacted and consequently adoption rates increased every single year thereafter (Howell and Magazinnik, 2017). According to the study, the rate of increase appears higher than what was predicted if you simply projected the slope of change observed between 2007 and 2008. By looking at preexisting trends, it would have taken states multiple decades to accomplish what, in the wake of Race to the Top, was accomplished in less than 5 years (Howell and Magazinnik, 2017).

The study found that even states that did not win additional funding through the initial Race to the Top competition, still had heightened policy activity. Howell and Magazinnik (2017) surmise that to improve their chances of winning additional funding, all participating states had incentives to adopt new education policies to demonstrate their willingness to comply with Race to the Top mandates and hopefully increase their chances of winning. They found that in a competition with private bid, but an undeclared number of winners, many participating states had incentives to seek higher scores by first

enacting some of the policies that Race to the Top touted (Howell and Magazinnik, 2017).

These findings are important not only because they demonstrate an increase in policy adoption, but in the aftermath of Race to the Top, policy adoption rates began to increase in all states regardless of funds rewarded (Howell and Magazinnik, 2017). Additionally, their study fits in with existing theories of policy diffusion and demonstrates further reasons for enactment of education policies. It is through better understanding of this coercion component, which is tied to federal purse strings, that there is a more complete picture of why certain educational policies, such as value-added models, are implemented within states.

Policy Entrepreneurs

In addition to looking at federal coercion, policy entrepreneurs also need to be looked at further as a supplement to this research on policy adoption. Policy entrepreneurs, defined as individuals whose actions promote dynamic policy change (McLendon et al., 2006), also need to be considered. In their study concerning policy change, Baumgartner and Jones (1993) found that the structural features of American federalism help shape policy trajectories, as previously suspected, but that individual actors also matter. Further, entrepreneurs facilitate change by redefining problems, refashioning policy images, developing potential solutions, and mobilizing (Baumgartner and Jones, 1993).

One study that was similar in nature used Event History Analysis along with survey data to examine the roles played by entrepreneurs in advancing school-choice

legislation (Mintrom, 1997). It was found that the presence and actions of entrepreneurs significantly raised the probability of a state enacting the reforms (Mintrom, 1997). In my study the policy entrepreneur would be the Obama Administration and their roll out of Race to the Top initiatives. The administration sought to build upon prior policy (No Child Left Behind) with additional meritocratic principles. By also tying federal education funds to the policy, the administration helped to facilitate the diffusion of the policy.

With the myriad of national organizations of policy specialists and governmental officials currently in existence, this may lead to a decrease in communication among the states, and in turn would lead to less opportunity for states to exchange policy ideas amongst themselves (Welch and Thompson, 1980). The example policy used within this study is skewed toward those that have completely diffused. Because of this, the sample is biased in favor of the fast diffusers. A spurious relationship might be introduced between time of invention and time of diffusion. (Welch and Thompson, 1980).

Adding to this, the federal government is partly responsible for the “time” effect in two ways. Over time the federal government has entered into more and more policy areas. Since the nineteenth century there has been a marked increase in the number of policies that have federal incentives (Welch and Thompson, 1980). Additionally, as the federal government has entered into so many policy arenas it has stimulated the development of “interstate organizations of public officials” within all levels of government. These organizations facilitate policy diffusion by providing additional opportunities for state officials to exchange ideas. Some may have been established

specifically to help with federal programs. It is because of this that the effects of the federal incentive and of the time of invention would be very difficult to distinguish (Welch and Thompson, 1980). This shows evidence of the influence of both internal determinants and regional diffusion which is seen in both the Welch and Thompson (1980) study as well as mine.

For my study, the internal determinants would comprise of variables that look at Common Core, performance-based funding, and teacher incentive adoption. Regional diffusion entails using cue and border states, along with a state belonging to specific federal circuit court region. It is the combination of these determinants that I have a full scope of why education policy may diffuse.

Accountability Policies and Where We are Headed

In addition to policy entrepreneurs being a catalyst for policy adoption, accountability policies may also further diffuse due to state compromise and internal determinants, along with compliance with federal mandates. There has been an emergence of these meritocratic policies (McLendon et al., 2006; Mintrom, 1997; Welch and Thompson, 1980) and that value-added model mandates are artifacts of failed reform efforts toward either accountability or budgeting. In other words, a political compromise becomes acceptable to both proponents and opponents of accountability and budgeting (McLendon et al., 2006).

The findings lend support to the general understanding in the literature that both state and federal influence to adopt an education policy is moderated by political partisanship and a state's region. I argue that federal coercion, along with the regional

and political contexts, matters in the adoption of meritocratic policies such as value-added models (Jha et al., 2019). The findings are consistent with the argument by some scholars that states having already implemented prior meritocratic policies are more likely to continue that trend. Therefore, it is important to separate the role of meritocratic practices from political and regional contexts in the adoption of policies. However, the results of the regressions do not conclusively confirm the assertion that regional influences play an outsized role in the adoption of reforms in public education. Further, the findings do not fully support the widely held view that “cue” states inform their counterparts on the initial adoption of meritocratic policies, and that they are politically aligned only with the Republican Party (Walker, 1969). The results however show that federal education funding could be important in policymakers’ efforts in reforming public education including the adoption of future meritocratic policies.

Additionally, in this research I saw performance-based funding (one of the first state and federal implementation efforts of tying funds to specified outcomes within education) first piloted in higher education. This eventually morphed into a variety of accountability measures, including value-added models, a few decades later in K-12. Accountability measures in K-12 were widely spread under the No Child Left Behind Act implemented in 2002. Such measures have been expanded upon under subsequent presidential administrations (Race to the Top with Obama and Every Student Succeeds Act under Trump). What I am currently finding is a further diffusion of these accountability policies found in K-12, such as learning targets, are being implemented in higher education.

The diffusion began as student learning outcome assessments and in 1997 many academic units were required to develop plans to assess student achievement in their undergraduate and graduate programs. In the Unit Assessment Plans all academic programs identified: (a) the process followed to develop the plan, (b) desired student outcomes, (c) measures and methods for assessing student achievement of the outcomes, and (d) plans to use the assessment results for program improvement (Matuga and Turos, 2018; Gartner et al., 2010; Drinka et al., 2005). This was further adopted by many universities in the early 2000's with many departments being required to report their assessment activities and any changes/improvements made in the units based upon their assessment results (Learning Outcomes, 2022). Using this data, those within higher education were encouraged to continue implementing their assessment plans. Follow up of these assessments did not regularly occur and some schools cited that it became obvious that without a recurring process that required assessment updates, such as a program review, formal student assessment activities get lost in administration changeovers, budgetary concerns, and other competing priorities (Student Outcomes, 2022).

The response in higher education currently to this monitoring problem is for many universities to ask their departments to implement and track on a yearly basis specific student outcomes based on courses taken, which included learning outcomes detailed by the professors and curriculum maps. Curriculum maps are used to see how the learning outcomes identified and developed are met within the program (Learning Outcomes, 2022).

Example of learning outcomes include (Learning Outcomes, page 1, 2022):

“Students will be able to communicate both orally and verbally about music of all genres and styles in a clear and articulate manner (comprehension). Students will be able to analyze and interpret texts within a written context (analysis). Students will be able to demonstrate an understanding of core knowledge in the course taught (application). Students will be able to judge the reasonableness of obtained solutions (evaluation). Students will be able to evaluate theory and critique research within the discipline (evaluation). Students will be able to work in groups and be part of an effective team (synthesis)”.

The achievement of outcomes will be measured through direct and indirect means. Direct measures require students to represent, produce or demonstrate their learning (Assessment Overview, 2022). Indirect measures capture information about students’ perceptions about their learning experiences and attitudes toward the learning processes (Assessment Overview, 2022).

These measures mirror standards that were implemented under Common Core by many states in K-12 beginning in 2010 (Read the Standards, 2022). An extension of this research paper would be to look at the diffusion between K-12 to higher education, considering factors that lead to the implementation of student achievement outcomes (an extension of accountability measures).

In sum, incentives provided by the federal government do help to stimulate the diffusion of policies through the states. Further, incentives of direct fiscal aid are found to be more effective than indirect incentives in stimulating quick early diffusion, but both

types of federal incentives are about equal in promoting diffusion throughout all states. (Welch and Thompson, 1980).

What Can be Further Explored

Little empirical work exists that comprehensively explores which incentives are used at the school district level, and the characteristics and demographics of the districts that offer incentives of different types. Moreover, limited attention has been paid to district policies being implemented heterogeneously across a single state with a high demand for specific kinds of teachers. The exception is work by Balter and Duncombe (2008), who use results from a survey of district superintendents to examine the use of compensation incentive policies in New York. They find that districts that use a range of recruitment practices hire more highly qualified teachers, and that larger districts tend to use a greater array of such practices. Aside from this work, the extant literature for the most part leaves us with scarce evidence about the prevalence of districts' use of economic incentive policies.

Additionally, this study focused on external determinants such as state groupings, cue states, and previous policies that could affect policy adoption. However, a more qualitative approach could also be taken to look at the push and pull that occurs between state officials and the federal government concerning the adoption of federal initiatives. This would focus on the interactive effect that occurs between leaders to see if policy mandates are simply handed down from the federal government or if states influence how the policy is created and eventually released through the initiatives.

Another external determinant to consider is the partisan role of the executive. Both No Child Left Behind and Race to the Top were federal initiatives put forth by two separate presidential administrations and under different partisan leadership. It can be further examined to see if this presidential partisanship helped to determine the diffusion and adoption of the given policies.

This study, like many others that are similar, focuses specifically on the correlates of policy creation and adoption, and not on the determinants of policy decline and death. Building upon this study, one would want to look at the theoretical, as well as methodological perspectives in order to best expand the general understanding of meritocratic policies along with other education policy reforms. The average lifespan of similar policies may be a place to begin. It is by expanding the scope of the research that I would broaden to encompass the full lifespan of an education policy, including its birth, maturation, decline, and death.

WORKS CITED

A Nation At Risk: Summary & Effects on Education. (2014, August 25). Retrieved from <https://study.com/academy/lesson/a-nation-at-risk-summary-effects-on-education.html>.

Abella, M. (2006). Global competition for skilled workers and consequences. In *Competing for Global Talent*, ed. C. Kuptsch and P.E. Fong, 11–32. Geneva: ILO.

Achieve. (2013, November 1). *Closing the expectations gap - files.eric.ed.gov*. Retrieved February 7, 2022, from <https://files.eric.ed.gov/fulltext/ED546620.pdf>

Allard, S. (2004). Competitive Pressures and the Emergence of Mothers' Aid Programs in the United States. *Policy Studies Journal*, 32(4), 521–544.
<https://doi.org/10.1111/j.1541-0072.2004.00079.x>

Allen, A. (2011). Michael Young's The Rise of the Meritocracy: A philosophical critique. *British Journal of Educational Studies*, 59(4), 367–382.

Ampaw, F. & Jaeger, A. J. (2012). Completing the Three Stages of Doctoral Education: An Event History Analysis. *Research in Higher Education*, 53(6), 640–660.
<https://doi.org/10.1007/s11162-011-9250-3>

Amrein-Beardsley, A. & Holloway, J. (2019). Value-Added Models for Teacher Evaluation and Accountability: Commonsense Assumptions. *Educational Policy (Los Altos, Calif.)*, 33(3), 516–542. <https://doi.org/10.1177/0895904817719519>

Andersen, P. K., Borgan, Gill, R. D., & Keiding, N. (1993). *Statistical models based on counting processes*. New York: Springer-Verlag.

Andersen, L. & Pallesen, T. (2008). “Not Just for the Money?” How Financial Incentives Affect the Number of Publications at Danish Research Institutions. *International Public Management Journal*, 11(1), 28–47. <https://doi.org/10.1080/10967490801887889>

Aragon, S. (2018, March 1). *Targeted teacher recruitment - your education policy team*. Retrieved January 24, 2022, from https://www.ecs.org/wpcontent/uploads/Targeted_Teacher_Recruitment.pdf

Armour-Garb, A. (2009). Should “value-added” models be used to evaluate teachers? *Journal of Policy Analysis and Management*, 28(4), 692–693.
<https://doi.org/10.1002/pam.20462>

Arra Education. ARRA Education | Federal Funds Information for States. (n.d.). Retrieved January 12, 2022, from <https://ffis.org/452501/453743.html>

Aslantas, Ismail (2021) How good are value-added measures of teacher performance? A review and empirical investigation using data from Turkey, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/14158/>

Assessment overview. Assessment Overview | Institutional Research & Decision Support. (n.d.). Retrieved January 18, 2022, from <https://irds.stanford.edu/assessment/assessment-overview>

Augoustinos, M., Tuffin, K., & Every, D. (2005). New racism, meritocracy and individualism: constraining affirmative action in education. *Discourse & Society*, 16(3), 315–340. <https://doi.org/10.1177/0957926505051168>

Balter, D. & Duncombe, W. D. (2008). Recruiting highly qualified teachers: Do district recruitment practices matter? *Public Finance Review*, 36(1), 33–62. <https://doi.org/10.1177/1091142106293949>

Baum, S. & McPherson, M. (2011, January 18). Is education a public good or a private good? *The Chronicle of Higher Education*, online blog [Innovations: Insights and commentary on higher education].

Baumgartner, F. R. & Jones, B. D. (2009). *Agendas and instability in American politics*. Chicago: University of Chicago Press.

Baybeck, B., Berry, W. D., & Siegel, D. A. (2011). A strategic theory of policy diffusion via intergovernmental competition. *The Journal of Politics*, 73(1), 232–247. <https://doi.org/10.1017/S0022381610000988>

Becker, H. (1993). *Human capital*. 3rd ed. Chicago: University of Chicago Press.

Becker, H. (2006). The age of human capital. In *Education, globalization and social change*, ed. H. Lauder, P. Brown, J.A. Dillabough, and A.H. Halsey, 292–4. Oxford: Oxford University Press.

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

Berdahl, R. O. (1971). *Statewide coordination of higher education*. Washington, DC: ACE.

Berry, W.D. (2017). Micro-credentials: the badges of professional growth. *The Education Digest*, 82(9), 21.

- Berry, F. S. & Berry, W. D. (1990). State lottery adoptions as policy innovations: An event history analysis. *The American Political Science Review*, 84(2), 395–415. <https://doi.org/10.2307/1963526>
- Berry, F. S. & Berry, W. D. (1992). Tax innovation in the states: Capitalizing on political opportunity. *American Journal of Political Science*, 36(3), 715–742. <https://doi.org/10.2307/2111588>
- Berry, F. S. & Berry, W. D. (2014). “Innovation and diffusion models in policy research.” Sabatier, & Weible, C. M.. *Theories of the policy process* (Third edition.). Westview Press.
- Berry, B. & Byrd, A. (2019, June 1). *Micro-credentials and education policy in the United States: Recognizing learning and leadership for our nation’s teachers*. VOCED plus, the international tertiary education and research database. Retrieved January 24, 2022, from <https://www.voced.edu.au/content/ngv:86800>
- Beyle, T. (1999). The governors. In V. Gray, R. Hanson, & H. Jacob (Eds.), *Politics in the American states: A comparative analysis* (pp. 191–231). Washington, DC: CQ Press.
- Bhatt, A., Zhang, Y., Davis, R., Eberle, A., & Heath, G. (2016). Economic implications of incorporating emission controls to mitigate air pollutants emitted from a modeled hydrocarbon-fuel biorefinery in the United States. *Biofuels, Bioproducts and Biorefining*, 10(5), 603–622. <https://doi.org/10.1002/bbb.1666>
- Blake, S., Tchoshanov, M., & Della-Piana, C. K. (2001). Report on the Partnership for Excellence in Teacher Education: an NSF funded project (NSF due-9343612). *Mathematics & Computer Education*, 35(3), 255–263.
- Block, F. (1987). *Revising state theory*. Philadelphia: Temple University Press.
- Blossfeld, H. P. & Rohwer, G. (2002). *Techniques of event history modeling: New approaches to causal analysis*. Hillsdale, NJ: Erlbaum.
- Blumenfeld, W. (2015, December 27). “Meritocracy” as propaganda lie and false assumption. The Good Men Project. Retrieved October 19, 2021, from <https://goodmenproject.com/featured-content/meritocracy-as-propaganda-lie-and-false-assumption-wcz/>.
- Bourdieu, P. & Passeron, J.-C. (1977). *Reproduction in education, society, and culture* (1990 ed. / preface by Pierre Bourdieu.). Sage in association with Theory, Culture & Society, Dept. of Administrative and Social Studies, Teesside Polytechnic

- Boushey, G. (2010). *Policy diffusion dynamics in America*. Cambridge: Cambridge University Press.
- Boushey, G. (2012). Punctuated equilibrium theory and the diffusion of innovations. *Policy Studies Journal*, 40(1), 127–146. <https://doi.org/10.1111/j.1541-0072.2011.00437.x>
- Box-Steffensmeier, J. M. & Jones, B. S. (2011). *Event history modeling: A guide for social scientists*. Cambridge University Press.
- Bowles, S. & Gintis, H. (1976) *Schooling in capitalist America*, London, Routledge and Regan Paul, pp 10, 53-4, 129, 130-3.
- Brehm, M., Imberman, S. A., & Lovenheim, M. F. (2017). Achievement effects of individual performance incentives in a teacher merit pay tournament. *Labour Economics*, 44, 133–150. <https://doi.org/10.1016/j.labeco.2016.12.008>
- Broscheid, A. (2011). Comparing circuits: Are some U.S. courts of appeals more liberal or conservative than others? *Law & Society Review*, 45(1), 171–194. <http://www.jstor.org/stable/23011962>
- Broucker, B. & DeWit, K. (2015). New public management in higher education. In J. Huisman et al. (Eds.), *The Palgrave international handbook of higher education policy and governance* (pp. 57–75). London: Palgrave Macmillan.
- Brown, P. & Tannock, S. (2009). Education, meritocracy and the global war for talent. *Journal of Education Policy*, 24(4), 377–392. <https://doi.org/10.1080/02680930802669938>
- Burke, J. C. & Minassians, H. P. (2001). *Linking resources to campus results: From fad to trend, the 5th annual survey*. Albany, NY: Rockefeller Institute of Government.
- Burke, J. C. (2002). *Funding public colleges and universities for performance*. Albany: Rockefeller Institute Press.
- Burke, J. C. (2005). Reinventing accountability. In J. C. Burke (Ed.), *Achieving accountability in higher education*. (pp. 216–245). San Francisco, CA: Jossey-Bass.
- Burke, J. & Minassians, H. (2003). Real accountability or accountability “lite”: Seventh annual survey, 2003. Albany, NY: Rockefeller Institute of Government.
- Burke, J. C. & Modarresi, S. (2000). To keep or not to keep performance funding: Signals from stakeholders. *The Journal of Higher Education*, 71, 432–453. doi:10.2307/2649147

Burke, J. C. & Serban, A. (1997). *State performance funding and budgeting for public higher education*. Albany, NY: State University of New York, Rockefeller Institute

Butler, L. (2020). *Michael Young, social science, and the British left, 1945-1970* (First edition.). Oxford University Press.

Castilla, E. (2008). Gender, race, and meritocracy in organizational careers. *The American Journal of Sociology*, 113(6), 1479–1526. <https://doi.org/10.1086/588738>

Canon, B. & Baum, L. (1981). Patterns of adoption of tort law innovations: An application of diffusion theory to judicial doctrines. *The American Political Science Review*, 75(4), 975–987. <https://doi.org/10.2307/1962297>

Chen, R. (2007) Socioeconomic and racial/ethnic gaps in student dropout risks: A longitudinal model to examine the differential effects of financial aid. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI

Chetty, R., Friedman, J., & Rockoff, J. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633–2679.

Chevalier, A., & Conlon, G. (2003). Does it pay to attend a prestigious university? IZA Discussion Paper 848, Bonn, Germany.

Chou, M., Moffitt, B., & Bryant, O. (2020). Political meritocracy and populism. In *Political Meritocracy and Populism* (1st ed., Vol. 1, pp. 14–32). Routledge. <https://doi.org/10.4324/9780429294846-2>

Curriculum development grants. Ohio Space Grant Consortium. (2020, July 23). Retrieved February 1, 2022, from <http://osgc.biz/cip.html>

Clotfelter, C., Ladd, H. F., & Vigdor, J. (2004). Teacher quality and minority achievement gaps. Durham, NC: Terry Sanford Institute of Public Policy.

Clotfelter, C., Ladd, H. F., Vigdor, J. L., & Wheeler, J. (2007). High-poverty schools and the distribution of teachers and principals. *North Carolina Law Review*, 85(5), 1345.

Clotfelter, C. T., Glennie, E., Ladd, H. F., & Vigdor, J. (2008). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *Journal of Public Economics* 92(5-6): 1352-1370.

Cohen-Vogel, L., Ingle, W. K., Levine, A. A., & Spence, M. (2007). The spread of merit-based college aid: Politics, policy consortia, and interstate competition. *Educational Policy*, 22(3), 339–362. doi:10.1177/0895904807307059.

Cole, M. (1988). *Bowles and Gintis revisited: Correspondence and contradiction in educational theory*. Falmer Press.

Collins, T. (1999). *Attracting and retaining teachers in rural areas*. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools.

Conant, J. (1961). *Slums and suburbs*. New York: McGraw-Hill.

Court role and structure. United States Courts. (n.d.). Retrieved January 12, 2022, from <https://www.uscourts.gov/about-federal-courts/court-role-and-structure>

Cox, D.R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B, Methodological*, 34(2), 187–220. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>

Crampton, F. (2004). State funding for education technology and school infrastructure: competing demands and limited resources. *Educational Considerations*, 31(2). <https://doi.org/10.4148/0146-9282.1248>

Croninger, R. G. & Valli, L. (2009). “Where is the action?” Challenges to studying the teaching of reading in elementary classrooms. *Educational Researcher*, 38(2), 100–108.

Crozier, G. (2018). Race and education: Meritocracy as white middle class privilege. *British Journal of Sociology of Education*, 39(8), 1239–1246. <https://doi.org/10.1080/01425692.2018.1523354>

Cyert, R. M. & March, J. G. (1963). *A behavioral theory of the firm*. Prentice-Hall.

Daley, D. & Garand, J. C. (2005). Horizontal diffusion, vertical diffusion, and internal pressure in state environmental policymaking, 1989-1998. *American Politics Research*, 33(5), 615–644. <https://doi.org/10.1177/1532673X04273416>

Davies, L. (2014). State “shared responsibility” policies for improved outcomes: Lessons learned. HCM Strategists. Retrieved from <http://hcmstrategists.com/wp-content/uploads/2014/04/HCM-State-Shared-Responsibility-RADD-2.0.pdf>

Davis, J. H., Schoorman, F. D., & Donaldson, L. (1997). Toward a stewardship theory of management. *Academy of Management Review*, 22(1), 20–47.

de Boer, H., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., Lemmens- Krug, K., & Vossensteyn, H. (2015). Performance-based funding and performance agreements in fourteen higher education systems. Enschede: University of Twente, Centre for Higher Education Policy Studies.

Dee, T. & Jacob, B. A. (2010). The impact of No Child Left Behind on students, teachers, and schools [with Comments and Discussion]. *Brookings Papers on Economic Activity*, 2010(2), 149–207. <https://doi.org/10.1353/eca.2010.0014>

Demirbag, M. & Glaister, K. W. (2010). Factors determining offshore location choice for R&D projects: A comparative study of developed and emerging regions. *Journal of Management Studies*, 47(8), 1534–1560. <https://doi.org/10.1111/j.1467-6486.2010.00948.x>

DiMaggio, P. & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160. <https://doi.org/10.2307/2095101>

Dougherty, K. & Natow, R. S. (2019). Performance-based funding for higher education: how well does neoliberal theory capture neoliberal practice? *Higher Education*, 80(3), 457–478. <https://doi.org/10.1007/s10734-019-00491-4>

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

Dougherty, K. J. & Reddy, V. (2013). Performance funding for higher education: What are the mechanisms? What are the impacts? *ASHE Higher Education Report*, Vol. 39, No. 2. Hoboken, NJ: Wiley.

Dougherty, K. J., Natow, R. S., & Vega, B. E. (2012). Popular but unstable: Explaining why state performance funding systems in the United States often do not persist. *Teachers College Record*, 114(3), 1–41.

Downs, G. (1976) *Bureaucracy, innovation, and public policy*. Lexington, MA: D. C. Heath.

Downs, G.W. & Rocke, D. M. (1995). *Optimal imperfection?: domestic uncertainty and institutions in international relations*. Princeton University Press.

Downs, G. W. & Mohr, L. B. (1979). Toward a theory of innovation. *Administration & Society*, 10(4), 379–408. <https://doi.org/10.1177/009539977901000401>

Doyle, W. R. (2006). Adoption of merit-based student grant programs: An event history analysis. *Educational Evaluation and Policy Analysis*, 28(3), 259–285.

Drinka, D., Voge, K., & Yen, M. Y.-M. (2005). From principles to practice: Analyzing a student learning outcomes assessment system. *Journal of Cases on Information Technology*, 7(3), 37–56. <https://doi.org/10.4018/jcit.2005070103>

Duffrin, E. (2011). What's the value in value-added? *The Education Digest*, 77(2), 46–49.

Duncan, A. (2009, July 4). U.S. Department of Education. Retrieved March 17, 2022, from <https://www.ed.gov/news/speeches/race-top-begins%3A?page=5>

Eggertsson, T. (1990). *Economic behavior and institutions*. Cambridge: Cambridge University Press.

Eisner, E. A. & Meier, K. J. (1990). Presidential control versus bureaucratic power: Explaining the Reagan revolution in antitrust. *American Journal of Political Science*, 34(1), 269–287. <https://doi.org/10.2307/2111519>

Enders, J., deBoer, H., & Weyer, E. (2013). Regulatory autonomy and performance: The reform of higher education re-visited. *Higher Education*, 65(1), 5–23. <https://doi.org/10.1007/s10734-012-9578-4>

"Evaluating the Effectiveness of Teacher Pay-for-Performance." RAND Corporation. Accessed April 03, 2019. <https://www.rand.org/capabilities/solutions/evaluating-the-effectiveness-of-teacher-pay-for-performance.html>.

Evaluation of the teacher incentive fund - futureed. (n.d.). Retrieved November 9, 2021, from <https://www.future-ed.org/work/evaluation-of-the-teacher-incentive-fund/>.

Feller, I. & Menzel, D. C. (1977). Diffusion milieus as a focus of research on innovation in the public sector. *Policy Sciences*, 8(1), 49–68. <https://doi.org/10.1007/BF01727601>

"Fiscal Year 2005 Budget Summary — February 2, 2004." FY 2005 ED Budget Summary: Summary. Accessed April 03, 2019. <https://www2.ed.gov/about/overview/budget/budget05/summary/edlite-section1.html>.

Fowler, L. (2020). Best practices for implementing federal environmental policies: a principal-agent perspective. *Journal of Environmental Planning and Management*, 63(8), 1453–1469. <https://doi.org/10.1080/09640568.2019.1670627>

Friedel, J. N., Thornton, Z. M., D'Amico, M. M., & Katsinas, S. G. (2013). Performance-based funding: The national landscape. University of Alabama Education Policy Center.

Frolich, N. (2011). Multi-layered accountability: Performance-based funding of universities. *Public Administration*, 89(3), 840–859.

Gardner, M. M., Hickmott, J., & Bresciani Ludvik, M. J. (2010). *Demonstrating student success : A practical guide to outcomes-based assessment of learning and development in student affairs*. Stylus Publishing, LLC.

Goldhaber, D. (2015). Exploring the potential of value-added performance measures to affect the quality of the teacher workforce. *Educational Researcher*, 44(2), 87–95.
<https://doi.org/10.3102/0013189X15574905>

Glick, D. M. & Friedland, Z. (2014). How often do states study each other? Evidence of policy knowledge diffusion. *American Politics Research*, 42(6), 956–985.
[doi:10.1177/1532673X14521658](https://doi.org/10.1177/1532673X14521658).

Gordon, R., Kane, T. J., & Staiger, D. O. (2006). Identifying effective teachers using performance on the job (Discussion paper 2006-01). Washington, DC: Brookings Institution.

Gray, D., Harkreader, S., & Wagar, D. (2001). Program review: Workforce development education program. Tallahassee: Florida Legislature, Office of Program Policy Analysis and Government Accountability.

Gray, V. (1973). Innovation in the states: A diffusion study. *The American Political Science Review*, 67(4), 1174–1185. <https://doi.org/10.2307/1956539>

Greenstein, D. (2015, February 12). Outcomes-based funding: Getting back to basics. Retrieved from <http://www.impatientoptimists.org/Posts/2015/02/OutcomesBased-Funding-Getting-Back-to-Basics>

Gross, P. K. (2008). Promoting or perturbing access: An event history analysis of the effects of financial aid on Latino students' educational attainment. Unpublished doctoral dissertation, Indiana University, Bloomington.

Grupp, F.W., & Richards, A. R. (1975). Variations in elite perceptions of American states as referents for public policy making. *The American Political Science Review*, 69(3), 850–858. <https://doi.org/10.2307/1958394>

"Gubernatorial and Legislative Party Control of State Government." Ballotpedia. Accessed April 16, 2019.
https://ballotpedia.org/Gubernatorial_and_legislative_party_control_of_state_government

- Haas, P. (1992). Introduction: epistemic communities and international policy coordination. *International Organization*, 46(1), 1–35.
<https://doi.org/10.1017/S0020818300001442>
- Research, H. (2014, August 1). Review of teacher incentive programs - email ... - CCRESA. Retrieved February 2, 2022, from <https://www.ccrea.net/wp-content/uploads/2014/12/Review-of-Teacher-Incentive-Programs-1.pdf>
- Hanushek, E. A. (1970). The value of teachers in teaching. Santa Monica, CA: RAND Corporation. Retrieved from <http://files.eric.ed.gov/fulltext/ED073089.pdf>
- Hanushek, E. A. (1979). Conceptual and empirical issues in estimating educational production function issues. *Journal of Human Resources*, 14, 351–388.
doi:10.2307/145575
- Hanushek, E. A. (2009). Teacher deselection. In D. Goldhaber & J. Hannaway (Eds.), *Creating a new teaching profession* (pp. 165–180). Washington, DC: Urban Institute Press.
- Harris, J. K. & Mueller, N. L. (2013). Policy activity and policy adoption in rural, suburban, and urban local health departments. *Journal of Public Health Management and Practice*, 19(2), E1–E8. doi: 10.1097/PHH.0b013e318252ee8c.
- Hartgen, D. T. (1973). Adoption-diffusion models of mode choice. Report PRR45, Planning and Research Bureau, Planning Division, Research and Applied Systems Section, New York State Department of Transportation.
- Hartney, M. & Flavin, P. (2011). From the schoolhouse to the statehouse: Teacher union political activism and U.S. state education reform policy. *State Politics & Policy Quarterly*, 11(3), 251–268. <https://doi.org/10.1177/1532440011413079>
- Hays, S. P. (1996). Influences on re-invention during the diffusion of innovations. *Political Research Quarterly*, 49(3), 631–650.
- Hess, F. & McShane, M. (2013). *Common core meets education reform. what it all means for politics, policy, and the future of schooling*. Teachers College Press.
- Hiler, T. (2018). The Absurd way I report higher education data. Student Aid Perspectives. https://www.nasfaa.org/newsitem/14547/Student_Aid_Perspectives_The_Absurd_Way_I_Report_Higher_Education_Data#_edn1. Accessed 8 June 2018.
- Hill, H., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, 48(3), 794–831.
<https://doi.org/10.3102/0002831210387916>

Ho, Li-Ching (2021) “Our students do not get that equal chance”: teachers’ perspectives of meritocracy, *Cambridge Journal of Education*, 51:2, 173-193, DOI: 10.1080/0305764X.2020.1796927

Hochschild, J. (1995). *Facing up to the American dream: Race, class, and the soul of the nation*. Princeton University Press.

Hoekstra, V. (2009). The Pendulum of precedent: U.S. state legislative response to supreme court decisions on minimum wage legislation for women. *State Politics & Policy Quarterly*, 9(3), 257–283. <https://doi.org/10.1177/153244000900900301>

Hoff, D. J. (2020, June 30). *The race to space rocketed NSF into classrooms*. Education Week. Retrieved January 31, 2022, from <https://www.edweek.org/teaching-learning/the-race-to-space-rocketed-nsf-into-classrooms/1999/05>

Hofferbert, R. I. (1966). The relation between public policy and structural and environmental variables in the American states. *American Political Science Review*, 60(1), 73–82.

Holyoke, T. T. (2003). Choosing battlegrounds: Interest group lobbying across multiple venues. *Political Research Quarterly*, 56(3), 325–336. <https://doi.org/10.1177/106591290305600307>

Hood, C. (2006). Gaming in target world: The targets approach to managing British public services. *Public Administration Review*, 66(4), 515–521.

Hook, J. G. & Cook, T. D. (1979). Equity theory and the cognitive ability of children. *Psychological Bulletin*, 86(3), 429–445. <https://doi.org/10.1037/0033-2909.86.3.429>

Hostetler, A., Sengupta, P., & Hollett, T. (2018). Unsilencing critical conversations in social studies teacher education using agent-based modelling. *Cognition and Instruction*, 36(2), 139–170.

Howell, W. G., Bio, W. G. H. A., & Bio, A. (2020, August 5). *Results of president Obama’s race to the top*. Education Next. Retrieved November 10, 2021, from <https://www.educationnext.org/results-president-obama-race-to-the-top-reform/>.

Howell, W.G. & Magazinnik, A. (2017). Presidential prescriptions for state policy: Obama’s race to the top initiative: Presidential prescriptions for state policy. *Journal of Policy Analysis and Management*, 36(3), 502–531. <https://doi.org/10.1002/pam.21986>

- Hunt, Carter, R., Zhang, L., & Yang, S. (2020). Micro-credentials: the potential of personalized professional development. *Development and Learning in Organizations*, 34(2), 33–35. <https://doi.org/10.1108/DLO-09-2019-0215>
- Illing, S. (2017, May 31). *To save the welfare state, liberals need a new narrative about personal responsibility*. Vox. Retrieved February 15, 2022, from <https://www.vox.com/conversations/2017/5/31/15680726/welfare-state-responsibility-liberalism-conservatism-entitlement-health-care>
- Illing, S. (2019, October 21). *How meritocracy harms everyone - even the winners*. Vox. Retrieved October 21, 2021, from <https://www.vox.com/identities/2019/10/21/20897021/meritocracy-economic-mobility-daniel-markovits>.
- Ingersoll, R. & Merrill, L. (2010). Who’s teaching our children? *Educational Leadership*, 67(8), 14–20.
- Ishitani, T. T. & DesJardins, S. L. (2002). A longitudinal investigation of dropout from colleges in the United States. *Journal of College Student Retention*, 4(2), 173–201. doi:10.2190/V4EN-NW42-742Q-2NTL.
- Jackson, C., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, 131(1), 157–218. <https://doi.org/10.1093/qje/qjv036>
- Jenkins, D., Ellwein, T., & Boswell, K. (2009). Formative evaluation of the Student Achievement Initiative “learning year”. New York: Columbia University, Teachers College, Community College Research Center Retrieved from <https://ccrc.tc.columbia.edu/publications/student-achievement-initiative-learning-year.html>.
- Jha, N., Banerjee, N., & Moller, S. (2019). Assessing the role of teachers’ unions in the adoption of accountability policies in public education. *The Urban Review*, 52(2), 299–330. <https://doi.org/10.1007/s11256-019-00529-y>
- Jochim, A. & Lavery, L. (2015). The evolving politics of the common core: policy implementation and conflict expansion. *Publius*, 45(3), 380–404. <https://doi.org/10.1093/publius/pjv015>
- Johnston, R.C. (2006). Teacher bonuses get high priority. *Education Week*, 25(20), 18–18.

- Jones, G. A. (2004). Ontario higher education reform, 1995–2003: From modest modifications to policy reform. *Canadian Journal of Higher Education/Revue Canadienne d'Enseignement Supérieure*, 34(3), 39–54.
- Jorgensen, M., & Hoffman, J. (2003), (rep.). http://images.pearsonclinical.com/images/PDF/assessmentReports/History_of_NC_LB.pdf.
- Jost, J. T. & Hunyady, O. (2005). Antecedents and consequences of system-justifying ideologies. *Current Directions in Psychological Science: A Journal of the American Psychological Society*, 14(5), 260–265. <https://doi.org/10.1111/j.0963-7214.2005.00377.x>
- Karch, A. (2006). National intervention and the diffusion of policy innovations. *American Politics Research*, 34(4), 403–426. <https://doi.org/10.1177/1532673X06288202>
- Kaufman, H. (1960) *The forest ranger: A study in administrative behavior*. Baltimore: Johns Hopkins University Press.
- Kelchen, R. (2018). *Higher education accountability*. Baltimore: Johns Hopkins University Press.
- Key, V. O. (1949). *Southern politics in state and nation* ([1st ed.]). A.A. Knopf.
- Kingdon, J. W. (1984). *Agendas, alternatives, and public policies*. Boston, MA: Little Brown.
- Klein, A. (2020, December 4). *Race to the top's impact on student achievement, state policy unclear, report says*. Education Week. Retrieved November 10, 2021, from <https://www.edweek.org/policy-politics/race-to-the-tops-impact-on-student-achievement-state-policy-unclear-report-says/2016/10>.
- Kolbe, T. & Strunk, K. O. (2010). Economic incentives as a strategy for responding to teacher staffing problems: An organizational typology. Presented at the Annual Conference of the American Education Finance Association. Richmond, VA.
- Kolbe, T. & Strunk, K. O. (2012). Economic incentives as a strategy for responding to teacher staffing problems: A typology of policies and practices. *Educational Administration Quarterly*, 48(5), 779–813. <https://doi.org/10.1177/0013161X12441011>
- Koretz, D. M. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *The Journal of Human Resources*, 37(4), 752–777.

- Kivisto, J. A. (2007). Agency theory as a framework for the government-university relationship. Tampere: University of Tampere Available from: <http://tampub.uta.fi/bitstream/handle/10024/67724/978-951-44-6969-5.pdf;sequence=1>.
- LaMagna, M. (2017). Placing digital badges and micro-credentials in context. *Journal of Electronic Resources Librarianship*, 29(4), 206–210. <https://doi.org/10.1080/1941126X.2017.1378538>
- Lane, J. E. & Kivisto, J. A. (n.d.). Interests, information, and incentives in higher education: Principal-agent theory and its potential applications to the study of higher education governance. In *Higher Education* (pp. 141–179). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6959-8_5
- Lave, C. A., & March, J. G. (1993). *An introduction to models in the social sciences*. Lanham, MD: University Press of America. (Reprinted from 1975.)
- Lazear, E. P. (1999). *Economic imperialism*. Boston, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w7300.pdf>
- Layzell, D. (1999). Linking performance to funding outcomes at the state level for public institutions of higher education: Past, present, and future. *Research in Higher Education*, 40(2), 233–246. <https://doi.org/10.1023/A:1018790815103>
- Learning outcomes and assessment methods*. Missouri State University. (n.d.). Retrieved January 18, 2022, from <https://www.missouristate.edu/assessment/learning-outcomes-and.htm>
- Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, 31, 3-12.
- Lee, J. & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990—2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, 34(2), 209–231. <https://doi.org/10.3102/0162373711431604>
- Levy, J. (1994). Learning and foreign policy: sweeping a conceptual minefield. *International Organization*, 48(2), 279–312. <https://doi.org/10.1017/S0020818300028198>
- Li, A. (2017). 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>
- Liberto, D. (2021, May 19). *Hazard rate definition*. Investopedia. Retrieved February 16, 2022, from <https://www.investopedia.com/terms/h/hazard-rate.asp>

- Light, A. (1978). Intergovernmental sources of innovation in state administration. *American Politics Research*, 6(2), 147–166.
<https://doi.org/10.1177/1532673X7800600203>
- Littler, J. (2018). *Against meritocracy: Culture, power and myths of mobility*. Routledge, an imprint of the Taylor & Francis Group.
- Lockwood, J. R., McCaffrey, D. F., Mariano, L. T., & Setodji, C. M. (2007). Bayesian methods for scalable multivariate value-added assessment. *Journal of Educational and Behavioral Statistics*, 32(2), 125–150.
- Lowry, W. (2005). Policy reversal and changing politics: State governments and dam removals. *State Politics & Policy Quarterly*, 5(4), 394–419.
<https://doi.org/10.1177/153244000500500405>
- Madsen, J. (1994). *Educational reform at the state level: The politics and problems of implementation*. Falmer Press.
- Marginson, S. (2009). *The limits of market reform in higher education*. Research Institute for Higher Education, Hiroshima University. Melbourne: University of Melbourne.
 Available from: https://melbournecshe.unimelb.edu.au/_data/assets/pdf_file/0009/1491660/RIHE_HigherEducationForum_Marginson_0909.pdf.
- Marianno, B. (2020). Compared to what? Changes in interest group resources and the proposal and adoption of state teacher policy. *Policy Studies Journal*, 48(4), 982–1022.
<https://doi.org/10.1111/psj.12345>
- Markovits, D. (2020). *The meritocracy trap: How America's foundational myth feeds inequality, dismantles the middle class, and devours the elite*. Penguin Books, an imprint of Penguin Random House LLC.
- Matuga, J. & Turos, J. M. (2018). The promise of an all-in-one assessment system for student learning outcomes: bowling green state university's excellence in assessment designation (EIA). *Assessment Update*, 30(1), 4–15. <https://doi.org/10.1002/au.30120>
- McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND.
- McCrystal, L. (2018, October 29). *How much does it cost - or save - workers to commute across state lines?* <https://www.inquirer.com>. Retrieved February 9, 2022, from <https://www.inquirer.com/philly/news/philadelphia-new-jersey-delaware-income-taxes-commuting-across-state-lines-20181029.html-2>

McDonald, J. (2014, December 4). *Handbook of biological statistics*. Independence - Handbook of Biological Statistics. Retrieved February 16, 2022, from <http://www.biostathandbook.com/independence.html>

McGuinn, P. (2012). Stimulating reform: Race to the top, competitive grants and the Obama education agenda. *Educational Policy (Los Altos, Calif.)*, 26(1), 136–159. <https://doi.org/10.1177/0895904811425911>

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

McLendon, M., 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/10.3102/01623737028001001>

McMurray, R. (2019, February 14). *Don't tie teacher merit pay to test scores*. The Eagle. Retrieved March 17, 2022, from https://theeagle.com/opinion/columnists/dont-tie-teacher-merit-pay-to-test-scores/article_7d48a72e-2f40-11e9-b02a-2bb50ddcce74.html

McNamee, & Miller, R. K. (2013). *The meritocracy myth*. (3rd ed.). Rowman & Littlefield Publishers.

Menand, L., Heller, N., & Hsu, H. (2019, September 17). *Is meritocracy making everyone miserable?* The New Yorker. Retrieved October 21, 2021, from <https://www.newyorker.com/magazine/2019/09/30/is-meritocracy-making-everyone-miserable>.

Meroe, A. (2014). Democracy, meritocracy and the uses of education. *The Journal of Negro Education*, 83(4), 485–498. <https://doi.org/10.7709/jnegroeducation.83.4.0485>

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

Mills, C., & Keddie, A. (2012). “Fixing” student deficit in contexts of diversity: Another cautionary tale for pre-service teacher education. *International Journal of Pedagogies and Learning*, 7(1), 9–19.

Mintrom, M. (1997). Policy entrepreneurs and the diffusion of innovation. *American Journal of Political Science*, 41, 738–770.

Mintrom, M. & Vergari, S. (1996). Advocacy coalitions, policy entrepreneurs, and policy change. *Policy Studies Journal*, 24(3), 420–434. <https://doi.org/10.1111/j.1541-0072.1996.tb01638.x>

Moe, T. (1984). The new economics of organization. *American Journal of Political Science*, 28(4), 739–777. <https://doi.org/10.2307/2110997>

Mohr, L. (1969). Determinants of innovation in organizations. *The American Political Science Review*, 63(1), 111–126. <https://doi.org/10.1017/S0003055400261510>

Mooney, C. Z., & Lee, M. H. (1995). Pre-Roe abortion regulation reform in the U.S. states. *American Journal of Political Science*, 39, 599–627.

Mooney, C. Z. (2001). Modeling regional effects on state policy diffusion. *Political Research Quarterly*, 54(1), 103–124.

Murnane, R. & Olsen, R. J. (1989). The effects of salaries and opportunity costs on duration in teaching: evidence from Michigan. *The Review of Economics and Statistics*, 71(2), 347–352.

Murphy, E J. & DeArmond, M. M. (2003). From the headlines to the frontlines: The teacher shortage and its implications for recruitment policy. Seattle, WA: Center on Reinventing Public Education, University of Washington.

Naidoo, R., Shankar, A., & Veer, E. (2011). The consumerist turn in higher education policy: Policy aspirations and outcomes. *Journal of Marketing Management*, 27(11–12), 1142–1162.

National Center for Education Evaluation and regional assistance: Evaluation of the Teacher Incentive Fund: Final report on implementation and impacts of pay-for-performance across four years. (n.d.). *NCEE publications: Evaluation of the teacher incentive fund: Final report on implementation and impacts of pay-for-performance across four years*. Institute of Education Sciences (IES) Home Page, a part of the U.S. Department of Education. Retrieved November 9, 2021, from <https://ies.ed.gov/ncee/pubs/20184004/>.

National Center for Education Statistics. (2017). “Undergraduate retention and graduation rates.” *The Condition of Education*, https://nces.ed.gov/programs/coe/indicator_ctr.asp. Accessed 14 September 2021.

Nationsonline.org, K. K. (n.d.). *Map of the United States*. Nations Online Project. Retrieved February 7, 2022, from https://www.nationsonline.org/oneworld/map/usa_map.htm

- NCLB. New America. (n.d.). Retrieved November 9, 2021, from <https://www.newamerica.org/education-policy/topics/federal-education-legislation-budget/federal-education-legislation/essa/nclb/>.
- NCTQ: Publications: State-by-state evaluation timeline briefs. National Council on Teacher Quality (NCTQ). (n.d.). <https://www.nctq.org/publications/State--by--State-Evaluation-Timeline-Briefs>.
- NCTQ state of the States 2019: Teacher and principal evaluation policy. National Council on Teacher Quality (NCTQ). (n.d.). Retrieved September 14, 2021, from <https://www.nctq.org/pages/State-of-the-States-2019:-Teacher-and-Principal-Evaluation-Policy>.
- Nelson, R. R. & Winter, S. G. (1974). "Neoclassical vs. evolutionary theories of economic growth: Critique and prospectus," *Economic Journal*, Vol. 83 (December), 886-905.
- Ness, E. C., Deupree, M.M., & Gandara, D. (2015). Campus responses to outcomes-based funding in Tennessee: Robust, aligned, and contested. Nashville: Tennessee Higher Education Coordinating Board Retrieved from <http://www.denisagandara.com/docs/FordFoundationPaper.pdf>
- Nishiyama, H. & Yamaguchi, M. (2013). Technological constraints, firm heterogeneity, and location choice of multinational enterprises. *Review of International Economics*, 21(5), 996–1005. <https://doi.org/10.1111/roie.12085>
- Northern, A., Scull, J., & Zeehandelaar Shaw, D. (2012, October 29). *How strong are U.S. teacher unions? A state-by-state comparison*. The Thomas B. Fordham Institute. Retrieved March 18, 2022, from <https://fordhaminstitute.org/national/research/how-strong-are-us-teacher-unions-state-state-comparison>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–257.
- Olcott, Don (2022). Micro-Credentials: A catalyst for strategic reset and change in U.S. higher education. *American journal of distance education*: 36(1), 19-35.
- Onosko, J. (2011). Race to the top leaves children and future citizens behind: The devastating effects of centralization, standardization, and high stakes accountability. *Democracy & Education*, 19(2), 1-11.
- Opfer, D. (2011). Defining and identifying hard-to-staff schools: The role of school demographics and conditions. *Educational Administration Quarterly*, 47(4), 582–619. <https://doi.org/10.1177/0013161X11400598>

Ortagus, J., Kelchen, R., Rosinger, K., & Voorhees, N. (2020). Performance-based funding in American higher education: A systematic synthesis of the intended and unintended consequences. *Educational Evaluation and Policy Analysis*, 42(4), 520–550. <https://doi.org/10.3102/0162373720953128>

Pacelle, R. (2018). Of political principals and legal principles: The solicitor general of the United States. In *Routledge Handbook of Judicial Behavior* (1st ed., pp. 196–220). Routledge. <https://doi.org/10.4324/9781315691527-13>

Pacheco, J. (2012). The social contagion model: Exploring the role of public opinion on the diffusion of antismoking legislation across the American states. *The Journal of Politics*, 74(1), 187–202. <https://doi.org/10.1017/S0022381611001241>

Paciotto, C. & Delany-Barmann, G. (2011). Planning micro-level language education reform in new diaspora sites: two-way immersion education in the rural Midwest. *Language Policy*, 10(3), 221–243. <https://doi.org/10.1007/s10993-011-9203-x>

Paige, M., Amrein-Beardsley, A., & Close, K. (2019). *Tennessee's national impact on teacher evaluation law & policy: An assessment of value-added model litigation*.

Papay, J. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*, 48(1), 163–193. <https://doi.org/10.3102/0002831210362589>

Pralle, S. (2003). Venue shopping, political strategy, and policy change: The internationalization of Canadian forest advocacy. *Journal of Public Policy*, 23(3), 233–260. <https://doi.org/10.1017/S0143814X03003118>

Prince, C. (2003). *Higher pay in hard-to-staff schools: the case for financial incentives*. Scarecrow Press.

Podgursky, M. (2006). Is there a “qualified teacher” shortage? *Education Next*, 6(2), 26–32.

Pusser, B. (2008). The state, the market and the institutional estate: Revisiting contemporary authority relations in higher education. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 23, pp. 105–139). Dordrecht: Springer.

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

- Rasmussen, T. (2000). State regulatory principals and local bureaucratic agents: The politics of local solid waste management. *American Review of Public Administration*, 30(3), 292–306. <https://doi.org/10.1177/02750740022064687>
- Reay, D. (2020). The perils and penalties of meritocracy: Sanctioning inequalities and legitimating prejudice. *The Political Quarterly*, 91(2), 405–412.
- Read the standards. Read the standards | Common Core state standards initiative. (2022). Retrieved January 18, 2022, from <http://www.corestandards.org/read-the-standards/>
- Reed, D., Rueben, K. S. and Barbour, E. (2006). *Retention of new teachers in California*. San Francisco, CA: Public Policy Institute of California.
- Regens, J. (1980). State policy responses to the energy issue: An analysis of innovation. *Social Science Quarterly*, 61(1), 44–57.
- Rein, M. & Schön, D. (1996). Frame-critical policy analysis and frame-reflective policy practice. *Knowledge and Policy*, 9(1), 85–104. <https://doi.org/10.1007/BF02832235>
- Review of recommendations for NSF project management reform (2017). House report 114-884 - report of activities of the Committee on Science, Space, and Technology U.S. House of Representatives together with minority views for the One Hundred Fourteenth Congress. (n.d.). Retrieved March 17, 2022, from <https://www.govinfo.gov/content/pkg/CRPT-114hrpt884/html/CRPT-114hrpt884.htm>
- Rhoades, G., & Sporn, B. (2002). Quality assurance in Europe and U.S.: Professional and political economic framing of higher education policy. *Higher Education*, 43(3), 355–390.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, 94(2), 247–252.
- Rogers, E. & F. Shoemaker (1971) *Communication of Innovations*. New York: Free Press.
- Ronfeldt, M., Loeb, S., & Wyckoff, J. (2012). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1), 4–36.
- Rowland, Erin. (2021). Who mediates matters - The impact of individual mediator competency skills on international and civil conflict mediation. [Unpublished doctoral dissertation]. University of Tennessee, Knoxville.

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

Sandel, M. (2021). *Tyranny of merit: What's become of the common good?* Penguin Books.

Sanders, W. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education*, 14(4), 329–339.

Seashore Louis, K. & Robinson, V. M. (2012). External mandates and instructional leadership: school leaders as mediating agents. *Journal of Educational Administration*, 50(5), 629–665. <https://doi.org/10.1108/09578231211249853>

Schulze-Cleven, T., & Olson, J. R. (2017). Worlds of higher education transformed: Toward varieties of academic capitalism. *Higher Education*, 73(6), 813–831

Science & Engineering State Indicators. State support for higher education per full-time equivalent student | State Indicators | National Science Foundation - State Indicators. (2020, December 31). Retrieved March 17, 2022, from <https://nces.nsf.gov/indicators/states/indicator/state-support-for-higher-education-per-fte-student>

Scott, M. & Kennedy, B. B. (2005). Pitfalls in pathways: Some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics*, 30(4), 413–442. <https://doi.org/10.3102/10769986030004413>

Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, 52(4), 840–857. doi:10.1111/j.1540-5907.2008.00346.x.

Shulock, N., & Snyder, M. (2013, December 5). Don't dismiss performance funding. Inside Higher Ed. Retrieved from <https://www.insidehighered.com/views/2013/12/05/performance-funding-isnt-perfect-recent-study-shortchanges-it-essay>

Simmons, B., Dobbin, F., & Garrett, G. (2006). Introduction: The international diffusion of liberalism. *International Organization*, 60(4), 781–810. <https://doi.org/10.1017/S0020818306060267>

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York: Oxford University Press.

- Son Hing, L., Bobocel, D. R., Zanna, M. P., Garcia, D. M., Gee, S. S., & Oraziatti, K. (2011). The merit of meritocracy. *Journal of Personality and Social Psychology*, 101(3), 433–450. <https://doi.org/10.1037/a0024618>
- Souto-Otero, M. (2010). Education, meritocracy and redistribution. *Journal of Education Policy*, 25(3), 397–413. <https://doi.org/10.1080/02680930903576396>
- Snyder, M. (2015). Driving better outcomes: Typology and principles to inform outcomes-based funding. Washington, DC: HCM Strategists. Retrieved from <http://hcmstrategists.com/drivingoutcomes/wpcontent/themes/hcm/pdf/Driving%20Outcomes.pdf>
- Spruance, S., Reid, J. E., Grace, M., & Samore, M. (2004). Hazard ratio in clinical trials. *Antimicrobial Agents and Chemotherapy*, 48(8), 2787–2792. <https://doi.org/10.1128/AAC.48.8.2787-2792.2004>
- Standards in your State. Standards in Your State | Common Core state standards initiative. (n.d.). Retrieved January 12, 2022, from <http://www.corestandards.org/standards-in-your-state/>
- State Education Practices (SEP). National Center for Education Statistics (NCES) Home Page, a part of the U.S. Department of Education. (n.d.). Retrieved January 11, 2022, from https://nces.ed.gov/programs/statereform/tab3_12.asp
- Statistical groupings of states and counties - census.gov. (n.d.). Retrieved February 9, 2022, from <https://www2.census.gov/geo/pdfs/reference/GARM/Ch6GARM.pdf>
- Steinhübel, L., Wegmann, J., & Mußhoff, O. (2020). Digging deep and running dry—the adoption of borewell technology in the face of climate change and urbanization. *Agricultural Economics*, 51(5), 685–706. <https://doi.org/10.1111/agec.12586>
- Stephanie, G. (2020, December 28). *Statistical noise: Simple definition*. Statistics How To. Retrieved November 5, 2021, from <https://www.statisticshowto.com/statistical-noise/>.
- Stone, D. A. (1989). Causal stories and the formation of policy agendas. *Political Science Quarterly*, 104, 281–300. doi:10.2307/2151585
- Stream, C. (1999). Health reform in the states: A model of state small group health insurance market reforms. *Political Research Quarterly*, 52(3), 499–525. <https://doi.org/10.1177/106591299905200303>

Strunk, K. & Zeehandelaar, D. (2011). Differentiated compensation: How California school districts use economic incentives to target teachers. *Journal of Education Finance*, 36(3), 268–293. <https://doi.org/10.1353/jef.2011.0002>

Student Outcomes Assessment: University of Illinois. CITL. (n.d.). Retrieved January 18, 2022, from <https://citl.illinois.edu/citl-101/measurement-evaluation/outcomes-assessments>

Subotnik, R. (1993). *Genius revisited: High IQ children grown up*. Norwood, NJ: Ablex.

Sundquist, C. (2002). Equal opportunity, individual liberty, and meritocracy in education: reinforcing structures of privilege and inequality. *Georgetown Journal on Poverty Law & Policy*, 9(1), 227–248.

Swift, A. & Marshall, G. (1997). Meritocratic equality of opportunity: Economic efficiency, social justice, or both? *Policy Studies*, 18(1), 35–48. <https://doi.org/10.1080/01442879708423720>

Taylor, M. J. (2020). Meritocracy. *TLS. Times Literary Supplement* (1969), 6137, 8–8.

Taylor, B., Cantwell, B., Watts, K., & Wood, O. (2020). Partisanship, white racial resentment, and state support for higher education. *The Journal of Higher Education (Columbus)*, 91(6), 858–887. <https://doi.org/10.1080/00221546.2019.1706016>

Teacher incentive fund. Office of elementary and secondary education. (n.d.). Retrieved November 9, 2021, from <https://oese.ed.gov/teacher-incentive-fund/>.

"Teacher merit pay." Ballotpedia. Accessed April 03, 2019. https://ballotpedia.org/Teacher_merit_pay.

The National Commission on Excellence in Education. (1983, April 1). *A nation at risk: The imperative for educational reform*. Retrieved March 16, 2022, from https://edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf

Teodoro, M. (2009). Bureaucratic job mobility and the diffusion of innovations. *American Journal of Political Science*, 53(1), 175–189. <https://doi.org/10.1111/j.1540-5907.2008.00364.x>

Tekwe, C. D., Carter, R. L., Ma, C., Algina, J., Lucas, M. E., Roth, J., et al. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29(1), 11–35.

The Trustees of Princeton University. (n.d.). *Stay connected for the latest books, ideas, and special offers*. Princeton University. <https://press.princeton.edu/ideas/a-belief-in-meritocracy-is-not-only-false-its-bad-for-you>

Thornton, Z. M., & Friedel, J. N. (2016). Performance based funding: State policy influences on small rural community colleges. *Community College Journal of Research and Practice*, 40(3), 188–203.

Torre, M. de la. (2012, April 23). *Changes in student populations and teacher workforce in low-performing Chicago schools targeted for reform*. National Center for Education Statistics (NCES) Home Page, a part of the U.S. Department of Education. Retrieved January 11, 2022, from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=REL2012123>

Tracking the COVID-19 economy's effects on food, housing, and employment hardships. Center on Budget and Policy Priorities. (n.d.). Retrieved January 27, 2022, from <https://www.cbpp.org/research/poverty-and-inequality/tracking-the-covid-19-economys-effects-on-food-housing-and>

United Teachers of Los Angeles. 2006-2009: UTLA/LAUSD Contract 2009. Available from <http://www.utla.net/node/2451>.

US Census Bureau. "Population and housing unit estimates datasets." *Datasets*, 18 Dec. 2018, www.census.gov/programs-surveys/popest/data/data-sets.html.

U.S. Department of Education. Race to the top fact sheet (2009), <https://www2.ed.gov/programs/racetothetop/factsheet.pdf> [<https://perma.cc/35GG-Y3HM>].

USAFacts. (n.d.). *Spending per student in k-12 public schools*. USAFacts. Retrieved September 14, 2021, from https://usafacts.org/data/topics/people-society/education/k-12-education/spending-per-student-in-k-12-public-schools/?gclid=CjwKCAjw7fuJBhBdEiwA2ILMYSOYTed3a2-UMqFvaNe1W_M9wxmxDTh8NvVRhZhYoZxutdXb9_yrBoCgZkQAvD_BwE.

USAFacts. (n.d.). *Percent of people in poverty*. USAFacts. https://usafacts.org/data/topics/people-society/poverty/poverty-measures/poverty-rate-of-all-persons/?utm_source=google&utm_medium=cpc&utm_campaign=ND-StatsData&gclid=Cj0KCQjw24qHBhCnARIsAPbdtlIVZh1ZGu97Oae15-NpylCEEaG2mkS5pjgwVLCIWZLsvLyoSXaiE_AaAnhIEALw_wcB

Volden, C. (2006). States as policy laboratories: Emulating success in the children's health insurance program. *American Journal of Political Science*, 50(2), 294–312. doi:10.1111/j.1540-5907.2006.00185.x.

- Walker, J. (1969). The diffusion of innovations among the American states. *The American Political Science Review*, 63(3), 88–46. <https://doi.org/10.2307/1954434>
- Walker, R. M., Avellaneda, C.N., & Berry, F.S. (2011). The diffusion of innovation: A longitudinal empirical rest of the Berry and Berry model. *Public Management Review* 13: 95-125.
- Welch, S., & Thompson, K. (1980). The impact of federal incentives on state policy innovation. *American Journal of Political Science*, 24(4), 715–729. <https://doi.org/10.2307/2110955>
- Weyland, K. (2004). *Learning from foreign models in Latin American policy reform*. Woodrow Wilson Center Press.
- What is race to the top and how will it benefit public schools?* Public School Review. (n.d.). Retrieved November 10, 2021, from <https://www.publicschoolreview.com/blog/what-is-race-to-the-top-and-how-will-it-benefit-public-schools>.
- Willner, J. and Grönblom, S. (2009). The impact of budget cuts and incentive wages on academic work. *International Review of Applied Economics*, 23(6), 673–689. <https://doi.org/10.1080/02692170903239853>
- Winter, S., Jr. (1968) Toward a neo-Schumpeterian theory of the Ffrm. Publication No. P-3802. Santa Monica: Rand Corporation.
- Wong, K. K., & Shen, F. X. (2002). Politics of state led education reform: Market competition and electoral dynamics. *Educational Policy*, 16(1), 161–192.
- Wooton, W. (1965). *SMSG; the making of a curriculum*. Yale University Press.
- Young, M. (2017). *The rise of the meritocracy*. (Second edition.). Taylor and Francis.
- Zaltman, G., Duncan, R., & Holbek, J. (1973). *Innovations and organizations*. Wiley.
- Zamani-Gallaher, E. & Yeakey, C. C. (2014). *The Obama administration and educational reform*. (First edition.). Emerald Publishing Limited.
- Zhao, Y. (2009). *Catching up or leading the way: American education in the age of globalization*. Alexandria, VA: ASCD.

Zumeta, W. (1998). Public university accountability to the state in the late twentieth century: Time for a rethinking? *Policy Studies Review*, 15(4), 5–22.

VITA

Originally from Connecticut, Kirsten England grew up in Branford, Connecticut. After high school, she attended Florida Southern College in Lakeland, Florida and received a Bachelor of Science degree in Political Science. Before graduating with her undergraduate degree, she knew she wanted to attend graduate school. She chose to attend the University of Tennessee, Knoxville to pursue a Doctor of Philosophy degree in Political Science with a concentration in Education Policy. Her research interest includes studying the how education policies diffuse and the US court system. After graduation,

she hopes to find a tenured track position at a local university in the Tampa, Florida area. She is incredibly grateful for all the support from her husband, fur-babies, and friends as she begins this next exciting chapter.