

Essays on State Business Tax Incentives and Policy Diffusion

Stephanie Leiser

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Justin Marlowe, Chair

Mark Long

Pamela Clouser McCann

Program Authorized to Offer Degree:

Evans School of Public Affairs

© Copyright 2014

Stephanie Leiser

University of Washington

Abstract

Essays on State Business Tax Incentives and Policy Diffusion

Stephanie Leiser

Chair of the Supervisory Committee:

Justin Marlowe

Associate Professor of Public Affairs

Endowed Professor of Public Finance and Civic Engagement

Daniel J. Evans School of Public Affairs

States' use of tax incentives to lure business investment has drawn a lot of attention from policymakers, the public, and academics. Despite mixed assessments of their efficacy in creating net increases in jobs and investment, business tax incentives continue to be highly popular among state policymakers. This dissertation draws on theories of policy diffusion and tax competition to explore why states adopt tax incentives for business, with a focus on how states influence each other in their policy choices.

The first essay is a mixed methods case study of the spread of incentives for film and video production. It examines qualitative data on interviews with state policymakers, as well as quantitative data on the adoption of film incentives in all 50 states. The analysis suggests that state adoptions of film incentives are primarily driven by two factors: the size of the existing film industry in the state and a "bandwagon" effect based on the total number of adopters. The second essay investigates whether similar adoption patterns hold for four other state tax incentive

policies: Investment Tax Credits, apportionment formula changes, R&D tax credits, and Job Creation Tax Credits. The quantitative event history analyses show that factors that influence adoption decisions are largely inconsistent across the different incentive types, but the evidence is consistent with the idea that the adoption of business tax incentives is a zero-sum game or “race to the bottom.”

One theme that emerges in both of the first two essays is the importance of modeling and interpreting how diffusion processes change over time. The third essay is a methodological discussion of duration dependence—how the hazard of adoption changes over time—in the context of policy diffusion. It argues that the most commonly used methods for modeling diffusion are inadequate for detecting certain system-level diffusion dynamics and should be complemented by a more thorough analysis of duration dependence. It discusses how duration dependence relates to other methodological issues and provides a list of recommendations for researchers on how to properly model and interpret duration dependence in quantitative policy diffusion studies.

Acknowledgements

I am grateful to my committee members for their enthusiasm, patience, and encouragement. I would especially like to thank my committee chair, Justin Marlowe. If he ever had any doubts about me, he didn't let on, and that has made more of a difference than he realizes. I also owe a special thanks to the Evans School PhD program coordinator, Bill Zumeta. He is a passionate advocate for the PhD program, and more importantly, for its students.

For many acts of friendship and support both large and small, thank you to my fellow PhD students: Moroni Benally, Grant Blume, Ryan Bodanyi, Anne Buffardi, Colleen Chrisinger, Tyler Davis, Kate Destler, Maya Dietz, Laura Fricke, Dani Fumia, Lily Hsueh, Shalini Jain, Shane Johnston, Allison Kelly, Alec Kennedy, Helen Kilber, Yuta Masuda, Jacqueline Meijer-Irons, Kathleen Moore, Sarah Oppenheimer, Travis Reynolds, Tim Scharks, Ryan Scott, Tyler Scott, Pradeep Singh, Jason Williams, Sarah Williamson, and Ziad Zaghrouit.

I also wish to acknowledge my former colleagues in and around Michigan state government. I've learned more from them than could ever be codified in a dissertation, and it is their influence that gave me the courage to try for a PhD. While academics have the luxury of keeping the daily work of government at arm's length, someone must go about making the decisions, fixing the problems, and completing the endless thankless tasks that constitute public service.

The most important and heartfelt thanks must go to my family. Thank you to Lynne Leiser for making our family feel at home in Seattle and being a grandmother to our girls. I can't thank her enough for all of the babysitting, shared meals, and words of encouragement.

To my parents, Laura and Steve Schmidt, and my sister, Erin Tabb, this PhD is just the latest in the series of life events for which I know they will always be by my side no matter what. Even as I sit here putting the finishing touches on my dissertation, my mom is taking care of my kids, having flown across the country and taken a month of her own time to help me.

To say “thank you” to my husband, Scott, would be entirely inadequate. To finish my PhD, I’ve had to ask more of him than I ever expected. In addition to everything else he does every day, for me he exemplifies what it means to be a scientist and a scholar, to persevere through all of the alternating setbacks and victories of research because he knows his work serves a greater purpose.

Finally, I am grateful to my daughters. Whether this work in some small way contributes to making a better world for them is the standard by which any of it matters.

Table of Contents

Introduction	1
References	9
Essay One: The Diffusion of State Film Incentives	10
Introduction	11
Policy Diffusion, Tax Competition, and Mechanisms of Interstate Influence	13
Research Design	16
Adoption of Film Incentives in Michigan	18
Adoption of Film Incentives in Washington	21
Adoption of Film Incentives in Mississippi	22
Case Comparison and Hypothesis Generation	25
Event History Analysis	27
Discussion and Conclusions	38
References	42
Tables and Figures	45
Appendix	51
Essay Two: The Diffusion of State Business Tax Incentives	52
Introduction	53
State Tax Competition	54
Policy Diffusion and State Tax Policy	58
State Business Tax Incentives	59
A Model of Policy Adoption and Diffusion	67
Data and Hypotheses	71
Results	77
Discussion and Conclusions	88
References	91
Tables and Figures	95

Essay Three: Duration Dependence in Policy Diffusion Studies	105
Introduction	106
Defining and Operationalizing Policy Diffusion Mechanisms	108
Event History Analysis and Duration Dependence	112
System-level Dynamics in Policy Diffusion	120
Modeling Strategies and Recommendations	125
Conclusions	131
References	133
Figures	136
Conclusions	138

Introduction

Why do states adopt tax incentives for business? This is the question that motivates the three essays in this dissertation. States have a unique ambivalence toward tax incentives for business and economic development. When asked about specific tax incentives—for example, the state of Washington’s willingness to give Boeing nearly \$9 billion in tax breaks to build the new 777X airplane in the state—state policymakers are often overwhelmingly supportive. Landing a major project with a tax incentive is usually just as politically advantageous as securing pork barrel earmarks in the budget process. They are both tangible reminders of policymakers’ influence and ability. But when asked about the use of tax incentives generally as an economic development strategy, state policymakers play the role of a reluctant party to a game they suspect is rigged against them, frequently invoking the metaphor of an “arms race” or “race to the bottom.” As one former state legislator I interviewed laments, “I would certainly hope that there would be a time when every state would unilaterally disarm from the use of tax incentives.” In addition, a study by the Pew Center on the States finds that “... no state regularly and rigorously tests whether those investments [incentives] are working and ensures lawmakers consider this information when deciding whether to use them...” (Pew Center on the States, 2012, p. 1). Tax incentives are one of the biggest components of state economic development policies, but state policymakers seem so disillusioned with them that they do not dedicate adequate resources to evaluating their effectiveness. As one state tax policy expert remarked, it is a “love/hate relationship” between state policymakers and tax incentives.

Tax incentives can include credits, deductions, exemptions, and various other provisions that are intended to reduce the total tax burden on a particular type of taxpayer or activity, such as capital investment or research and development. A recent investigation of state business incentive programs by the New York Times counts 1,874 programs on which states implicitly

spend about \$80.4 billion per year (Story, Fehr, & Watkins, 2012). These are not 1,874 unique programs—many are very similar from state to state. For example, in 2000, six states had adopted tax incentives for film and video production. Ten years later, *all but* six states had them. This pattern of rapid and widespread adoption is suggestive of a diffusion or interdependence process in which states’ decisions to adopt incentives are not merely coincidental. The adoption of an incentive by one state likely influences other states’ decisions about whether to adopt the policy.

The nature of policy interdependence, however, is rarely clear. As mentioned above, the “race to the bottom” metaphor is often used with regard to tax policy issues. The race to the bottom implies a kind of reverse auction where states bid up the generosity of incentives in order to attract jobs and investment. The race to the bottom is a zero-sum game—one state’s gain is another state’s loss. This view of interdependence implies that states will be influenced by the actions of other states they view as their competitors. Alternatively, some view interstate competition as healthy and argue that states are “laboratories of democracy,” experimenting with innovative policy ideas and developing varied responses to common problems. This view of interdependence implies that states look for opportunities to learn from each other.

This dissertation asks three questions: how exactly do these channels of influence work when states are adopting tax incentives? Which states influence each other and why? How can these influences be detected in the data? The essays in this dissertation explore these questions with a mixture of qualitative and quantitative data, analyzed in both in-depth and comparative case studies of five different types of state business tax incentives.

Similar questions have been addressed in the tax competition literature in economics and the policy diffusion literature in political science and public administration. The tax competition

literature examines strategic interactions between jurisdictions in the setting of tax rates (for a review, see Wilson, 1999). Instead of examining tax rates, I focus on more discrete tax policy choices, the adoption of certain tax incentives. As such, it is similar to the policy diffusion literature, which examines a wide variety of discrete policy innovations and attempts to model multiple mechanisms of interjurisdictional influence, including competition, learning, coercion, and other mechanisms (Berry & Berry, 2007). However, even within the policy diffusion literature, there is considerable disagreement about how to define and categorize various mechanisms of diffusion (Elkins & Simmons, 2005), besides the issues involved in operationalizing and measuring them (Maggetti & Gilardi, 2013). This dissertation attempts to find common ground between the approaches of the tax competition and policy diffusion literatures by taking a detailed look at the mechanisms of interstate influence in state business tax incentive policies.

Recognizing that literatures on tax competition and policy diffusion provide conflicting guidance on the causal mechanisms of diffusion, in the first essay, I take an inductive approach and conduct an intensive case study of the adoption of state tax incentives for film and video production. The adoption of state film incentives can be considered an extreme case of diffusion because it resulted in near universal adoption among states in a very short period of time. As Gerring (2007) explains, the greatest strength of an extreme case study is its ability to generate new hypotheses and refine existing theory. Extreme cases are usually considered to be “prototypical” or “paradigmatic,” highlighting potential causal relationships in their starkest context (Gerring, 2007).

The case study begins by analyzing qualitative data from interviews of policymakers in three states that adopted film incentives: Michigan, Washington, and Mississippi. I use these

data to develop a synthetic narrative for each state that describes the political processes leading to adoption and analyze these summaries to try to identify key variables that affect states' decisions. Based on these results, especially details about how states influence one another in their decisions, I develop three hypotheses about how diffusion mechanisms operate for film incentives: a state is more likely to adopt film incentives if 1) its neighboring states have done so, 2) there is a greater total number of adopting states (i.e. a "bandwagon" effect), and 3) it has a large local film industry.

I then test these hypotheses using event history methods and quantitative data on the timing of film incentive adoption in all 50 states. I find support for the second and third hypotheses, but not the first. Specifically, I find a prominent "bandwagon" effect—a statistically significant quadratic relationship between the total number of prior adopters and the hazard of adopting film incentives, suggesting that the hazard peaks when about 34 states have adopted. This finding is important because it suggests that states can be influenced by the total number of adopters, even while controlling for channels of influence that most policy diffusion studies focus on, from neighbors and ideologically similar states. I also find that a 1% increase in the size of a state's film industry is associated with a 40-50% increase in the hazard of adopting film incentives, all else equal. Furthermore, the interview data suggest that the idea for incentives spread from state to state through contacts between members of the film industry, not through state policymakers, as is normally assumed in cases of diffusion.

To see whether the patterns of adoption for film incentives are similar for other types of tax incentives, in the second essay, I examine factors that affect states' adoptions of four common business tax incentives: Investment Tax Credits, apportionment formula changes, R&D tax credits, and Job Creation Tax Credits. I analyze quantitative state data spanning the years

1970-2010 using event history methods. Surprisingly, I find few commonalities among the factors that significantly influence states' adoptions of these different tax incentives. For example, the influence of the size of the manufacturing industry on the hazard of adopting these incentives should be similar to the influence of the size of the film industry on the hazard of adopting film incentives because these incentives tend to target manufacturing. However, I find that while having a larger manufacturing industry significantly increases the hazard of adopting Investment Tax Credits and Job Creation Tax Credits, it does not appear to have an effect on decisions to adopt apportionment changes or R&D credits. Moreover, unlike many policy diffusion studies, but consistent with my findings about film incentives, I do not find strong evidence that states are consistently influenced by their neighbors in deciding whether to adopt tax incentives.

The most important findings that emerge from this analysis are not about particular variables, but rather about the importance of modeling how the diffusion process changes over time. Event history methods include several tools to model time dynamics, such as non-proportional hazards (i.e. time-varying effects) and baseline hazard plots, but these tools are underutilized in the policy diffusion literature. I find, for example, that when the effects of covariates are modeled as time-constant, none of the coefficients in the apportionment change model are significant at conventional levels. However, when the effects are allowed to vary over time, several coefficients attain significance. Also, the baseline hazard plots, which illustrate how the hazard of adoption changes over time controlling for the effects of covariates (also called duration dependence), follow an inverted U-shape for all four tax incentives, meaning that over time, the general popularity of each incentive first rises, peaks, then falls. As I explain in the essay, this pattern is consistent with the idea that states view the use of tax incentives as a

competitive zero-sum game or race to the bottom. At first, states race to gain the “first mover advantage,” and the hazard is rising. As more states adopt over time, however, the comparative advantage of having the tax incentive lessens, reducing the urgency to adopt them, and causing the hazard to fall.

The third essay is more methodological in orientation and expands upon an important issue that arose while I was working through the first two essays. Specifically, when working with the film incentive data, I noticed that if I tried to include a variable measuring the total number of prior adopters (i.e. the bandwagon effect) in a Cox proportional hazards model, the algorithm failed to converge and the model could not be estimated. As I explain in detail in the third essay, the mathematics of the Cox model prevent the total number of prior adopters from being included as a variable in the regression because it is so closely related to time. In effect, it is absorbed into the effect of time, which is illustrated by the Cox baseline hazard plot. I had to find a way to disentangle the bandwagon effect from the baseline duration dependence by switching away from the Cox model to a different type of event history model.

The key insight to arise out of this experience is that a full understanding of the policy diffusion process is incomplete without an analysis of the baseline hazard. The third essay explores the meaning of duration dependence in quantitative policy diffusion studies and discusses the various reasons why one may observe a baseline hazard that changes over time. Most importantly, I explain why certain system-level diffusion dynamics—including bandwagon effects, critical mass effects, competitive races, policy outbreaks (Boushey, 2010; 2012), and system-wide generalized learning processes—may *only* be detected in the baseline hazard. I argue that by focusing only on individual-level characteristics of states or other governments and treating duration dependence as a statistical nuisance to be controlled for, policy diffusion

scholars are overlooking important dynamics in diffusion processes. The essay also discusses the relationship between duration dependence and other modeling issues like non-proportional hazards and unobserved heterogeneity, and includes a list of recommendations policy diffusion researchers can follow to properly model and interpret duration dependence.

Modeling these system-level diffusion effects is especially important in studying the diffusion of tax policies, which are often characterized by competitive dynamics like races to the bottom or bandwagon effects, as I find in the first and second essays. In fact, given the inconsistent findings on the influences of neighboring states or state-level political and economic characteristics, system-level dynamics may be the most important predictors of state adoptions of tax incentives for business.

References

- Berry, F. S., & Berry, W. D. (2007). Innovation and diffusion models in policy research. In P. Sabatier (Ed.), *Theories of the policy process* (pp. 223-260). Boulder, CO: Westview Press.
- Boushey, G. (2010). *Policy diffusion dynamics in America*. New York, NY: Cambridge University Press.
- Boushey, G. (2012). Punctuated equilibrium theory and the diffusion of innovations. *Policy Studies Journal*, 40(1), 127-146.
- Elkins, Z., & Simmons, B. (2005). On waves, clusters, and diffusion: A conceptual framework. *The Annals of the American Academy of Political and Social Science*, 598(1), 33-51.
- Gerring, J. (2007). *Case study research : Principles and practices*. New York: Cambridge University Press.
- Maggetti, M., & Gilardi, F. (2013). *How policies spread: A meta-analysis of diffusion mechanisms*. Unpublished manuscript.
- Pew Center on the States. (2012). *Evidence counts: Evaluating state tax incentives for jobs and growth*. (). Washington, D.C.: Pew Center on the States.
- Story, L., Fehr, T., & Watkins, D. (2012, December 1). United States of subsidies. *New York Times*. Retrieved from <http://www.nytimes.com>.
- Wilson, J. D. (1999). Theories of tax competition. *National Tax Journal*, 52(2)

Essay One: The Diffusion of State Film Incentives

In 2000, only six states had programs to provide monetary incentives for the TV and movie industry to shoot and produce films in their states. By the end of 2010, *all but* six states had enacted some kind of film incentive legislation. Why the sudden proliferation of incentives for the film industry? One possibility is that states were merely responding to the problem of “runaway production,” which refers to the film industry’s increased activity in Canada and other locations outside the US in the late 1980s and 1990s (see McDonald, 2011). During this time, the Canadian federal government and most of the provincial governments began actively recruiting the film industry using generous tax incentives and subsidy programs (Department of Canadian Heritage, 2009). In fact, Oklahoma, one of the earliest states to adopt film incentives in 2001, titled its legislation the “Compete with Canada Film Act” (SB 674 of 2001). Others suggest that while Canadian competition may have provided the initial impetus for the adoption of incentives in the United States, the widespread adoption of film incentives has been driven by interstate competition, not international competition. McDonald (2011) argues that “the weapon [of state film incentives] is not currently employed to protect the nation, but it is being used by states to fight each other.” Critics have argued that film incentives just move production from one state to another, and there are no net economic gains once the effects on all states are considered (Luther, 2010). Moreover, Button (2014) finds that state film incentives have largely failed to create local film industries, as measured by the number of business establishments, in states with incentives.

Whichever account is correct, there is no question that state film incentive policies have spread rapidly and reached a high level of saturation among states. This pattern of widespread adoption in such a short time period is suggestive of a diffusion process that involves interstate influence, which can for example involve states learning from or competing with one another.

State film incentives provide an interesting, if extreme, case study of how policy ideas diffuse within the context of American federalism. In particular, the case study approach can provide insight about how interstate influence works in an applied policy context, exposing weaknesses in and suggesting refinements to existing policy diffusion theory.

This essay uses qualitative elite interview methods combined with quantitative event history analysis to develop a detailed account of why states adopt film incentives. First, I use elite interview data and process tracing techniques to describe the political processes that led to film incentive adoption in three states: Michigan, Washington, and Mississippi. The analysis reveals that in each of these states, the reasons for adopting film incentives were quite different. For example, with regard to interstate influence, Washington and Mississippi noted their key competitors as their geographic neighbors while Michigan viewed itself as competing with all 49 other states. Drawing on these in-depth accounts, I develop testable hypotheses about which factors affect whether a state adopts film incentives, focusing specifically on potential patterns of interstate influence. Next, I test these hypotheses using quantitative data on the timing of adoption for all 50 states. Unlike most studies that use a discrete time setup, I use exact-time data, which allows for more precision in detecting patterns related to the timing of adoptions. I find that state adoptions of film incentives are primarily driven by two factors: the size of the existing film industry in the state and a “bandwagon” effect based on the number of states that had already adopted film incentives. I find limited evidence that states are influenced by their geographic neighbors and no evidence that they are influenced by their ideological similarity to prior adopters. Finally, taking into account the results of both the qualitative and quantitative analyses, I discuss the implications that these empirical case study findings have for policy

diffusion theory. Specifically, I argue that existing theory under-emphasizes the role of duration dependence and bandwagon effects in state policy diffusion.

Policy Diffusion, Tax Competition, and Mechanisms of Interstate Influence

Policy diffusion is a process that Strang (1991, p. 325) concisely defines as “any process where prior adoption of a trait or practice in a population alters the probability of adoption for remaining non-adopters.” Therefore, diffusion excludes process of coincidental policy convergence, in which states act independently but happen to arrive at similar policies. Interstate influence and the spread of policies among states and other governments has been extensively studied in the literature on policy diffusion (see Berry and Berry, 2007 for a review). However, policy diffusion researchers have largely focused on expenditure-side policies like public goods production and regulation, leaving questions of tax policy and tax incentives to the tax competition literature in economics (Shipan and Volden, 2012). The few studies of the spread of certain tax policies include studies of the adoption of state income and sales taxes (Berry and Berry, 1992), enterprise zones (Mossberger, 2000), tax apportionment policies (Omer and Shelley, 2004), research and development tax credits (Miller and Richard, 2010) and sales tax exemptions (Fletcher and Murray, 2006).

Nevertheless, policy diffusion scholars have made significant advances in elucidating and distinguishing the causal mechanisms involved in diffusion processes. As Shipan and Volden (2008) explain, policy diffusion models posit four basic mechanisms of diffusion:

- *Learning*. States learn about when other states adopt a certain policy and perhaps even whether or not that policy was successful.

- *Competition.* States may emulate the policies of other states in order to gain an economic advantage or avoid being disadvantaged. Competition is usually associated with economic spillovers or externalities, where the actions of one state have economic effects on other states.
- *Imitation.* States can adopt policies simply because they are imitating certain other states (e.g. states they view as “leaders”)—no learning about the actual policy is necessary. This mechanism is sometimes also called mimicry.
- *Coercion.* Coercion involves other states or levels of government purposively trying to influence a state’s policy choice.

A rapidly growing body of research has focused on trying to empirically differentiate the mechanisms of learning and competition. For example, Berry and Baybeck (2005) approach this problem by arguing that policy learning and competition imply influences from different sets of states. Specifically, they contend that policy learning is a simpler process in which states are likely to be influenced by the number of nearby states adopting a policy. Competition, however, requires that competitor states be specified according to context-specific economic relationships between states, such as the locations of specific populations or industries. They use data on lottery adoptions and welfare benefits to demonstrate that estimating models that test for only learning or only competition lead to different results than models that include both learning and competition. Similarly, in studying state tax apportionment changes, Omer and Shelley (2004) find evidence that states compete with their neighbors that share similar industries and cross-border Metropolitan Statistical Areas while controlling for learning with a variable that measures the total number of neighboring states that previously changed apportionment formulas. Other

studies that empirically differentiate between learning and competition include Baybeck, Berry, and Siegel's (2011) analysis of lottery adoptions, Shipan and Volden's (2008) analysis of antismoking policies, and Boehmke and Witmer's (2004) analysis of Indian gaming.

Questions of interstate influence in tax policy are frequently discussed in the economics literature on tax competition. However, the tax competition literature differs from the policy diffusion literature in how it conceptualizes the mechanisms of interstate diffusion/influence. Unlike the policy diffusion literature, the tax competition literature implicitly assumes a broader theoretical conception of "competition" and interprets all interstate influence as "competitive." For example, studies in both the tax competition literature and the policy diffusion literature have found evidence that states are influenced by their geographic neighbors (e.g. Rork, 2003 and Berry and Berry, 1992). The tax competition literature interprets this neighborly influence as evidence of competition, while the policy diffusion literature generally interprets it as evidence of learning or mimicry. In the policy diffusion literature, competition is generally viewed as a zero-sum situation in which states compete for some finite benefits (e.g. business investment) (Shipan and Volden, 2008). In the economics tradition, competition is not necessarily zero-sum. For example, in the Tiebout model, the theoretical basis for much of the tax competition literature, governments are analogous to firms in a perfectly competitive market (Tiebout, 1956). Citizens (customers) can choose to live (buy) wherever they want, and governments (firms) compete for citizens by offering desirable public goods for the lowest possible tax price. The desire to retain and attract citizens gives governments an incentive to learn, innovate, experiment, and ultimately adopt policies that help them become more effective and efficient. In the Tiebout model, competition necessarily involves governments learning from one another, and results in a positive-sum outcome.

Research Design

This research design uses a multi-method approach to explaining how states influence one another in the adoption of film incentives. The first part of this research consists of qualitative case studies of three states intended to describe why they adopted film incentives and how film incentive policies in other states may or may not have influenced the policy process. Elite interviews were conducted with policymakers and other stakeholders involved in the adoption of film incentives in their respective states. Following the suggestion of Beamer (2002), the interview data are triangulated with secondary data sources, such as official records, newspaper articles, and other interviews, to maximize reliability.

Sets of interviews were conducted in three different states, Michigan, Washington, and Mississippi. Of the 43 total adopters, Mississippi was the 10th state to adopt film incentives in May of 2004, Washington was the 24th state to adopt incentives in March of 2006, and Michigan was the 32nd state to adopt incentives on the very last day of 2006. The sample therefore covers a broad range in the timing of adoption, the dependent variable in this study. The states in the sample were also chosen to exhibit variation in geographic, economic, and political characteristics as well. Both Washington and Michigan, though far apart, share a border with Canada, while Mississippi is in the south. Mississippi is generally smaller, poorer, and more politically conservative than the other two states. Washington is richer and more liberal than the other two states. Michigan has the largest population and most professionalized legislature, but falls between Mississippi and Washington in terms of wealth and political ideology. I selected initial interviewees in each state after reviewing legislative documents and newspaper articles

that mentioned specific individuals as key players. From there, I used a “snowball sampling” method by asking interviewees who else they thought I should interview. In total, ten interviews in Michigan, six in Washington, and three in Mississippi were conducted with legislators, legislative staff, agency officials, lobbyists, and representatives from the film industry.

The interviews themselves were semi-structured and open-ended elite interviews, which treat interviewees as experts and allow them to provide their own understanding of events (Leech, 2002). Although the interviewees were generally encouraged to lead the conversation, all interviews included discussions of specific interview questions, listed in Figure 1. Following the logic of the policy diffusion literature, I assume that different mechanisms of diffusion will imply influences from different sets of states, which is specifically addressed in questions 7 through 10. However, recognizing that the literatures on policy diffusion and tax competition provide inconsistent guidance on the conceptual nature of diffusion mechanisms, I do not assume *a priori* that a particular pattern of diffusion—say, influence by geographic neighbors—is associated with a particular definition of competition, learning, etc. Therefore, I rely on interviewees to convey their own understandings of why and how certain states influence one another.

The interview data are first used to synthesize coherent accounts of the adoption of film incentives in each state. I then use the case data to identify specific variables that seem to be important in one or more states’ decisions to adopt film incentives using process tracing techniques. Process tracing emphasizes the linking of case data to elements of theory (George and Bennett, 2005) and can therefore serve as a useful tool to inductively generate hypotheses about how film incentive policies spread from state to state. I focus especially on diffusion-related variables that describe how the actions of other states may have influenced the decision to

adopt film incentives in the case study states. The purpose of the qualitative analysis, therefore, is to generate possible causal explanations rather than to adjudicate between them.

The analysis of the case data for Washington, Michigan, and Mississippi yields three hypotheses related to policy diffusion, and the second part of this research tests their generalizability in a quantitative statistical context using event history analysis techniques and data on all 50 states. As Berry and Berry (2007) explain, event history analysis is a technique that is well-suited for testing hypotheses about policy diffusion patterns. Evidence from a Cox proportional hazards model as well as parametric models is presented to explore which factors affect the timing of state film incentive adoption.

Adoption of Film Incentives in Michigan

The initial catalysts for the adoption of film incentives in Michigan were two individuals with film industry connections who could see that film incentives were becoming a major factor in the industry. In the early 2000s, Michigan's Film Commissioner, Janet Lockwood, joined several other state and regional film commissioners to create a group called Film US. Film US's goal was to address the problem of "runaway production" created by Canada's adoption of federal and provincial-level film incentives in the 1990s. Film US lobbied Congress for federal-level film incentives in the U.S., but after little success, the group disbanded, and the film industry turned its lobbying efforts to the state level. In Michigan, Lockwood floated the idea of film incentives internally, but then-Governor John Engler, a Republican, was always hostile to the idea.

In 2003, the political landscape changed with the election of Governor Jennifer Granholm, a Democrat. Also beginning his first term in office that year was Representative Bill

Huizenga, a Republican from West Michigan, who also sat on the board of Compass Film Academy (now called the Compass College of Cinematic Arts) in Grand Rapids. Shortly after he began his term, Huizenga introduced himself to Lockwood and said he was interested in introducing film incentive legislation. They began working on the legislation, which they reportedly modeled based on Louisiana's incentive program because members of the Film Advisory Council indicated that Louisiana was a "hotbed" of incentive-related productions. Huizenga introduced the bills in June 2004. The bills passed the House but died in the Senate Finance Committee at the end of the 2003-2004 Session.

In the next session, Huizenga, who had developed a reputation as an up-and-coming leader, was named Chair of the Commerce Committee. He and his staff continued to work with Lockwood and the Film Council on more "comprehensive" film incentive legislation. The package now included more bills, and therefore more bill sponsors with a stake in the matter, and it also covered commercials, an element designed to win the support of the Big 3 auto companies. The bills were also supported by film industry labor unions, including the Teamsters, pivotal players in Michigan politics. As Chair of the Commerce Committee, Huizenga was able to shepherd the bills through the House and had more bargaining leverage in getting the Senate to take up the bills. Nevertheless, the bills were again sent to the Senate Finance committee, led by Senator Nancy Cassis, a Republican who had always expressed opposition to film incentives. It again came down to the last days of the Session in 2006, and Huizenga was able to recruit a key ally in Dick DeVos, son of billionaire Amway-founder Richard DeVos, Sr. and prominent figure in Michigan Republican politics, who reportedly had a son involved in the film industry. With DeVos's support, Huizenga was able to convince the Senate

Majority Leader to approve a single bill from the multi-bill package on the very last day of Session.

Though the bill became law, the package had been significantly watered down, and Michigan “did not become a player at all” in the film industry. According to one interviewee, “If you’re not on the top of these incentives—if you’re not the best state, then you might as well not be in the business.” The legislation did, however, attract the attention of Governor Granholm and several key legislators, including the incoming Chairman of the Commerce Committee, Democrat Andy Meisner and Chairman of the Senate Commerce and Tourism Committee, Republican Jason Allen. Another workgroup formed in 2007 to draft new and improved film incentives, this time including key members of the Governor’s staff, and explicit directions from the Governor to create the best film incentive in the country. According to several interviewees, this workgroup systematically examined film incentive statutes in every other state (and a few Canadian provinces), and followed a strategy to make each component of Michigan’s film incentive program as good as or better than any other state. As one interviewee put it, “The essence of the package was to look at some of the key elements of the other front-runner states and to one up them—to one up them in each category and by being the most comprehensive of any of the states...” Also during this time, Michigan’s economic situation continued to worsen, leading the nation in unemployment. After years of reliance on the failing auto industry, Michigan’s economy had been anticipating the recession long before it officially began in December 2007, and several interviewees indicate that politics in Michigan had a feeling of desperation and urgency. As one interviewee explained, the film incentives were seen as a “Hail Mary.”

Early in 2008, the workgroup finalized the bill package. The Governor mentioned the film incentives in her 2008 State of the State address, and the bills were introduced in February. Representative Meisner and Senator Allen had agreed to hold joint committee hearings on the bills. Numerous members of the Michigan film industry, including some famous names, came to support the bills, and no one registered opposition to the bills in committee testimony. The new and expanded film incentive package moved quickly through the legislative process, registering a single “no” vote from Senator Cassis, and was signed by Governor Granholm on April 7, 2008.

Adoption of Film Incentives in Washington

The process of adopting film incentives in Washington included much less political fanfare than in Michigan and was driven more by a sense of maintaining Washington’s place in the industry rather than launching a new industry. Once the Canadian federal government and the province of British Columbia adopted film incentives in 1997, the state of Washington reportedly began to lose film industry business to their northern neighbors. According to a legislative report, expenditures on motion picture and video productions in Washington fell by 74% between 2001 and 2006 (Senate Bill Report, SB 6558 of 2006). One interviewee from the Washington film industry said that there had been talks and meetings within the industry about what to do about this loss of business since the 1990s.

In 2004, a group of film industry leaders officially came together to form the Washington Entertainment Industry Players Association (WEIPA), whose goal was to create a film incentive program in Washington state. WEIPA hired an Olympia lobbyist who specializes in tax issues to help them draft a proposal. Meanwhile, Oregon adopted its film incentive program in July of 2005, providing another regional threat to Washington’s film business. As one film industry

representative testified before a Senate committee, “In order to attract productions, we already have to discount services by 10 to 20 percent to compete with Oregon and Vancouver.”

WEIPA continued to develop its policy proposal and hired another veteran lobbyist to help with political strategy. In the Spokane area, which was represented by the Democratic Senate Majority Leader Lisa Brown, a nascent but quickly growing film community was developing, anchored by a full-service film and video production company called North by Northwest. WEIPA and its lobbying team were able to convince Senator Brown to be the lead sponsor of their legislative proposal. The strategy gave instant credibility to the proposal as it is relatively unusual for legislative leaders to sponsor substantive policy bills. WEIPA also successfully recruited the support of film industry labor unions, a key political ally, by including provisions to protect jobs and benefits for in-state workers in the legislation.

Senator Brown introduced the bill in January of 2006, and legislative committee hearings began in February. Similar to Michigan, there was broad support and very little opposition to the proposal. Unlike Michigan, however, arguments for the film incentive program centered around helping Washington regain its fleeing film industry. In fact, the legislation itself states that “in-state producers are taking their projects to more competitive economic climates, such as Oregon and Vancouver, British Columbia, where compelling tax incentive packages and subsidies are already in effect” (SB 6558 of 2006, as enrolled). The bill moved quickly through the legislative process, and Governor Christine Gregoire, who had been quietly but not actively supportive, signed the bill on March 27.

Adoption of Film Incentives in Mississippi

Similar to Michigan, the story of film incentives in Mississippi starts with the state's governor-appointed Film Commissioner. Ward Emling began his career working on film crews, became Mississippi's Film Commissioner in the early 1980s, and was the president of the Association of Film Commissioners International (AFCI) by 1998. Emling's involvement with AFCI not only raised the stature of Mississippi as a filming location, but also gave Emling a unique perspective on what was going on in the film industry on an international level. As mentioned, the most important development in the 1990s was the creation of film incentives in Canada. Emling notes that for the first time, the Canadians "proved that it [film incentives] could be developed like an industry."

From witnessing the success of incentive programs in Canada and elsewhere, Emling realized that the use of incentives represented a shift to an "economic development mentality," in which film production must be lured with financial incentives, just like any other industry. As Emling explains in a 2004 newspaper article, "Everything I do is exactly what other developers do. Economic developers are looking at quality of life, transportation, access to airports and the workforce – the same things we look at as film commissioners." (Lofton, 2004).

As a state, Mississippi was well-positioned to develop its film industry. Mississippi's Film Commission, created in 1973, was one of the first government film commissions not only in the US, but in the world. Mississippi also raised its profile as a film location in the 1990's with the success of "A Time to Kill," "The Client," and "Ghosts of Mississippi." Two events in the early 2000's provided opportunities to elevate the idea of film incentives in Mississippi to the public and legislative agenda. The first was a high-profile tax incentive deal between the state and Nissan Motor Co. to build a \$950 million auto assembly plant in Canton, MS, for which Nissan received \$295 million in state incentives (Lyne, 2000). Emling re-framed his argument

for film incentives to say they could be “just like Nissan.” Supporters of film incentives could then advocate for their issue using the terminology of economic development and a concrete analogy with a wildly popular economic development deal. The other event to create an opening for public consideration of film incentives in Mississippi was the July 2002 expansion of film incentives in Louisiana, which proved to be extraordinarily successful in making Louisiana a top film production destination. According to interviewees, Louisiana’s film incentives inspired both Mississippi’s sense of rivalry with its “sister state,” as well as an opportunity for learning about how to craft a successful film incentive policy.

In the meantime, the idea of film incentives as economic development had caught the attention of the Mississippi Tourism Association, and Emling teamed up with their lobbyist, Donna Echols Mabus, to develop a political strategy. The two began discussions with policymakers, including legislators and agency officials, in 2003. They were able to recruit the enthusiastic support of Rep. Diane Peranich, who became Chair of the newly-created House Tourism Committee in the 2004 session and introduced the film incentive legislation (House Bill 1780 of 2004). In addition, film incentive advocates found support from other legislators who had family members who were involved in the film industry. In making their case to legislators, film incentive supporters reinforced the economic development framing of the issue as well as the regional learning/competition angle. “It made sense to tell them what our neighbors were doing,” one advocate explained, “You know your neighbors. Your neighbors are like you.” Peranich fueled the competitive angle, adding that “Our sister states have been sites for a lot of film activity, and now we will be able to steal the march.” (Lofton, 2004). In drafting the legislation itself, Peranich, Mabus, and Emling drew heavily from Louisiana’s statutes, as well as using other states like North Carolina as models. Peranich introduced the legislation in March of

2004, and by May 12, it had been signed by Governor Haley Barbour, who, like Governor Gregoire in Washington, had quietly supported the legislation.

Case Comparison and Hypothesis Generation

Understanding the diffusion processes involves identifying the channels of influence between adopters and non-adopters. In terms of Strang's (1991) definition, which sets of states alter the probability of adoption for non-adopters? As Berry and Berry (2007) explain, influence from geographic neighbors is one of the most common patterns detected in policy diffusion studies. This also appears to be of possible relevance for film incentives. In Washington, for example, interviewees were very clear that Washington needed to adopt film incentives to counteract the loss of business activity to Oregon and Vancouver, B.C. Similarly, the qualitative data imply that the adoption of film incentives by Louisiana increased the probability that Mississippi would adopt incentives. In these cases, the nature of the influence of prior over potential adopters is mediated by regional proximity. That is, Washington was most greatly influenced by the actions of Oregon and British Columbia (and Mississippi by Louisiana) because they are geographic neighbors and presumably share characteristics that are important to film production, such as climate and scenery. This idea of regional influence leads to the first hypothesis: *States are more likely to adopt film incentives if their geographic neighbors have previously adopted film incentives.*

Unlike Washington and Mississippi, the process of regional diffusion did not appear to be as strong of a factor for Michigan. None of the Michigan interviewees mentioned that they were influenced by the actions of regional states, and while some Michigan interviewees mentioned Canadian competition, they mentioned Vancouver as often as Toronto. What Michigan

interviewees did report is that they felt they were engaged in competition with *all* other states, particularly those that offered the most generous and successful incentives. In other words, they were as likely to see Florida or Arizona as a competitor as Indiana or Ohio. Therefore, a second diffusion hypothesis encompasses the idea of nationwide competition: *States are more likely to adopt film incentives the more other states have previously adopted film incentives.*

These two hypotheses are fairly typical of quantitative policy diffusion studies, but the qualitative case data suggest that there may be another diffusion mechanism at work. Policy diffusion is fundamentally about the flow of information and ideas between political jurisdictions, and it is usually implicitly assumed that this information flows at the level of policymakers, government actors, and others directly involved in government decision-making. For example, Shipan and Volden (2012) explain that low-cost communication, low-cost travel, professional organizations, and informal personal networks all function to facilitate the flow of information among policymakers in different states, often resulting in policy diffusion. The idea that policymakers would have more contacts with policymakers in neighboring states and face lower costs of travel or communication with neighboring states then serves as the basis for a hypothesis of regional influence.

However, in the present case of film incentives, the flow of information in the film industry itself also appears to play an important role. In Washington, the initial impetus for film incentive legislation came from the film industry when it realized that it needed a way to compete with other jurisdictions that offered incentives. One Washington interviewee reported that since the 1990s, incentives have become an integral part of the film business. She explained that for film producers considering filming in a particular state, “The entrance into the conversation is ‘what is the incentive available in this state and how does that fit into my

business model?’’ Similarly, in both Michigan and Mississippi, the state film commissioners had long tenures and were well-connected to industry trends, and were therefore well-informed and well-positioned to advocate for film incentives in their respective states.

As in any competitive industry, each business must keep track of what its competitors are doing to gain a competitive advantage. In the film industry in particular, successfully competing has come to mean recognizing the importance of incentives, which are often as large as the equivalent of a 30-40% subsidy on production expenditures. The larger and more sophisticated a state’s film industry is, the more contacts it will have with film producers in other states, and the more likely (and quickly) it will become aware of the growing importance of film incentives. In addition, a larger industry translates to a larger political constituency and more political power to convince states to adopt incentives. Interviewees in both Michigan and Mississippi emphasized that legislators with familial or personal ties to someone in the film industry—which would be more common with a larger film industry in the state—became key supporters. Similarly, Best and Teske (2002) found that the size of the retail industries in different states had a significant impact on how aggressively those states decided to tax Internet transactions. The size of the film industry, therefore, is a measure of both its competitive capacity and its political power. In this way, the actual mechanism of policy diffusion is the flow of information within the film industry (i.e. the object / target of the policy), not the flow of information among policymakers or government officials (i.e. the subject / enactors of the policy). A third diffusion hypothesis which represents the idea of diffusion at the policy-object level is: *States are more likely to adopt film incentives the larger their respective film industries are.*

Event History Analysis

Data

The quantitative data used in this study span the years 1996 to 2009 and are continuous time data. This means that each row of data (i.e. each observation) represents an interval of time over which values of the variables do not change, and when the value of one (or more) does change, there is a new row of data. Therefore, not all observations are of equal duration, and there are frequently multiple observations of a single state in a single year. Once a state has adopted a film-incentive, it is no longer observed. Table 1 provides a summary of the variables, definitions, and data sources.

In continuous time event history data, the dependent variable is actually three variables: one measuring the start of the duration, one measuring the end, and a “failure” indicator. This indicator takes a value of one if the state adopted a film incentive policy at the end of the specified duration and a value of zero otherwise. I measure duration in days and record the exact day that film incentives became law in each state.¹ The first state to adopt a film incentive policy was Louisiana in 1992, but the idea did not start spreading to other states until 1997, which I use as the “start of risk,” the time at which I consider other states to be “at risk” of adopting film incentives.² By December 31, 2009, the last day in the study, all but six states (Delaware, Nebraska, Nevada, New Hampshire, North Dakota, and Vermont) had adopted film incentives. Figure 2 is a Kaplan-Meier cumulative failure plot, which shows the cumulative proportion of states adopting film incentives over the course of the study. The S-shape created by the plot is typical of diffusion processes.

¹ I use the adoption date rather than the effective date because it better reflects timing in the political decision-making process. Effective dates often vary dramatically from adoption dates, with some states even making the statutes retroactive. Effective dates, therefore, do not accurately represent the political conditions under which adoption decisions are made.

² It is common for the first adopter to be excluded from policy diffusion studies because it obviously would not have been influenced by adoptions in other states. Therefore, a phenomenon of diffusion cannot be studied until there is at least one adopter.

One limitation of this methodology is that it only examines the timing of first adoptions. It does not account for differences in incentive programs across states or subsequent changes to film incentive programs. As I noted in the Michigan case, Michigan first adopted film incentives in December 31, 2006 but its incentive was not very generous compared to other states, and it was not until 2008 that Michigan revamped its program to offer larger incentives. The state of Washington has always had one of the smallest incentive programs in terms of cost. These differences in film incentive programs could account for some of the variation in timing of adoption. For example, a more expensive program would likely be more difficult to pass. Unfortunately, data that would allow for the direct, valid comparison of state film incentive programs does not exist. Nevertheless, the initial political choice to join the growing group of states with film incentives provides sufficient evidence of a diffusion process at work, as the results below will show.

To operationalize the hypothesis of regional influence, I use two variables to measure how states are influenced by their contiguous neighbors. First, I include a variable that measures the percentage of bordering states that have already adopted film incentives at each point in time, *PCTBORDERADOPT*.³ I also include a dummy variable for whether a state shares a physical border with Canada, *CAN_BORDER*, which adopted film incentives at both the federal and provincial levels in the 1990s, before the start of risk in this study. I expect that increases in these two variables will increase competitive pressures, increasing the hazard of adoption. To operationalize the hypothesis of nationwide influence, I include a variable that simply measures the total number of states previously adopting a film incentive policy, *NUM_PREV*. I expect that

³ I use the percentage of border states adopting instead of the number of border states adopting because not all states have the same number of contiguous neighbors. Therefore, I expect an adoption by a border state to have a larger impact on a state with, for example, just two neighbors as opposed to a state with six neighbors.

as the total number of states adopting film incentives increases, it will create a “bandwagon” effect and increase the hazard of adoption for remaining states.

The third major hypothesis is that the size of the existing film industry facilitates the adoption of film incentives. BEA state-level data on real GDP associated with the “motion picture and sound recording industries” are used to create the variable *LN_FILM_IND* (taking the natural log)⁴. Because GDP data for a particular year is not actually available until the end of the year, the data are lagged so that during 1999, for example, states would be making decisions based on 1998 GDP figures. I expect that larger existing film industries will increase the hazard of adopting film incentives.

To control for another possible channel of diffusion, Grossback, Nicholson-Crotty, and Peterson’s (2004) measure of ideological distance is also included in the model. Ideological distance is intended to capture the level of ideological (liberal/conservative) similarity between a potential adopter and previous adopters, the presumption being that states are more likely to adopt a policy if it has already been adopted by states with similar ideological leanings. For example, conservative states could interpret the adoption of film incentives by other conservative states as a signal that the policy conforms to conservative political principles. Underlying ideology scores are from Berry, et al. (2010), and are described in more detail below.

Ideological distance, *IDEO_DIST*, is equal to the absolute value of the difference between the ideology of the potential adopter and a weighted average of the ideologies of all previous

⁴ Since the start of risk begins in 1997, this means 1996 data is needed, but unfortunately, the BEA changed the way it classified industries in 1997, and there is no way to bridge pre- and post-1997 data. Therefore, I generate 1996 figures through linear extrapolation. I extrapolate based on 1997-2002 data, which is before most states adopted incentives. Therefore, the extrapolated numbers should reflect existing industry size trends prior to any effects of adoption (which presumably enlarge the industry).

adopters. The most recent adopter is given a 50% weight, and the other 50% encompasses all other previous adopters.⁵

Finally, a set of variables is included to control for state-specific economic and political characteristics. Recall that in Michigan, several interviewees attributed policymakers' willingness to adopt incentives to a sense of desperation in the face of unprecedented economic troubles. To represent the overall wealth and level of economic hardship in each state, both the unemployment rate (*UNEMP*) and per capita GDP (*LN_PERCAPGSP*) are included. Like the film industry data, I use annual numbers and lag them by a year. Political variables include current (i.e. non-lagged) measures of state government ideology (*IDEOLOGY*)⁶, legislative professionalism (*PROFESS100*)⁷, a divided government dummy (*DIVIDED*), and the corporate income tax rate (*RATE*)⁸.

Cox Model Results

For the first analysis of the data, I estimate a Cox proportional hazards model, which allows for a nonparametric estimation of the underlying duration dependence or baseline hazard. The mathematical properties of the Cox model's partial likelihood prevents the variable measuring the total number of previous adopters (*NUM_PREV*) from being included because its value does not vary across states at a given time t . The effect of the total number of previous

⁵ Results of the models reported below are robust to alternative specifications of ideological distance, including giving all prior adopters equal weight and squaring the difference instead of using the absolute value.

⁶ I use Berry, et.al.'s (2010) *NOMINATE* measure of state government ideology, which ranges from zero (most conservative) to 100 (most liberal). The *NOMINATE* measure is an aggregate measure that accounts for partisan affiliation and power in the governor and state legislature.

⁷ I use Squire's (2007) aggregate index of state legislative professionalism, which is intended to capture a state's "capacity... to generate and digest information in the policymaking process." States with legislatures that are more "professional" have more session days, larger staffs, and higher pay. Professionalism scores range from 0.0 (least professional) to 1.0 (most professional), which I multiply by 100 to get *PROFESS100*. Because the index is only calculated for intermittent years, I linearly extrapolate/interpolate the data so there is an annual measure.

⁸ I use the maximum rate for states with graduated rate structures. States without a corporate income tax are coded as zeroes.

adopters is instead absorbed into the baseline hazard. While the inability to directly test the effect of *NUM_PREV* is a drawback of the Cox model, it is nevertheless important to do this preliminary analysis before specifically parameterizing the baseline hazard because model estimates can be very sensitive to how duration dependence is parameterized (Box-Steffensmeier and Jones, 2004). In the next section, the Cox model results are used to guide in the selection of a parametric model for duration dependence.

The Cox model estimates the hazard of adoption for state *i* at time *t* based on the baseline hazard $h_0(t)$ and the values of that individual's covariates at time *t*:

$$h_i(t) = h_0(t) \exp[\beta_1 PCTBORDERADOPT_i(t) + \beta_2 IDEO_DIST_i(t) + \beta_3 LN_FILM_IND_i(t-1) + \beta_4 CAN_BORDER_i + \beta_5 UNEMP_i(t-1) + \beta_6 LN_PERCAPITAGSP_i(t-1) + \beta_7 RATE_i(t) + \beta_8 PROFESS100_i(t) + \beta_9 IDEOLOGY_i(t) + \beta_{10} DIVIDED_i(t)]$$

Results from the Cox model estimation are reported in Table 2. Hazard ratio coefficients greater (less) than one means that a unit increase in the covariate increases (decreases) the hazard of adopting film incentives. The Cox model being a “proportional hazards” model, the effects are modeled to be constant over time, and the proportional hazards property is statistically confirmed using Grambsch and Therneau’s (1994) test. In addition, examination of DfBeta influence statistics indicate that California has a disproportionate impact on several of the model’s parameter estimates. This makes sense theoretically because California is the historical home of the film industry and did not adopt film incentives until 2009. It is reasonable to believe that California’s decision process was significantly different from other states because it was in a much more defensive position competitively. As a result, I drop California from the model.

In examining the results related to regional influence, the hazard ratios for *PCTBORDERADOPT* and *CAN_BORDER* are both significant at conventional levels ($p = 0.032$ and $p = 0.097$, respectively) and surprisingly less than one. The hazard ratio for *CAN_BORDER* implies that sharing a border with Canada reduces the hazard of adoption by almost 60% at any given time, all else equal. Similarly, increases in *PCTBORDERADOPT* also reduce the hazard of adoption by a rather large amount. For example, for a state with five neighbors, adoption by one of its neighbors results in a 20 percentage point increase in *PCTBORDERADOPT*, which reduces the hazard of adoption by over 40% [$20 \times (1.00 - 0.979) = 0.42$], all else equal. A possible explanation for this result is that states believe there is a “first mover advantage,” in which states that act quickly and adopt film incentives before their neighbors (whether other states or Canadian provinces) will stand to gain the most. Being one of the later states means that much of the potential benefits have already been dissipated, therefore reducing the urgency to adopt film incentives.

The model results regarding the size of state film industries provide support for the hypothesis that a larger film industry will create an impetus for adoption of film incentives. The hazard ratio for *LN_FILM_IND* is 1.504 ($p = 0.007$), meaning that a 1% increase in the size of the existing film industry increases the hazard of adoption by 50%, all else equal, which is quite a dramatic increase. The only other covariate that has a significant effect on the hazard is *IDEOLOGY* ($p = 0.058$), but the effect is relatively small. A one point increase in the *IDEOLOGY* score—becoming slightly more liberal on a scale of 0 = most conservative and 100 = most liberal—increases the hazard of adoption by 1.5%. Although this effect is statistically significant, it is not very large, which makes the non-significant result for ideological distance

(*IDEO_DIST*) unsurprising. As the qualitative interview data corroborates, the politics of film incentive adoption does not appear to have a strongly partisan aspect.

Other than *IDEOLOGY*, none of the economic or political control variables appear to have a significant effect on the hazard of adopting film incentives. In other words, a state with high unemployment, low corporate taxes, and a part-time legislature, all else equal, is just as likely to adopt film incentives as a state with low unemployment, high corporate taxes, and a highly-paid full-time legislature. It is actually quite surprising that the economic variables do not appear to influence whether states adopt film incentives because arguments about film incentives are often framed in terms of economic development. For example, one interviewee in Michigan explained that “the economy was so in the tank, anything was good if it was going to create a job.”

Figure 3 provides a kernel-smoothed non-parametric estimate of the baseline hazard. The baseline hazard can be interpreted as the underlying hazard or duration dependence, purged of the effects of the covariates.⁹ A quick examination of the baseline hazard plot shows a strong pattern of positive duration dependence. Ideally, a Cox model baseline hazard will not show any strong trends, which calls for an explanation of the apparent tendency for the hazard to rise over time in this study.

Three possible explanations for the observed positive duration dependence include: a bandwagon process in which states are influenced by the total number of previous adopters unobserved heterogeneity/omitted variables, or exogenous influences. The explanation that finds most support in the interview data is that there was simply a bandwagon effect, where states felt

⁹ Usually this means estimating the model by setting the values of covariates to zero, but to avoid nonsensical covariate values (e.g. a 0% unemployment rate) I obtained the baseline hazard by mean-centering the variables (except the dummy variables *CAN_BORDER* and *DIVIDED*) and setting them equal to zero.

increasing pressure to adopt film incentives as more and more other states did so (i.e. *NUM_PREV*). This would be an example of a “homogeneous mixing process,” which used to be a common assumption in diffusion studies but has more recently faced criticism (Strang, 1991). Homogenous mixing models are common in epidemiology and assume that “non-infected” states (i.e. states that have not adopted incentives) are equally likely to be “infected” by any of the prior adopters. As a result, the proportion of infected individuals increases over time in the familiar S-shaped curve. Strang (1991) argued that scholars should focus on modeling “heterogeneous mixing” processes in which the probability that prior adopters “infect” non-adopters is structured by political and economic relationships. The *PCTBORDERADOPT* variable is one example that assumes states are more likely to be “infected” by their neighboring states. However, with an industry as mobile as the film industry, competition can be felt equally from any other states. One interviewee related a story of how Washington recently lost a major motion picture production to Alabama, a state with which it shares little in common politically, economically, or socially.

Another possibility is that there are one or more omitted variables that create unobserved heterogeneity in the model. Of course, omitted variables can always be a problem in empirical studies, but as Zorn (2000) explains, in event history models, unobserved heterogeneity appears as negative duration dependence. In cases of observed positive duration dependence, such as the current case, this implies that the observed duration dependence is actually too flat. Given that the task is to explain the existence of positive duration dependence in the first place, the possibility of unobserved heterogeneity provides no explanatory insight.

A third possible explanation for the observed positive duration dependence is that states were responding similarly to some exogenous influences. In 2005, perhaps partly in response to

the first wave of state film incentives, several Canadian provinces, including Ontario and British Columbia, increased the generosity of their film incentives (PriceWaterhouse Coopers, 2006). The provinces may also have been responding to increases in the value of the Canadian dollar compared to the US dollar. Throughout the 1990s, the exchange was falling, making Canada an attractive place for US investment, but beginning in 2002, the exchange rate began to climb again. Although none of the interviewees specifically mention this narrative, it is possible that the 2005 increases in Canadian film incentives may have triggered a second wave of state film incentives. However, one interviewee with a long history in the industry argues that while Canadian competition may have been a factor for the first wave of film incentives, it was no longer relevant compared to competition from other states: “Around 2003, I think I stopped caring about Canada, when I saw Louisiana and New Mexico...” two states widely regarded as the vanguard of state film incentives.

Parametric Model Results

In order to explore the possibility that states are influenced by the total number of previous adopters, parametric event history models must be used. Parametric event history models specify the form of the baseline hazard or duration dependence using various parametric distributions. The exponential model, for example, has a constant baseline hazard, the Weibull and Gompertz models each allow monotonic baseline hazards, and the log-normal and log-logistic models each allow non-monotonic hazards. I use these five distributions, as well as the same set of covariates from the Cox model above, and estimate ten total parametric models: in the first five models, I add *NUM_PREV*, the total number of states previously adopting film

incentives; in the next five, I also add a quadratic term for *NUM_PREV* to capture any non-linearities in the effect of *NUM_PREV* over time.

Details about the model selection process are given in Appendix A. Of the ten models considered, an exponential model with the quadratic specification of *NUM_PREV* fit the data best. Specifically, the model equation is:

$$h_i(t) = \exp [\beta_0 + \beta_1 PCTBORDERADOPT_i(t) + \beta_2 IDEO_DIST_i(t) + \beta_3 LN_FILM_IND_i(t-1) + \beta_4 CAN_BORDER_i + \beta_5 UNEMP_i(t-1) + \beta_6 LN_PERCAPITAGSP_i(t-1) + \beta_7 RATE_i(t) + \beta_8 PROFESS100_i(t) + \beta_9 IDEOLOGY_i(t) + \beta_{10} DIVIDED_i(t) + \beta_{11} NUM_PREV_i(t) + \beta_{12} NUM_PREV2_i(t)]$$

The constant baseline hazard in the exponential model is expressed by the term $\exp(\beta_0)$. The results for this exponential model are given in Table 3.

The model provides strong evidence that the total number of previous states that have adopted film incentives affects the hazard of adoption for states. Both terms for *NUM_PREV* are significant at the 1% level ($p = 0.000$ and $p=0.004$, respectively). The coefficients for *NUM_PREV* and its quadratic term suggest that the effect of *NUM_PREV* is positive but diminishing and eventually negative (maximum hazard is reached when *NUM_PREV* is just over 34 states)¹⁰, which is expected given the general S-shape of diffusion curves. It is also notable that the regional competition variables, *PCTBORDERADOPT* and *CAN_BORDER*, are no longer significant at conventional levels.

Results for the remaining covariates are roughly comparable to the Cox model results. *LN_FILM_IND* is still highly significant ($p=0.007$) and has a hazard ratio of 1.393, which is smaller than its hazard ratio from the Cox model. All else equal, a 1% increase in the size of the

¹⁰ The vertex of the parabola can be calculated using non-exponentiated coefficients. For *NUM_PREV*, the coefficient is 0.3049511 and for *NUM_PREV2*, the coefficient is -0.0044387. The parabola reaches its maximum when $NUM_PREV = \text{abs}[-0.3049511 / 2(-0.0044387)] = 34.3514$.

existing film industry increases the hazard of adoption by about 40%, instead of the Cox model's 50%, which is still quite a large effect. *IDEOLOGY* is still significant ($p=0.043$) and has almost exactly the same hazard ratio as it did in the Cox model, providing additional evidence that more liberal states are slightly more likely to adopt film incentives. Again, the remaining economic and political control variables do not appear to significantly affect the timing of state film incentive adoptions.

Discussion and Conclusions

What factors make states more likely to adopt film incentives? The two most important factors seem to be how large a given state's existing film industry is and how many states have already adopted film incentives. This study tests three hypotheses about the possible state-to-state diffusion processes underlying the spread of state film incentives and finds support for two of them. Both the Cox model and the parametric exponential model provide evidence that small increases in the size of a state's existing film industry translate to large increases in the hazard of adoption. This means that state film industries were highly-successful advocates for a policy that would directly benefit them, and they drew upon their size to exert political influence on state politics. The qualitative case study evidence further suggests that it was the film industry itself that served as the conduit for the diffusion of the idea of state film incentives. In Washington, it was very clearly the film industry that took the initiative to organize itself, develop a policy proposal, and form a political strategy to get the policy adopted. In Mississippi, the key leader was the well-respected and well-connected Film Commissioner, and in Michigan, the impetus for film incentives came from a legislator with strong ties to the local film industry. Those within state film industries were simply responding to increased competitive pressure resulting from the

adoption of film incentives in other states and Canadian provinces. In some cases, a state incentive essentially translates to a 30-40% subsidy for all production expenditures. It makes sense, therefore, that states with larger film industries adopted film incentives more quickly, not only because a larger industry creates a larger constituency with more political power, but because the larger, more sophisticated industries probably caught on to the trend of film incentives faster.

The evidence also suggests that there was a bandwagon style diffusion process happening at the same time, as shown by the statistically significant quadratic relationship between the hazard of adoption and the total number of adopters. States are influenced simply by the total number of previous adopters. Perhaps they see adoptions by other states as signals that film incentives are good policies, perhaps they wish to replicate the perceived success of early adopters, or perhaps they feel increased competitive pressure to keep their existing film business from leaving for other states. Whatever the reason, this diffusion pattern appears to be happening at the level of state policymakers, as opposed to within the film industry.

Importantly, I find little evidence that this “homogenous” mixing process—where states are influenced by the total number of previous adopters—would be better modeled, as Strang (1991) suggests, by a “heterogeneous” mixing process—where states are influenced more strongly by states with which they share some connection or characteristic. For example, because state ideology appears to be a significant predictor of timing of film incentive adoption, one might expect that ideological distance might mediate the diffusion process. That is, states with ideologies similar to previous adopters might be more likely to also adopt film incentives. However, the ideological distance variable was significant in neither of the models. Another heterogeneous mixing process would be my first hypothesis that states are more likely to adopt

film incentives if their neighbors have. In the Cox model, I actually find that having additional neighbors who have adopted film incentives (looking at *PCTBORDERADOPT* and *CAN_BORDER*) actually *reduces* the hazard that a given state might adopt them. This result is somewhat puzzling and certainly not consistent with the stories of Mississippi and Washington, which saw the adoption of incentives by neighboring jurisdictions as adding pressure for these states to adopt their own incentives. Nevertheless, once the homogenous mixing process is explicitly modeled in the exponential model, *PCTBORDERADOPT* and *CAN_BORDER* are no longer significant (though the coefficient values are similar). The role of regional diffusion, therefore, is unclear. If indeed having neighbors adopt a policy reduces a given state's hazard, it could suggest that states perceive a first mover advantage at the regional level—that is, the gains dissipate as more nearby states adopt incentive policies.

Despite the strong evidence for a bandwagon-style homogenous mixing process, it is always difficult to rule out other undetected diffusion patterns or exogenous influences to explain the pattern of positive duration dependence seen in the Cox baseline hazard. The total number of adopters, as well as many “heterogeneous” diffusion variables such as *PCTBORDERADOPT*, is strictly increasing with time, so it can be difficult to separate out the effects of these variables from each other and from time itself (i.e. exogenous influences). It could be the case that *NUM_PREV* is simply absorbing the effects of other diffusion variables and/or exogenous influences that affect all states. For example, one possibility is that after Canadian provinces increased their film incentives in 2005, it set off a second wave of state film incentive adoptions. While this alternative interpretation cannot be ruled out, the qualitative evidence does not support it.

The diffusion processes suggested by this study of film incentives do not fit neatly into the diffusion mechanism categories of learning, competition, imitation, and coercion proposed by Shipan and Volden (2008) and others. Part of the issue is that there are two levels of information dissemination and diffusion: the film industry level and the state policymaker level. The flow of information at the film industry level is consistent with the workings of a competitive market with highly-mobile production abilities. Film producers learn about the increasing importance of incentives. At the state policymaker level, states learn about film incentive programs enacted in other states. But this learning by both the film industry and by state policymakers clearly has a competitive element to it—they learn about what others are doing in order to compete with them. This study suggests that future studies should push to further clarify and distinguish the key differences between the mechanisms of learning and competition.

What is relevant and even alarming about this study of film incentives is that state decisions are driven so strongly by a combination of an outside interest group influence and a bandwagon style political trend. This analysis provides no evidence that states account for their individual characteristics or even their similarities or relationships to other states when deciding whether to adopt film incentives. Either state film incentives are perceived as unambiguously good policies for every state, which would be surprising given that Button (2014) has found that film incentives are ineffective at creating local film industries, or the bandwagon effect has trumped any political or economic considerations that might make states hesitate to adopt them. Resolving this question is beyond the scope of this analysis, but understanding the factors that do and do not have an effect on the timing of state film incentive adoption is an important first step.

References

- Baybeck, B., Berry, W. D., & Siegel, D. A. (2011). A strategic theory of policy diffusion via intergovernmental competition." *Journal of Politics*, 73(1), 232-247.
- Beamer, G. (2002). Elite interviews and state politics research. *State Politics & Policy Quarterly*, 2 (1), 86-96.
- Berry, F. S. & Berry, W. D. (2007). Innovation and diffusion models in policy research. In Sabatier, P. (Ed.), *Theories of the Policy Process* (pp. 223-260). Boulder, CO: Westview Press.
- Berry, W. D., Fording, R. C., Ringquist, E. J., Hanson, R. L., & Klarner, C. E. (2010). Measuring citizen and government ideology in the U.S. states: A re-appraisal. *State Politics & Policy Quarterly*, 10(2), 117-135.
- 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.
- Berry, W. D., & Baybeck, B. (2005). Using geographic information systems to study interstate competition. *American Political Science Review*, 99(4), 505-519.
- Best, S. J., & Teske, P. (2002). Explaining state Internet sales taxation: New economy, old-fashioned interest group politics. *State Politics & Policy Quarterly*, 2(1): 37-51.
- Boehmke, F. J., & Witmer, R. (2004). Disentangling diffusion: The effects of social learning and economic competition on state policy innovation and expansion. *Political Research Quarterly*, 57(1), 39-51.
- Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling : A guide for social scientists*. Cambridge; New York: Cambridge University Press.
- Button, P. (2014). *Can Motion Picture Production Incentives Create a Local Film Industry?* Unpublished manuscript.
- Department of Canadian Heritage. (2012) Canadian film policy: History of film incentives. Retrieved from <http://www.pch.gc.ca/pgm/em-cr/evaltn/2008/prt3-eng.cfm>.
- Fletcher, J., & Murray, M. (2006). Competition over the tax base in the state sales tax. *Public Finance Review*, 34(3), 258-281.
- George, A. L., & Bennett, A. (2005). *Case Studies and theory development in the social sciences*. Cambridge, Mass.: MIT Press.
- Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3), 515-526.

- Grossback, L. J., Nicholson-Crotty, S., & Peterson, D. A. M. (2004). Ideology and learning in policy diffusion. *American Politics Research*, 32(5), 521-545.
- Klarner, C. (2013) Klarner Politics. Retrieved from <http://www.indstate.edu/polisci/klarnerpolitics.htm>.
- Leech, B. L. (2002). Interview methods in political science. *Political Science and Politics*, 35(4), 663-664.
- Lofton, L. (2004, May 10). New incentives enhance film industry opportunities in state. *Mississippi Business Journal*, Retrieved from <http://www.msbusiness.com>.
- Luther, W. (2010, January). *Movie production incentives: Blockbuster support for lackluster policy*. Washington, D.C.: Tax Foundation.
- Lyne, J. (2000, November 13). Nissan driving \$950 million, 4,000-employee plant to Mississippi delta. *Site Selection*, Retrieved from <http://www.siteselection.com>.
- McDonald, A. (2011). Down the rabbit hole: The madness of state film incentives as a "solution" to runaway production. *University of Pennsylvania Journal of Business Law*, 14(1), 85-165.
- Miller, C., & Richard, B. (2010). The policy diffusion of the state R&D investment tax credit. *State and Local Government Review*, 42(1), 22-35.
- Mossberger, Karen. (2000). *The politics of ideas and the spread of Enterprise Zones*. Washington, DC: Georgetown University Press.
- Omer, T. C., & Shelley, M. K. (2004). Competitive, political, and economic factors influencing state tax policy changes. *The Journal of the American Taxation Association*, 26(Supplement), 103-126.
- PriceWaterhouse Coopers. (2006). *The big table of film and video incentives in Canada*. Retrieved from <http://www.pwc.com/ca/en/entertainment-media/film-video-tax-incentives-canada.jhtml>.
- Rork, J. C. (2003). Coveting thy neighbor's taxation. *National Tax Journal*, 56(4), 775-788.
- Shipan, C. R., & Volden, C. (2012). Policy diffusion: Seven lessons for scholars and practitioners. *Public Administration Review*, 72(6), 788-796.
- Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, 52(4), 840-857.
- Squire, P. (2007). Measuring state legislative professionalism: The Squire index revisited. *State Politics and Policy Quarterly*, 7(2), 211-227.

Strang, D. (1991). Adding social structure to diffusion models. *Sociological Methods & Research*, 19(3), 324-353.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, 64(5), 416-424.

Zorn, C. J. W. (2000). Modeling duration dependence. *Political Analysis*, 8(4), 367-380.

Table 1: Variables and data sources

Variable	Description	Data Source
Dependent variable: film incentive adoption dates	Exact day of approval for each state incentive program	State legislative websites
<i>PCTBORDERADOPT</i>	Percentage of bordering states that have already adopted incentives	Author calculations
<i>IDEO_DIST</i>	$ABS([(MostRecentAdopterIdeology + AllOtherAdoptersIdeology) / 2] - PotentialAdopterIdeology)$	Author calculations based on Grossback, Nicholson-Crotty, and Peterson's (2004) and <i>IDEOLOGY</i> measure
<i>LN_FILM_IND</i>	Natural log transformation of real GDP associated with the "motion picture and sound recording industries" for each state. 1996 figures are extrapolated.	BEA
<i>CAN_BORDER</i>	Dummy variable where 1 = state shares a border with Canada	Author
<i>UNEMP</i>	Annually-averaged unemployment rate for each state	BLS
<i>LN_PERCAPGSP</i>	Natural log transformation of real GDP per capita by state	BEA
<i>RATE</i>	Top state corporate income tax rate as of January 1 of each year. States without a corporate income tax rate are coded as 0.	Tax Foundation, Council of State Governments
<i>PROFESS100</i>	Index of state legislative professionalism * 100. Raw scores range from 0.0 (least professional) to 1.0 (most professional). Annual scores are interpolated or extrapolated from Squire's (2007) data.	Squire (2007)
<i>IDEOLOGY</i>	Measure of the partisan ideology of state governments based on Berry, et.al.'s (2010) <i>NOMINATE</i> measure. Scores range from 0 (most conservative) to 100 (most liberal).	Berry, et.al. (2010)
<i>DIVIDED</i>	Dummy variable where 1 = a single party does not control the governorship and both branches of the legislature	Klarner (2013) State Partisan Balance Data http://www.indstate.edu/polisci/klarnerpolitics.htm
<i>NUM_PREV</i>	Total number of states previously adopting film incentives	Author calculations

Table 2: Cox model estimates for state film incentive adoptions

Variable	Hazard Ratio	Robust Std. Error	P-value
<i>PCTBORDERADOPT</i>	0.979	0.010	0.032
<i>IDEO_DIST</i>	1.007	0.013	0.602
<i>LN_FILM_IND</i>	1.504	0.227	0.007
<i>CAN_BORDER</i>	0.419	0.220	0.097
<i>UNEMP</i>	0.755	0.170	0.211
<i>LN_PERCAPGSP</i>	0.146	0.183	0.125
<i>RATE</i>	0.945	0.054	0.321
<i>PROFESS100</i>	1.009	0.017	0.577
<i>IDEOLOGY</i>	1.015	0.008	0.058
<i>DIVIDED</i>	0.973	0.386	0.945
Log-pseudolikelihood	-124.62995		
Observations	1565		

Notes: Results are based on observations of 48 states (excluding LA and CA), of which there were 42 failures. Standard errors are clustered by state.

Table 3: Exponential model estimates for state film incentive adoptions

Variable	Hazard Ratio	Robust Std. Error	P-value
<i>PCTBORDERADOPT</i>	0.989	0.009	0.250
<i>IDEO_DIST</i>	1.009	0.013	0.484
<i>LN_FILM_IND</i>	1.393	0.170	0.007
<i>CAN_BORDER</i>	0.530	0.272	0.217
<i>UNEMP</i>	1.032	0.186	0.860
<i>LN_PERCAPGSP</i>	0.274	0.268	0.185
<i>RATE</i>	0.974	0.054	0.632
<i>PROFESS100</i>	1.003	0.018	0.880
<i>IDEOLOGY</i>	1.015	0.008	0.043
<i>DIVIDED</i>	0.984	0.349	0.963
<i>NUM_PREV</i>	1.357	0.111	0.000
<i>NUM_PREV2</i>	0.996	0.002	0.004
Constant	481.481	4937.921	0.547
Log-pseudolikelihood	-27.821855		
Observations	1565		

Notes: Results are based on observations of 48 states (excluding LA and CA), of which there were 42 failures. Standard errors are clustered by state.

Figure 1: Interview Questions

1. Tell me about your role your state's adoption of film incentives.
2. Where did the initial idea for film incentives come from?
3. Why did the legislature decide to take up the film incentives?
4. How were the details of the incentives policy decided upon?
5. What types of arguments were made (and by whom) in support or in opposition?
6. Why do you think these incentives were adopted when they were? (i.e. why didn't it happen sooner, later, or not at all?)
7. Anywhere in this process, did you *hear about* what other states were doing?
8. Anywhere in this process, did you *seek out* information about what other states were doing?
9. Which states? Why those?
10. What types of information and where did you get it?
11. How did (didn't) this information influence the political process or content of the film incentive program?

Figure 2: Kaplan-Meier cumulative failure plot for state film incentive adoptions

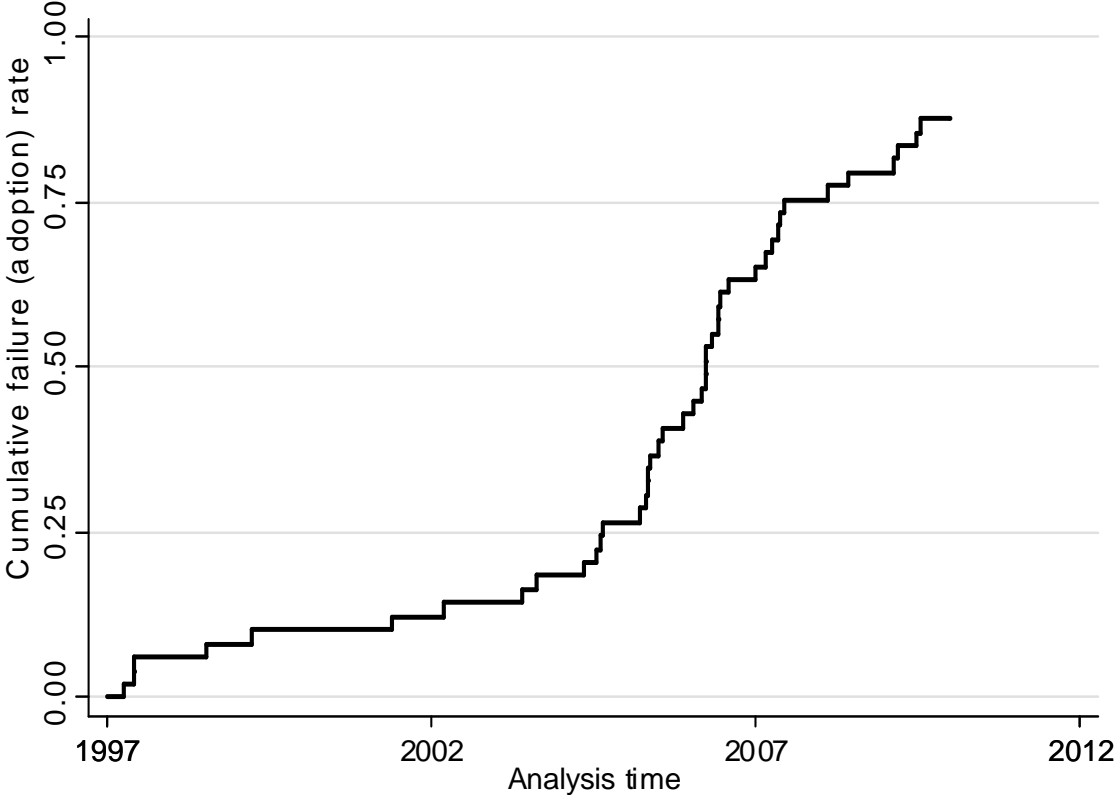
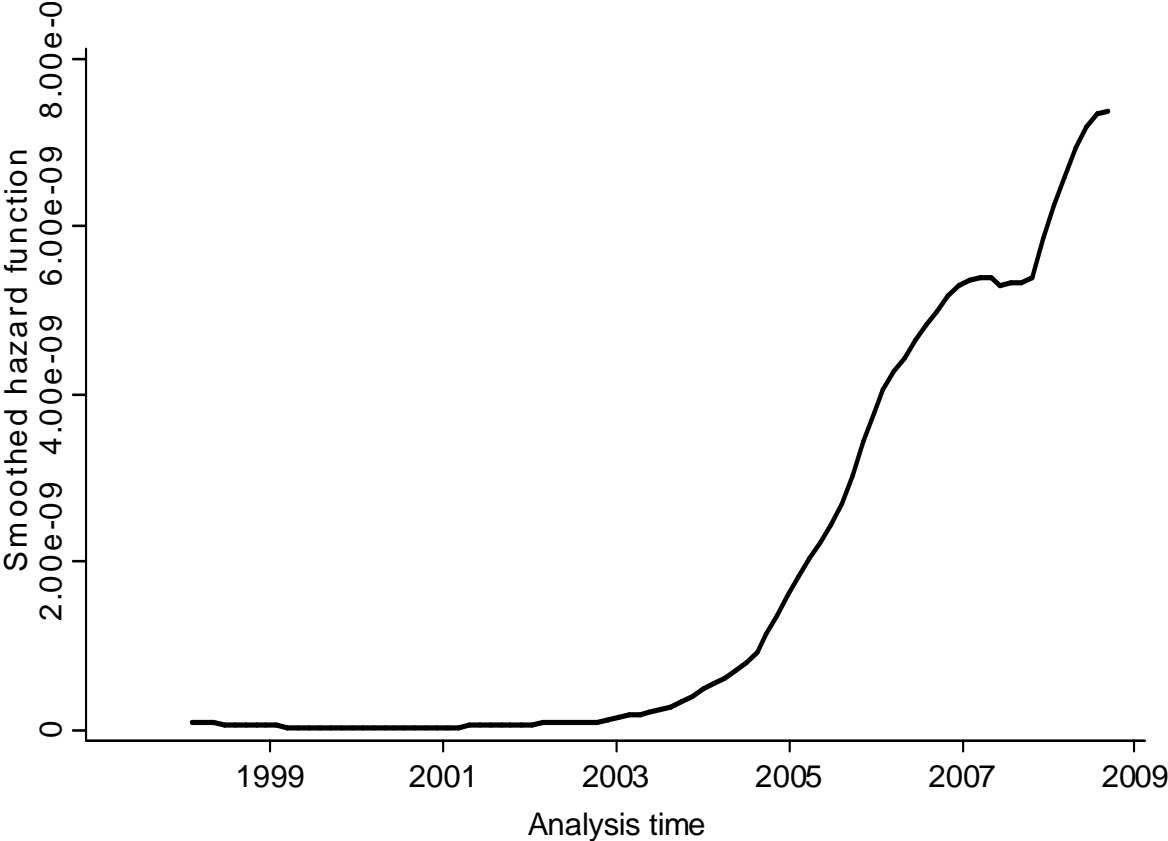


Figure 3: Cox baseline hazard for state film incentive adoptions



Appendix A: Adjudicating Between Parametric Models

To find the best fitting parametric model, I estimated results for five parametric distributions: exponential, Weibull, Gompertz, log-logistic, and log-linear. Each model includes all covariates included in the Cox model, plus either a linear or quadratic specification for *NUM_PREV*. For each model, I calculated the AIC and BIC:

Distribution	Specification of <i>NUM_PREV</i>	AIC	BIC
Exponential	linear	90.974	155.242
Exponential	quadratic	81.644	151.267
Weibull	linear	92.727	162.351
Weibull	quadratic	80.922	155.901
Gompertz	linear	89.435	159.059
Gompertz	quadratic	83.644	158.623
Log-logistic	linear	112.958	182.581
Log-logistic	quadratic	87.540	162.519
Log-linear	linear	102.568	172.192
Log-linear	quadratic	101.058	176.037

The AIC and BIC do not agree on which model fits the data best. The model with the lowest AIC is the quadratic Weibull model (AIC = 80.922), while the model with the lowest BIC is the quadratic exponential model (BIC = 151.267). Fortunately, the one-parameter exponential distribution is nested in the two-parameter Weibull distribution, so a likelihood ratio test can also be used to compare them. The likelihood ratio test essentially determines whether the extra parameter in the Weibull model (i.e. the “shape” parameter that allows for a non-monotonic hazard) significantly contributes to the fit of the model. The p-value for the likelihood ratio test is 0.099, which is not strongly compelling. Therefore, I err on the side of parsimony and determine that the quadratic exponential model has the best fit. Fortunately, model estimates for the quadratic Weibull model (not reported, available upon request) show nearly identical patterns of significance and hazard ratio values as the exponential model, so the choice between the exponential and Weibull models does not affect the substantive conclusions of this essay.

Essay Two: The Diffusion of State Business Tax Incentives

In the past several decades, state tax incentives for businesses have proliferated. A recent investigation of state business incentive programs by the New York Times counted 1,874 programs on which states spent an estimated \$80.4 billion per year (Story et al., 2012). Some interpret this trend optimistically, describing states as “laboratories of democracy” spreading innovative tax policy ideas throughout the country. Others argue that the spread of tax incentives is a pernicious “race to the bottom” where states undercut each other to compete for a finite amount of business investment or jobs. Although studies by Chirinko and Wilson (2008), Goolsbee and Maydew (2000), Neumark (2013), and Wilson (2009) have shown mixed assessments of whether business tax incentives are successful in creating new jobs and investment, we still know very little about why states choose to adopt them.

This essay focuses on four of the most popular types of tax incentive policies: investment tax credits (ITCs), apportionment formula changes, research and development (R&D) tax credits, and job creation tax credits (JCTCs) to determine what factors influence states’ decisions in adopting these incentives. I draw on the literatures in state tax competition and policy diffusion to develop a theoretical framework that models how states learn about and decide whether to adopt tax incentives. I use this framework to develop hypotheses about political, economic, and diffusion-related variables that influence the learning and adoption process and then test the hypotheses using event history analysis techniques.

The results of the analysis contribute to the empirical literature on state tax policy and economic development and provide a more nuanced understanding of the factors that influence states’ adoptions of tax incentives, including evidence that is consistent with the idea that states are “racing to the bottom”—that is, adopting tax incentives in response to an overall nationwide trend rather than adopting them based on state-specific circumstances such as political or

economic climate. Considering the influence of state-specific variables, patterns of adoption and diffusion are quite different across the four types of incentives. For example, while having a greater number of workers in the manufacturing industry significantly increases the hazard of adopting ITCs and JTCs, it does not appear to have an effect on decisions to adopt apportionment changes or R&D credits. In addition, unlike many previous studies in state tax competition and policy diffusion, I do not find that states are consistently influenced by whether their neighboring states have also adopted tax incentives. Neighbors have a significant effect only for R&D credits, but not for the other types of tax incentives.

The analysis also makes a methodological contribution to the literature on policy diffusion by demonstrating the importance of modeling how diffusion processes change over time. This study emphasizes a thorough investigation of the baseline hazard (i.e. duration dependence) and non-proportional hazards (i.e. time-varying effects). The baseline hazard in particular is the key to interpreting the general trends in states' adoptions of tax incentives while controlling for state characteristics, and provides the evidence that suggests states may be racing to the bottom.

State Tax Competition

Two largely separate strands of literature in economics and political science have attempted to discern whether and how states influence one another in adopting tax policy. In economics, tax policy interdependence is interpreted as competition and is studied in the literature on tax competition. The classic theoretical view, associated with (Oates, 1972) and formalized by Zodrow and Mieszkowski (1986; as reviewed in J. D. Wilson, 1999) is based on the intuition that when deciding whether to lower taxes, governments account for the marginal

benefits associated with the fact that some residents or businesses will move into the jurisdiction if taxes are lowered. However, they ignore the marginal costs associated with the fact that losing those residents and businesses reduces the tax bases of other jurisdictions. A social planner overseeing all jurisdictions would follow the standard Samuelson rule and take both effects into account, but because single jurisdictions only account for their own marginal benefits of an enhanced tax base, they have incentives to set their tax rates at inefficiently low levels. Once one jurisdiction has lowered its tax rate, for example, by adopting a tax incentive, other jurisdictions may try to prevent their tax bases from fleeing by also lowering their tax rates, triggering a “race to the bottom.” Researchers who support the classic view tend to see tax competition as destructive and efficiency-reducing.

Another view of tax competition combines the insights of Tiebout’s (1956) model of local public goods provision and Brennan and Buchanan’s (1980) “Leviathan” model (J. D. Wilson, 1999). In Tiebout’s model, under a restrictive set of assumptions (including assuming away inter-jurisdictional spillovers), perfectly mobile residents will efficiently sort themselves into homogenous communities that reflect their preferences for various public goods. Brennan and Buchanan’s model argues that governments are motivated to maximize their budgets instead of social welfare (Niskanen, 1971), suggesting that competition acts to restrain the otherwise excessive growth of government. Combining these two perspectives in a “hybrid” model, Tiebout-style sorting acts as a disciplining mechanism for local governments to minimize their costs—if another jurisdiction can provide the same level of public goods for a lower tax price, then all residents will move there. Pressures for cost minimization then act as a check on Leviathan-style government growth. Under this “hybrid” model, tax competition is lauded as the essential element that forces state and local governments to be innovative and efficient.

The main difference between these models is in their assumptions about government motivations and behavior. Both models, however, predict that tax rates will be lower under tax competition than they would otherwise be, and that competitive pressures force jurisdictions to mirror the policies of their competitors. The empirical work on horizontal tax competition focuses on this last feature: studies attempt to determine whether a “reaction function”—the effect of an increase/decrease in one jurisdiction’s tax rate on another jurisdiction’s tax rate—is actually positive. These studies use spatial lag variables to measure the effect of other states’ tax rates. Several studies have estimated reaction functions for state business tax rates. For example, Rork (2003) finds a positive reaction coefficient of 0.16 for bordering state corporate income tax rates between 1967 and 1996. In other words, a 10% increase in the average corporate income tax rate of border states results in a 1.6% increase in the home state’s rate. Rork also finds that higher corporate income tax rates are positively associated with single party Democratic control of state government and larger elderly populations and negatively associated with state per capita income.

Deskins and Hill (2010) investigate state tax reaction functions from the late 1970s to 2006. For state corporate income tax rates, they find a significant positive reaction coefficient of 0.12¹¹ between the home state rate and the population-weighted average rate of bordering states. The reaction coefficient is not significant if border state rates are weighted by the distance between the states’ largest cities or if no weights are used. They also find that state corporate income tax rates are positively associated with personal income and sales tax rates and negatively associated with Republican state government. Deskins and Hill try to explain the weak evidence for interdependence in corporate income tax rates by pointing out that businesses

¹¹ They also interact the spatial lag terms with a quadratic time trend to see if the reaction coefficient changes over time, but find all interaction terms to be insignificant at conventional levels.

have more opportunities to reduce their taxes through tax planning or tax incentives, such as the ones studied in this essay, so statutory tax rates are of diminished policy importance.

Fox, Hill, and Murray (2014) estimate reaction functions for state business taxes, including the business share of property taxes, while also controlling for vertical tax competition from local governments. They estimate that a 10% increase in the unweighted average tax rate of neighboring states results in a 3.55% increase in the home state's rate, but the reaction coefficient is not significant when border state tax rates are weighted by geographic distance between the states' largest cities. Among other results, they find that state business taxes are affected by population, per capita income, and political party control.

In contrast to other empirical findings, Chirinko and Wilson (2013) find evidence of negative reaction functions for state corporate income tax rates, Investment Tax Credit rates, and capital apportionment weights, the latter two of which will be discussed in more detail below. They argue that the appearance of a positive reaction function (i.e. a race to the bottom) is an illusion created by all states responding similarly to exogenous shocks such as foreign competition. They suggest that their observed negative reaction function implies that if State A's neighbor lowers its ITC rate (reduces its generosity), it will cause capital to flow out of State A and into State B. State B then uses this "windfall" increase in the tax base to reduce tax rates and keep overall revenues level.¹² In Chirinko and Wilson's account, therefore, interdependence in state tax policy is not a matter of states reacting to policy changes in other states but of states reacting to changes in their own tax bases (which are only incidentally caused by tax rate changes in other states).

¹² Chirinko and Wilson argue that one state's reaction to a tax increase/decrease in another state depends on the income elasticity of preferences for public versus private goods. If State A decreases its tax rate, drawing tax base away from State B, State B can respond to decreased tax revenue either by cutting public spending (if the relative preference is for private goods) or increasing taxes to make up the difference (if the relative preference is for public spending). The latter instance is a case of a negative reaction function.

In general, there is no clear consensus in the tax competition literature on the nature of interdependence in state business tax policy. What is clear from the tax competition literature, however, is that results can be very sensitive to model specifications. For example, as Chirinko and Wilson (2013) show, the reaction function can flip signs depending on how cross-sectional and time dependence are modeled. In addition, the use of spatial lags makes it very tempting to try multiple complex weighting schemes to find one that produces significant results, but this can quickly turn into a “fishing expedition” without theoretical guidance.

Policy Diffusion and State Tax Policy

The other major areas of scholarship that focus on how states influence each other in adopting policies are the policy interdependence and policy diffusion literatures in political science. Unlike the tax competition literature in economics, which implicitly assumes that the nature of tax policy interdependence is competitive, political scientists posit a variety of interdependence mechanisms, including learning, imitation, coercion, common norms, and taken-for-grantedness, in addition to competition (Braun, Gilardi, Fuglister, & Luyet, 2007; Shipan & Volden, 2008). A major emphasis in the current literature is to examine how multiple mechanisms may operate in a single policy context. For example, Basinger and Hallerberg (2004) and Gilardi and Wasserfallen (forthcoming) show that competitive pressures in international tax competition are mitigated, respectively, by domestic political conditions and the socialization and development of common norms among policymakers.

However, the research on policy interdependence and diffusion has been criticized for being inconsistent and sometimes contradictory in how various diffusion mechanisms are operationalized and measured in empirical research (Graham, Shipan, & Volden, 2013; Maggetti

& Gilardi, 2013). For example, many studies estimate the effect of a variable that captures policies in bordering states. This kind of variable usually represents a process of regional learning (e.g. F. S. Berry & Berry, 1992), but could just as easily be interpreted as evidence of competition or imitation. As a result, there has been a renewed emphasis on clarifying the theoretical underpinnings of policy interdependence and diffusion studies (Shipan & Volden, 2012).

Related to the broad literature on policy interdependence, policy diffusion research usually studies the adoption of discrete policy innovations. The dependent variable, therefore, is usually a dichotomous indicator of whether a particular policy was adopted by the state (or other observational unit) during some interval of time (often a year), or alternatively the duration of time until the state adopts the policy.¹³ Policy diffusion models have been used to study a handful of state tax-related policies, including the adoption of state income and sales taxes (F. S. Berry & Berry, 1992; Hayes & Dennis, 2014); enterprise zones (Mossberger, 2000), tax and expenditure budgets (Metcalf, 2012), the deductibility of federal income taxes (Hayes and Dennis, 2014), apportionment formula changes (Omer & Shelley, 2004), and R&D tax credits (Hearn, Lacy, & Warshaw, forthcoming; Miller & Richard, 2010). I review the last three, which are specifically related to the taxation of business, in detail in the next section.

State Business Tax Incentives

As Makse and Volden (2011) and Boehmke and Skinner (2012) have noted, there is a relative lack of policy diffusion studies that examine more than one policy, making it difficult to generalize their findings. Makse and Volden's study provides evidence that characteristics of the

¹³ Time-series—cross-section data are similar to panel data except that they generally have more time periods (T) and fewer cross-sections (N). As Beck, Katz, and Tucker (1998) explain in detail, time-series—cross-section data with a binary dependent variable are identical to duration data grouped by time intervals.

policies themselves, such as levels of complexity and observability, affect how quickly and through which channels diffusion occurs. An implication of their work is that policies with similar characteristics should show similar patterns of diffusion. This study examines four of the most popular types of state business tax incentives: Investment Tax Credits, apportionment formula changes, R&D tax credits, and Job Creation Tax Credits. All four of these incentives are designed to reduce the tax burden on in-state business and encourage increases in in-state jobs and investment, and they are all broadly available to most or all businesses, as opposed to narrowly targeted at a specific firm or industry. Although it should be expected that there will be differences in the factors that affect the adoption of, say, an R&D credit versus an ITC, these four incentives should be similar enough to reflect any major themes in states' general approaches to tax incentives for business. In addition, the individual incentive policies are similar enough from one state to the next so that adoption by one state is comparable to adoption in another state. This does not mean, for example, ITCs are identical in all states, but as discussed in more detail below, the four types of incentives in this study have been recognized by other authors and tax professionals as belonging to the same general categories.

The tax incentives in this study were chosen partly because they are relatively widespread among states—in other words, I oversample cases of tax policy in which diffusion is most likely to have occurred as part of a “most-likely” case selection strategy. If only a handful of states had adopted a policy, the sample size would be too small to find significant relationships between variables. Using this sampling strategy requires a slight qualification of the research question from: how do tax incentives diffuse to how do tax incentives diffuse, *given that they diffuse at all?* The former question is an interesting one and some authors have explored how attributes of

the policies themselves affect the speed and extent of diffusion (Makse & Volden, 2011), but the question is beyond the scope of this essay for business tax incentives.

Investment Tax Credits

State investment tax credits (ITCs) provide businesses with a credit against their corporate income taxes typically equal to some percentage of in-state capital investment during the prior year and are intended to encourage businesses to invest in-state instead of elsewhere. As Chirinko and Wilson (2006) describe, state ITCs have proliferated over the past several decades, and they have also increased in average generosity. The first state to adopt an ITC was New York in 1969, and by 2010, 25 total states had adopted ITCs.¹⁴ Based on visual inspection of US maps, Chirinko and Wilson (2006) suggest that there have been four geographically-clustered “episodes” of ITC adoptions: in the Northeast between 1969 and 1975, in the Central Midwest between 1975 and 1986, in the Rust Belt between 1995 and 2003, and in the Southeast between 1995 and 1996.

Chirinko and Wilson (2008) find evidence that while state ITCs do seem to increase investment in states that adopt them, this increased investment appears to be at the expense of investment in other states, suggesting that states are in a “race to the bottom.” However, as mentioned above, Chirinko and Wilson (2013) find a negative reaction function for state ITC rates, meaning that one state’s ITC rate is inversely related to nearby states’ rates. They also find evidence that more generous ITCs are associated with Democratic state governments and smaller state populations, but they argue that most of the downward pressure on state tax burdens on

¹⁴ California and Maine subsequently repealed their credits. For the purposes of this study, however, I am only interested in the initial adoption. Why states abandon business incentives would be an interesting area for future research.

capital is a result of states responding similarly to exogenous shocks, including macroeconomic conditions and global competitive pressures.

Apportionment Formula Changes

In addition to using various tax credits, states are also increasingly manipulating their corporate income tax apportionment formulas to attract jobs and investment (Edmiston, 2002; Goolsbee & Maydew, 2000). Formula apportionment is a method used by states to determine what proportion of a business' total income can be taxed by that state. It requires firms to calculate their total national income and multiply it by an "apportionment factor," which is supposed to be a measure of the firm's presence in a particular state. The apportionment factor ϕ in state j is determined by three underlying factors: the property factor, the payroll factor, and the sales factor:

$$\phi_j = \alpha_j^P \frac{P_j}{P} + \alpha_j^L \frac{L_j}{L} + \alpha_j^S \frac{S_j}{S}$$

P , L , and S are the firm's total property, payroll and sales respectively, and P_j , L_j , and S_j are property, payroll, and sales in state j . The α terms are weights assigned to each of the factors by state j and usually sum to one.

For this study, the policy change of interest is a state moving from an equally weighted formula ($\alpha^P = \alpha^L = \alpha^S = 1/3$) to a formula that more heavily weights the sales factor ($\alpha^S > 1/3$, while $\alpha^P = \alpha^L < 1/3$). The logic is that having too much weight on the property and payroll factors punishes businesses that have made investments and created jobs in a state. For example, automotive manufacturers have lobbied Michigan to abandon the property and payroll factors in favor of a single sales factor ($\alpha^S = 1.0$, while $\alpha^P = \alpha^L = 0$) because they have a higher

concentration of their property and payroll in Michigan, while their sales are spread out over the whole world.

Prior to 1978, most states had equally-weighted three-factor apportionment formulas as a result of a push for uniformity by the Multistate Tax Commission. However, the Supreme Court's decision in *Moorman Manufacturing Company v. Bair* (437 U.S. 267) in 1978 upheld Iowa's single sales factor apportionment formula and essentially gave states free rein to abandon equally-weighted formulas and set their own apportionment policies. In 1978, only Iowa, Florida, Massachusetts, New York, and Wisconsin used apportionment formulas that put greater weight on the sales factor. As of 2010, 33 states had increased the weight on their sales factors (Bernthal et al, 2012).

There have been several studies of the effects of apportionment changes on employment (for a review, see Bernthal et al., May 2012). One influential study by Goolsbee & Maydew (2000) shows that a state's increasing the weight on its sales factor (with corresponding decreases on the property and payroll factor weights) significantly increases employment in that state, at the expense of employment in other states. Similarly, Edmiston's (2002) simulation model suggests that once one state has moved to single sales factor apportionment, as Iowa did, it is in the interest of other states to do so as well to try to capture some of the economic benefits of increased employment or to prevent the loss of jobs to other states. Edmiston also shows that there is a distinct first mover advantage, and the economic benefits dissipate as more states adopt single sales factor apportionment policies.

Omer and Shelley (2004) use an event history model to examine the diffusion of apportionment formula changes. They hypothesize that diffusion could take two forms: "strategic tax competition" and "regional diffusion." To measure competition, they use two

variables that measure the number of previous adopters among neighboring states that also share similar industries or that also share cross-border MSAs. They control for regional diffusion with a variable that measures just the number of neighboring states previously adopting the change, which is intended to reflect neighbor influence regardless of competitive relationships. They find that both competition variables have significant positive effects on the likelihood of adoption and that their regional diffusion variable is not significant. They also find evidence that poorer states with higher unemployment rates are also more likely to change their apportionment formulas.

In contrast to Omer and Shelley's findings, Chirinko and Wilson (2013) find a negative reaction function for the capital apportionment factor weight, which states usually reduce at the same time they increase the sales factor weight. As with state corporate income tax rates and ITC rates, they argue that decline in capital apportionment weight since the late 1970's is not the result of interstate tax competition but rather exogenous shocks to which all states respond.

R&D Tax Credits

Similar to ITCs, state R&D credits allow businesses to offset their corporate income taxes by an amount equal to some percentage of their in-state R&D expenditures in the previous year. States began adopting R&D credits shortly after the creation of a federal R&D credit in August of 1981. The first state to adopt an R&D credit was Minnesota in 1982, and as of 2010, 38 states had adopted them. Wilson (2009) found that R&D credits do increase R&D spending in states with the credits, but the overall pattern is again a zero-sum game among states.

Two recent studies have examined the spread of state R&D credits from an economic development policy perspective. Miller and Richard (2010) use both interview data and

quantitative event history data to test hypotheses that R&D credit adoption is related to various factors including economic conditions, existing R&D activity, and adoption by other states. They model diffusion in two different ways. In their first model, they include a variable for the percentage of neighboring states that have adopted. In their second model, they develop a measure to match each state with five most “similar” states in terms of population, patent activity, manufacturing share, and income and operationalize the diffusion variable as the percentage of “similar” states that have adopted. They find evidence that an increased hazard of adoption is associated with higher unemployment, larger manufacturing share of total jobs, and higher corporate income tax rates. In both models, the diffusion variable is significant, but the coefficients suggest that increases in adoptions by neighboring or “similar” states actually decrease the likelihood of adoption. This finding is akin to negative reaction function results from tax competition studies.

Hearn, et al. (forthcoming) also look at state adoptions of R&D credits using event history analysis and seek to test hypotheses related to higher education, political and economic conditions, and diffusion. They find that adoption is associated with higher unemployment rates, higher patenting activity, and Republican governors, among other findings. They find no significant effect for their diffusion variable, the number of neighboring states adopting¹⁵, and interpret this as evidence against geography-based interstate competition. Instead, they suggest a “leader-laggard” type diffusion process where the leader states are “well known for preexisting economic development, strong research universities, and technological innovation, including California, Minnesota, Massachusetts, and Illinois” (p. 12).

¹⁵ This is a slightly different formulation than the *proportion* of neighboring states adopting, which is used in Miller and Richard (2010) and this essay. I discuss this point in more detail below.

Job Creation Tax Credits

State Job Creation Tax Credits (JCTC) provide corporate income taxpayers credits for net job creation. Like the R&D credit, the federal government started the trend when it adopted a “New Jobs Tax Credit” in 1977. While the overall design of state JCTCs is similar, state credits vary by how they define “new” jobs, whether the credit is based on employee compensation or withholding taxes, and on the credit amount or rate itself (Chirinko & Wilson, September 2010). Connecticut and Ohio were the first states to adopt a JCTC in 1992, and by 2010, a total of 24 states had adopted a JCTC.

Chirinko and Wilson note that the total number of adopters over time seems to follow an S-shaped pattern, with few states adopting at first, followed by a growing rate of adoption, and then tapering off. They also note that states with JCTCs tend to be concentrated in the East. Their preliminary analysis indicates that JCTCs do have a modest positive effect on in-state employment, but they have not yet conducted any analyses to determine whether there is any net job growth when considering all 50 states. A recent review article by Neumark (2013) concludes that state JCTCs (he refers to them as hiring credits) probably have modest positive effects on job creation, but using this policy, states spend between \$9,100 and \$75,000 per job created. Neumark points to a study by Bartik and Erickcek (2010), which estimates that at least 92% of credits are for jobs that would have been created anyway.

As this summary of four major state tax incentives shows, the prevailing thinking among tax policy researchers is that state tax incentives have small positive effects on jobs and/or investment, but are not particularly cost-effective for state governments. Moreover, as many have suggested and Chirinko and Wilson (2008) and Wilson (2009) have confirmed, tax

incentives are likely a zero-sum game between states—one state’s gain is another state’s loss. Of the studies that have attempted to determine why states adopt or change these policies, the results are mixed. Some point to interstate competition as a significant factor (Omer & Shelley, 2004), while others argue that state characteristics or exogenous factors matter more (Chirinko & Wilson, 2013; Hearn et al., forthcoming). Nevertheless, states continue to adopt these incentives, and if interstate influence is not a factor, it is remarkably coincidental that states are adopting such similar policies to reduce tax burdens on business.

A Model of Policy Adoption and Diffusion

In developing a theoretical framework from which to derive hypotheses about the adoption of state business tax incentives, I try to respond to several criticisms of the policy diffusion literature: vague conceptualization and operationalization of diffusion mechanisms (Maggetti & Gilardi, 2013), failure to recognize the conditional nature of diffusion (Neumayer & Plumper, 2012; Shipan & Volden, 2012) and failure to analyze and interpret the baseline hazard (Bennett, 1999; Carter & Signorino, 2010; Zorn, 2000).

The first step in developing a theoretical framework is to recognize the differences between conceptualizations of learning and competition in the tax competition literature in economics and the policy diffusion literature in political science. As mentioned above, learning is not explicitly recognized as a mechanism of interdependence in economics. The traditional economics assumption is one of “perfect information,” in which any information is instantly available to all actors, who immediately analyze it and respond rationally—that is, competitively or strategically. Therefore, in neoclassical economic theory, learning is an automatic homogenous function of all actors. In political science, however, learning is an imperfect and

heterogeneous process that takes place over time. Actors are usually assumed to be “boundedly rational,” relying on heuristics and satisficing in order to make decisions with limited resources of attention, cognitive capacity, and information (Elkins & Simmons, 2005). In this way, the process of learning will be different for actors with different resources. As Mossberger (2000) describes, states vary widely in what kinds of information they are exposed to, what kinds of information they seek out, and how they use that information to make decisions. In the policy diffusion literature, the definition of learning is usually limited to learning about the success or failure of a policy in other adopting states (Shipan & Volden, 2008), a process sometimes called instrumental learning (May, 1992). However, more recently, Shipan and Volden (2012) have pointed out that actors may also seek to learn about the political aspects of a policy, also known as political learning (May, 1992). May (1992) also proposes that policymakers engage in “social learning,” which involves changes in the social construction of the policy problem, expectations, and goals.

For this study, I attempt to find a middle ground between the conceptions of learning in economics and political science. Drawing from economics, I argue that learning is ubiquitous and automatic and that most new information about a given policy is quickly and widely available to all actors in a system. An implication of this argument is that it is highly improbable that the same policies would arise in two different states entirely independently. It is unrealistic to assume that a state developing a policy idea would not at a minimum have heard if another state has already adopted the policy. Such a situation would require willful ignorance. Even policy diffusion scholars are starting to acknowledge that with modern communication technology and interstate organizations like the National Conference of State Legislatures acting

as clearinghouses of information, today's policymakers have access to vast amounts of policy information (Shipan & Volden, 2012).

However, while neoclassical economic models usually assume that actors can instantly evaluate these vast amounts of information, I argue that bounded rationality is a more realistic view of the decision-making capacity of policymakers. The sheer quantity of information available to policymakers makes the development of decision heuristics and shortcuts almost necessary. As evidenced by the growth and popularity of behavioral economics, there is some indication that economists are conceding that people don't always behave as strict utility maximizers. Even to the extent that they can make decisions rationally, policymakers are often faced with what economists call uncertainty and asymmetric information. For example, if a particular policy has proven successful in one state, that is not a guarantee that it will be successful in another (uncertainty), or an adopting state may attempt to obfuscate unsuccessful policy experiments.

Figure 1 outlines a model of policy adoption and diffusion based on this conception of learning. The first step in a diffusion process occurs when the first state considers adopting a policy, and this provides the raw material from which other states can learn. Information about state policy adoptions may come from media coverage, the legislation itself, policy analysis documents, documents created by interest groups, etc. (Karch, 2007). Early in the diffusion process, it is likely that the only information available to other states is that a policy exists and that certain states have adopted it. Over time, as more states adopt the policy, new information will likely emerge as various stakeholders evaluate the success of the policy, as implementation challenges are met, and as interest groups or political parties take positions on the policy. However, as Elkins and Simmons (2005) point out, sometimes even after the policy starts to

spread, states may still have little information beyond knowing which other states have already adopted the policy.

Information relevant to a particular policy can also be generated by events outside of the diffusion process—in methodological terminology, exogenous shocks. For example, the rate of adoption of a state tax incentive may be influenced by national or global macroeconomic conditions such as the business cycle or changes in capital mobility. These exogenous factors will tend to affect all states in similar ways and may create the illusion of interdependence when in reality states are independently reacting to a common stimulus, an issue known as Galton’s problem (Franzese & Hays, 2008).

All of this policy-relevant information, whether generated endogenously through accumulated information about state adoptions or by some exogenous event, is publicly available and easily accessible to state policymakers. Obviously, however, states vary in whether and how they access and respond to this information. As illustrated in the second box of Figure 1, I argue that political and economic characteristics affect how quickly states learn, how they interpret information, which pieces of information are deemed important, and/or how quickly states react to information. For example, a state that is struggling economically may be actively searching for information about policy ideas designed to increase jobs and may be more willing to experiment with policies, whereas a state that is doing well may only hear about a policy idea in passing. States may also be unsure whether a policy idea is compatible with their ideological predisposition and look for cues from the ideologies of previous adopters (Grossback, Nicholson-Crotty, & Peterson, 2004). If states perceive that there is a competitive element to a policy, as is often the case with tax incentives, then they may be particularly responsive to information about whether their competitors, however defined, have adopted the policy. It is

important to note that this process of learning does not require that what is learned be true in an objective sense. States are simply satisficing (Simon, 1956), so the learning process is expected to be imperfect. The final step in Figure 1 is for the state to decide, given what it has learned, whether to adopt the policy and whether it has the capacity to do so. In the next section, I develop hypotheses about how certain state-level variables condition the learning process when states are deciding whether to adopt business tax incentives.

The learning process in this model of policy adoption and diffusion can also be visualized in terms of the baseline hazard in event history analysis, which illustrates how the hazard of adopting a policy changes over time, controlling for the effects of covariates (i.e. holding covariates at their mean values). In other words, the baseline hazard is an illustration of duration dependence. It traces the hazard for an average or representative state and reflects how the general impression of the policy changes over time. A period of rising hazard would show that the policy is generally becoming more popular over time—perhaps adopters have had some success or macroeconomic conditions encourage adoption. A period of falling hazard would show that the policy is generally becoming less popular over time—perhaps there is growing doubt about the policy’s effectiveness or macroeconomic conditions discourage adoption. This last point is worth reiterating because it allows for the possibility for learning to result in non-adoption. As Elkins and Simmons (2005) have noted, the diffusion literature can sometimes wrongly define diffusion as an outcome—the widespread adoption of a policy—instead of diffusion as a process involving interdependence.

Data and Hypotheses

Table 1 summarizes the variables, data sources, and hypotheses. Each row for the dependent variables shows the year that is used for the start of risk, which is when the diffusion process is considered to begin. For the ITC, the start of risk is 1970, the year after the first state (NY) adopted an ITC, and New York is excluded from the analysis. For apportionment formula changes, the start of risk is 1979, the year after the *Moorman* decision, and Iowa, Florida, Massachusetts, New York, and Wisconsin are excluded from the analysis. For the R&D credit, the start of risk is 1981, the year the federal R&D credit was enacted. For the JCTC, the start of risk is 1991.¹⁶ All models in this study exclude Nevada, South Dakota, Texas, Washington, and Wyoming, states that do not have corporate income taxes. As with all event history data, once a state has adopted the incentive, it is right-censored and contributes no further observations to the analysis. Figure 2 provides Kaplan-Meier failure plots to show the cumulative proportion of states adopting each incentive over the time of the analysis. Typical diffusion processes will have an S-shaped Kaplan-Meier plot (Rogers, 2003), which appears to be plausible for three out of the four plots. However, the JCTC plot appears to have a more concave shape, similar to Boushey's (2010) "policy outbreaks," which are diffusion patterns initiated by exogenous shocks.

Table 2 provides descriptive statistics for the independent variables described in this section. Statistics are shown for each incentive for two years: the start of risk and the last year in the study, with adopters and non-adopters separated for comparison.

The first independent variable measures the percentage of neighboring states that have already adopted the policy. This variable is very common in policy diffusion studies, but is used inconsistently by authors and is therefore difficult to interpret. It has been used to operationalize

¹⁶ While the federal government adopted the "New Jobs Tax Credit" in 1977, states did not start adopting JCTCs until the 1990s.

mechanisms of learning, emulation, and competition (Maggetti & Gilardi, 2013), and is probably best described as a proxy for many possible channels of interdependence. Whether a state is a neighbor is therefore a heuristic for whether a state is similar in terms of a variety of economic, political, social, or geographic characteristics. While more precise diffusion variables could be defined based on various spatial weighting schemes, for the present purposes of comparing interdependence across four different policies, a more broadly-defined measure is preferred. In addition, using this variable allows the results of this study to be compared to other studies. In the context of the theoretical framework used in this study, influence from geographic neighbors is not a source of learning, which assumes that somehow geographic neighbors share policy information that is not available to other states. Instead, I argue that states may be likely to actively seek out information about whether their neighbors have adopted a policy for two possible reasons:

- Neighbors are viewed as similar and therefore worthy of emulation. In this way, states can free-ride on the presumably well-researched policy decisions of neighbors.
- Neighbors are viewed as competitors, and states will feel competitive pressure to adopt the same incentives.

Both of these explanations predict that an increase in the neighbors variable will increase the hazard of adoption, but as Chirinko and Wilson (2013) have argued, if other states reduce their tax burdens, it may draw tax base from the home state, which will then raise its taxes to keep revenues unchanged. This explanation is consistent with the neighbors variable reducing the hazard but is *not* evidence of policymaking interdependence because policymakers are not necessarily aware why their tax base has shrunk.

While neighbors may be a proxy for overall similarity, state policymakers may also seek out information about states that are ideologically similar (Grossback et al., 2004). For example, if a conservative state is considering adopting a policy, it may help to know about whether the states that have already adopted the policy also tend to be conservative. If so, the state can take it as a signal that the policy aligns with conservative political values and strategies. The closer the ideological alignment with prior adopters, the more confidence the state will have in also adopting the policy, and vice versa for greater ideological distance, which is expected to reduce the hazard of adoption. The ideological distance variable, based on Grossback, et al. (2004) is equal to the absolute value of the difference between the state government ideology of the potential adopter and a weighted average of the ideologies of all previous adopters. The most recent adopter is given a 50% weight, and the other 50% encompasses all other previous adopters.¹⁷ Underlying ideology scores are from Berry, et al.'s (2010) *NOMINATE* measure of state government ideology, which ranges from 0 (most conservative) to 100 (most liberal). *NOMINATE* is an aggregate measure that accounts for partisan affiliation and power in the governor and state legislature.

I also control for party control of the governorship and state legislatures using data from Klarner (2013).¹⁸ The governor party variable is a dummy variable equal to one when the Governor is a Democrat and zero when he or she is a Republican.¹⁹ I expect that Democrats will generally be skeptical of policies that are considered tax breaks for business, so both of these variables are expected to reduce the hazards of adoption.

¹⁷ Results of the models reported below are robust to alternative specifications of ideological distance, including giving all prior adopters equal weight and squaring the difference instead of using the absolute value.

¹⁸ Using Berry's (2010) single measure of government ideology is an alternative, but there is some evidence that governors in particular take an active role in promoting tax incentives (Hearn et al., forthcoming).

¹⁹ Out of 2,050 total observations, there are 24, mostly in Maine and Minnesota, where the governor's party is "other" and the variable takes on a value of 0.5.

Another important variable that affects how states can learn about and react to information about policy ideas is legislative professionalism. I use Squire's (2007, p. 211) index of state legislative professionalism, which is intended to capture a state's "capacity... to generate and digest information in the policymaking process." States with legislatures that are more "professional" have more session days, larger staffs, and higher pay. Squire's professionalism scores range from 0.0 (least professional) to 1.0 (most professional) and are multiplied by 100. States with higher levels of professionalism should have greater capacity to make better-informed decisions about policy adoption and greater capacity to adapt policy ideas to their states' own specific circumstances. Therefore, I expect that higher professionalism will be associated with increased hazards of adopting tax incentives.

Manufacturing businesses are often the targeted recipients of business tax incentives because they often make highly visible investments in capital and labor and are relatively R&D intensive compared to other industries. I expect that the size of a state's manufacturing industry, measured as the logged total number of employees in the manufacturing sector, will influence the policy learning process in several ways. A larger manufacturing sector will be able to exert more political influence on state policymakers. They will promote a favorable interpretation of information on tax incentives and lobby and advocate for incentives over the course of the legislative process. Perhaps even more importantly, manufacturers, along with the business community in general, may actually act as a channel for the diffusion of ideas. Because they naturally feel competitive pressure within their own industries, they will always be seeking ways to gain a competitive advantage. Convincing states to give them tax breaks is one increasingly popular way to accomplish this. Larger and more robust state manufacturing industries will have a greater capacity to both generate and promote tax incentive policies in their states. Therefore,

manufacturing employment should be associated with increased hazards of adopting tax incentives.

Because larger states will tend to have more manufacturing employees simply because of their size, I also include logged population in the models. The expected effect of population, however, is ambiguous. On one hand, there is evidence that larger states tend to be more innovative, perhaps because they have more “slack resources” and therefore may be more willing to take risks in adopting policies (Boehmke & Skinner, 2012). However, studies on “asymmetric tax competition” have shown that jurisdictions with smaller populations have an incentive to be more aggressive in reducing taxes because the cost of capital is more sensitive to tax rate changes in smaller jurisdictions (J. D. Wilson, 1999).

State economic conditions will also affect how states learn about tax incentive policies. Difficult economic times may make state policymakers more likely to actively seek out information about policies that have the potential to increase jobs and investment. Therefore, I expect that the unemployment rate to be associated with increased hazards of adoption and logged real state per capita income to be associated with reduced hazards of adoption. Similarly, I expect that higher corporate income tax rates will tend to increase the hazard of adopting incentives. Facing a high tax rate, in-state businesses are more likely to suggest policy ideas and lobby policymakers to adopt incentive policies to reduce their tax burdens. Policymakers may also be more receptive to, or even actively seek out, policies that can reduce the tax burden on in-state businesses.

Finally, I hypothesize that the baseline hazards for each of these tax incentives—which show the rates of adoption controlling for covariates—will be shaped like an inverted U, with a period of rising hazard followed by a falling hazard. If prior studies about state business tax

incentives are correct in the characterization of tax incentives as a zero-sum game (Chirinko & Wilson, 2008; Goolsbee & Maydew, 2000; D. J. Wilson, 2009), it means that the potential benefits of adopting a tax incentive (e.g. increased jobs and/or investment) will fall as more states adopt them—a “first mover advantage”—but it may take time for some states to realize this. The hazard of adoption should rise at first, as information begins to spread and more states become aware of the zero-sum nature of the situation. Over time, however, as the benefits of adopting tax incentives are dissipated, the hazard of adoption should eventually peak and begin to fall.

Results

Cox Model Results

The data are first analyzed using a Cox proportional hazards model, in which the hazard h that state i adopts the incentive at time t is a function of the baseline hazard h_0 and time-varying covariate values $\mathbf{x}_i(t)$:

$$h_i(t) = h_0 \exp [\boldsymbol{\beta}' \mathbf{x}_i(t)]$$

The hazard can be thought of as a conditional rate of adoption. It expresses the probability of adoption at time t given that the state has not yet adopted the incentive. Coefficients are reported as hazard ratios, which are similar to the odds ratios reported in logit regression. Coefficients greater (less) than one indicate that a unit increase in the covariate corresponds to an increase (decrease) in the hazard of adopting the incentive. Standard errors are clustered by state and tied observations are handled using the Efron method (Efron, 1977). Results are presented in Table 3.

The first thing to notice upon inspection of Table 3 is that very few covariate results are consistent across all four types of business tax incentives. The set of variables that significantly affect the hazard for one kind of incentive is not the same set of variables that affect the hazard for another type of incentive, which is somewhat surprising because the tax incentive policies were chosen to be as similar as possible. In fact, for apportionment changes, none of the coefficients achieve significance, although a Wald test indicates that the model as a whole is an improvement on the null model with no covariates ($p = 0.001$).

Consistent with many other policy diffusion studies, a higher percentage of neighboring states adopting an incentive appears to be associated with a greater hazard of adoption for the home state. For all four incentive types, the coefficients for this variable are greater than one, although only for the R&D credit does it achieve statistical significance. This finding contrasts with the results of Miller and Richard (2010), who find that the percentage of neighbors adopting significantly decreases the hazard, and Hearn, et al. (forthcoming), who find no significant effect associated with neighbor adoptions.

The results for ideological distance are also contrary to expectations. For all four incentive types, the coefficients are greater than one (with statistical significance achieved for R&D credits and JCTCs), indicating that a larger ideological distance between prior and potential adopters is actually associated with an *increased* hazard of adoption.

Results for the political party control variables are mixed. Coefficients for the Democratic governor variable all suggest that it reduces the hazard of adoption, although none reach statistical significance. Coefficients for legislative control, however, suggest that Democratically-controlled legislatures are generally *more* likely than Republican-controlled legislatures to adopt business tax incentives, another surprising finding. The legislative control

variable is significant for ITCs and JCTCs, indicating that fully Democratically-controlled state legislatures are 5.7 times as likely to adopt an ITC and 3.7 times as likely to adopt a JCTC as fully Republican-controlled legislatures, all else equal.

The Cox model results also show that legislative professionalism has a relatively small effect on the hazard of adoption. It is significant for the ITC and JCTC, but a ten point increase in the professionalism score (on a 100 point scale) only corresponds to a 20-30% increase in the hazard of adoption, all else equal. As expected, the number of manufacturing employees in a state is associated with an increased hazard of adopting all four types of tax incentives, and significantly so for the ITC and JCTC. For example, a 1% increase in the number of manufacturing employees more than triples the hazard of adopting an ITC, all else equal. The results for the population variable suggest that for all four incentive types, smaller states are more likely to adopt, consistent with the hypothesis that smaller jurisdictions compete more aggressively in cutting taxes. The population variable is significant only for the ITC model, indicating that a 1% larger population is associated with an 84% reduction in the hazard of adopting and ITC, all else equal.

The economic conditions variables also have mixed results. A one percentage point increase in the unemployment rate, all else equal, significantly raises the hazard of adopting an R&D credit by 30%, a quite large effect, but does not significantly affect the hazard of adopting ITCs, apportionment changes, or JCTCs (and may actually lower the hazard of adopt for JCTCs). Real per capita income is not a significant predictor in any of the models. Finally, it is not entirely clear that higher state corporate income tax rates uniformly result in an increased hazard of adopting tax incentives. While a one percentage point increase in the tax rate, all else equal, is associated with a statistically significant 39% increase in the hazard of adopting an R&D credit,

it does not significantly affect the hazard for the other incentives, and may even reduce the hazard for ITCs.

In order to investigate the hypotheses about duration dependence, the baseline hazard must be inspected. The baseline hazard can be thought of as the underlying effect of time, controlling for the effects of the covariates in the model. As explained above, it traces the hazard of adoption for an average or representative state and reflects how the general impression of the incentive policy changes over time. A period of rising (falling) hazard indicates that the incentive is generally becoming more (less) popular over time. As Box-Steffensmeier and Jones (2004) explain, one advantage of the Cox model is that it does not require the researcher to make any restrictive assumptions about the shape of the baseline hazard or duration dependence. Therefore, the plot of the non-parametric Cox baseline hazard can give an idea of the overall shape of the duration dependence. However, because no parameters or standard errors are estimated for the Cox baseline hazard, it cannot be used for statistical inference.

Figure 3 shows kernel-smoothed plots of the baseline hazards for each tax incentive, with covariates set to their mean values (and *Democratic Governor* set to 0). Except for the JCTC, all of the hazard plots follow an inverted U-shaped pattern, indicating that the incentives initially grew in popularity and then declined. The JCTC hazard plot suggests perhaps a short period of rising or constant hazard through 1996-97, but mainly shows falling hazards over the course of the analysis. The hazard reaches a maximum around 1995 for the ITC, 1991 for apportionment changes, and 2000 for R&D credits. The inverted U-shape of the hazard is suggestive of a zero-sum game or “race to the bottom.” As states gradually realize they are engaged in zero-sum competition for jobs and investment, the hazard should initially rise. Then, as more states adopt the incentives, the first mover advantage dissipates and the comparative advantage of having an

incentive declines, causing the hazard to fall. Another possible reason for the decline in the hazards is that some states are simply resistant to adopt incentives for idiosyncratic reasons (akin to immunity from infection). Another possibility is that after several years of experience, states begin to learn that the incentives are not as effective as initially anticipated, making the remaining non-adopters less likely to adopt them.

The rise and fall of hazard rates may also be affected by exogenous factors such as macro political trends. Figure 3 suggests that hazards were generally rising during the 1980s and into the early 1990s for ITCs, apportionment changes, and R&D credits. This coincides with the general popularity of tax cuts as an economic policy during the Reagan era and Bush eras. Similarly, all four incentives show declining hazards by the late 90s during Clinton's presidency. However, if conservative presidential politics tend to increase the use of state business tax incentives, one would expect to see them become more popular after George W. Bush became president in 2001, but the hazards continue to fall for all four types of incentives. In general, the shape of the baseline hazard plots is consistent with a combination of explanations that are both endogenous and exogenous to the process of learning and diffusion.

Non-Proportional Hazards Model Results

The non-monotonic shape of the baseline hazard highlights the importance of dynamics in the learning and diffusion process and suggests that the factors influencing adoption of tax incentives may be changing over time. Therefore, it is important to test the assumption of proportional hazards, which implies that the effects of covariates are consistent over time. For example, in their study of R&D tax credits, Hearn et al. (forthcoming) find that having a Republican governor initially increases the hazard of adoption but the effect weakens over time.

I test the proportional hazards assumption using Grambsch and Therneau's (1994) procedure and find that three out of the four models have covariates that show statistically significant ($p < 0.15$) violations of the proportional hazards assumption, meaning that the proportional hazards assumption is not appropriate and the models are not properly specified.

To allow for non-proportional hazards, I follow the advice of Box-Steffensmeier and Jones (2004) and interact the offending covariates with a time variable measured in years.²⁰ Table 4 shows the results of the non-proportional hazards models, including time interactions for covariates that initially failed the proportional hazards test.

Starting with the ITC, many of the results are substantively similar to the original Cox results, except for the effect of legislative professionalism. The coefficient for professionalism in the original Cox model suggests that higher levels of professionalism are associated with an increased hazard of adopting ITCs, all else equal. However, in the non-proportional hazards model, both the primary effect and the time interaction are statistically significant, meaning that the effect of professionalism is actually time-varying. At the beginning of the analysis, in 1970, a ten point increase in the professionalism score (on a scale of 0-100) reduces the hazard of adopting an ITC by about 57%, all else equal. Each subsequent year, however, increases the hazard ratio by 0.004. Therefore, by the 15th year, 1985, a ten point increase in the professionalism score *increases* the hazard of adoption by 3% [for one point: $-(1 - 0.943) + 15(1 - 1.004) = 0.003$], and the effect continues to grow larger each subsequent year. A possible explanation for this finding is that more professional states were initially skeptical of ITCs because there was still very little evidence about the potential effectiveness of ITCs in increasing net jobs and investment in other states. Perhaps they could also more clearly see that ITCs

²⁰ While the convention is to interact covariates with logged time, as Box-Steffensmeier, Reiter, and Zorn (2003) explain, there really is no theoretical or mathematical reason to prefer any particular function of time. Therefore, favoring a straightforward interpretation, I simply interact covariates linearly with time.

would be expensive in terms of foregone revenue. Less professional states may have been initially naïve about the revenue expense. Over time, as the evidence supporting ITCs grew, initially reluctant professional states surpassed their less professional peers in the rate of adoption.

Looking at the results for apportionment changes, several coefficients are now significant and consistent with the hypotheses, whereas none of the coefficients were significant in the original Cox model results. Unlike the original Cox model, the results for ideological distance suggest that a greater ideological gap between prior and potential adopters reduces the hazard of adopting apportionment changes, all else equal, but as the time interaction variable indicates, the effect diminishes over time. The Democratic governor variable is also now significant, suggesting that having a Democratic governor instead of a Republican governor reduces the hazard of adopting apportionment changes by 56%, all else equal. The results also show that having a one percentage point higher unemployment rate or corporate income tax rate significantly increases the hazard of adopting apportionment changes, by 57% and 53% respectively, all else equal. Although not significant, the coefficients for their corresponding time interactions further suggest that the effects of unemployment rates and corporate income tax rates may diminish over time.

Turning to the R&D credit, including non-proportional hazards does not correct the interpretation of the original Cox model so much as it adds nuance. The effect of the percentage of border states already adopting the R&D credit is still large and significant, in contrast to the findings of previous studies, and surprisingly, increases in ideological distance significantly increase the hazard of adoption, all else equal. The results for Democratic governor provide yet another example of the importance of testing for and modeling non-proportional hazards. While

this variable was not significant in the original Cox model, both Democratic governor and its time interaction are significant in the non-proportional hazards model. The table shows that having a Democratic instead of Republican governor initially reduces the hazard of adoption by about 95%, all else equal, but this effect changes over time, similar to prior results (Hearn et al., forthcoming). By the 7th year of the analysis (1988), Democratic governors actually have an 11% higher hazard of adoption [$-(1 - 0.053) + 7(1 - 1.151) = 0.11$], and the hazard ratio continues to rise by 0.15 each year. Similar to the results from apportionment changes, I also find that higher unemployment rates and corporate income tax rates, all else equal, are associated with 62% and 101% greater hazards of adopting R&D credits. The corresponding time interactions are also significant, suggesting that the effects gradually diminish over time.

Finally, the non-proportional hazards model results for the JCTC are not dramatically different from the Cox proportional hazards results. Adoption of the JCTC appears to be primarily influenced by the number of manufacturing employees and ideological distance. Greater ideological distance between prior and potential adopters is still associated with an *increased* hazard of adopting JCTCs. A one percent increase in the number of manufacturing employees, all else equal, is associated with a 262% increase in the hazard of adopting a JCTC. The professionalism variable, while of similar magnitude as the original Cox model, has fallen just below statistical significance ($p = 0.11$). Similarly, Democratic control of the state legislature was significant in the original Cox model, but failed the test for proportional hazards. However, when the time interaction is included in the model, both variables fall short of statistical significance, perhaps because of lack of statistical power. The original Cox model suggests that a fully Democratically-controlled legislature has a 367% higher hazard of adopting a JCTC, all else equal. The non-proportional hazards model suggests that if the effect is time-

varying, it starts out smaller, increasing the hazard by 59% in 1991, but growing by ten percentage points per year.

For all four types of tax incentives, taking account of time-varying effects adds nuance and in some cases corrects the erroneous conclusions of the original Cox proportional hazards models. The most dramatic differences between the proportional hazards and non-proportional hazards models are for apportionment changes. While no variables achieve statistical significance under proportional hazards, several do once non-proportional hazards are appropriately modeled.

Testing for Unobserved Heterogeneity

Another important step in properly specifying event history models for which duration dependence is of substantive importance is accounting for unobserved heterogeneity. As Zorn (2000) explains, unobserved heterogeneity creates “spurious” negative duration dependence, making the baseline hazard flatter if it is rising and/or steeper when it is falling. The standard procedure for accounting for unobserved heterogeneity in event history analysis is by incorporating a “frailty” parameter into the models, which is analogous to random effects in panel data methods (Box-Steffensmeier & Jones, 2004).

While it is possible to incorporate a random frailty distribution into the Cox model, the Cox model will still derive the baseline hazard non-parametrically, and there may be remaining duration dependence. However, I wish to constrain the model to a constant hazard so that no remaining duration dependence is allowed and to see whether the random effects can explain the duration dependence observed in the Cox baseline hazard plots—that is, to see whether the

observed duration dependence is “spurious.” Instead of using a Cox model, therefore, I switch to a complementary log-log (cloglog) model.

Like the more familiar logit and probit models, the cloglog model is used for binary dependent variables. However, it is especially appropriate for event history data for two reasons. First, cloglog is the binary data analog to the continuous time Cox model, so its estimates are more directly comparable to Cox model estimates than logit or probit (Carter & Signorino, 2010). Like the Cox model, cloglog coefficients can be interpreted as effects on the hazard. Second, unlike the logit and probit distributions, which are symmetric around a probability of 0.5, the cloglog distribution is asymmetric, with a fat tail near probability = 0 and increasing rapidly approaching probability = 1. Because event history data are right-censored, observations of $y = 1$ are relatively rare compared to observations for which $y = 0$. Therefore, the cloglog link function is more theoretically appropriate than logit or probit for event history data (Buckley & Westerland, 2004).

To test for unobserved heterogeneity, I estimate cloglog models for each incentive with random effects and no specification for duration dependence, which effectively constrains the models to a constant baseline hazard. I include the same sets of covariates as were used in the non-proportional Cox models, including the time interactions. The results show that random effects parameters are not significant at conventional levels ($p < 0.10$) for three out of the four models (results not shown), suggesting that unobserved heterogeneity is not biasing estimates of duration dependence. For the apportionment change model, the random effects parameter was weakly significant with $p = 0.098$. However, the results of the cloglog random effects model are quite similar to the non-proportional Cox model, so they are not reported.

Cloglog Model Results and Duration Dependence

Now that I have accounted for non-proportional hazards and ruled out unmeasured heterogeneity, I use cloglog models to test for any remaining duration dependence. For each incentive, I run models with the following specifications for duration dependence: constant hazard, linear time trend, quadratic time trend, logged time, cubic splines (Beck, Katz, & Tucker, 1998), and a cubic polynomial (Carter & Signorino, 2010). Most studies only use one specification for duration dependence—time dummies and cubic splines are very common. While these specifications are flexible, they are not very conducive to interpretation, and should arguably not even be used if a more parsimonious specification for duration dependence fits the data better. Moreover, as argued above, duration dependence should be of primary interest in policy diffusion studies, not just a nuisance to be controlled for. Once the six specifications for duration dependence are estimated, model fit is compared using likelihood ratio tests, corroborated with AIC and BIC scores. Results for the best fitting models for each incentive are reported in Table 5. Coefficients are exponentiated to allow for easier comparison to the Cox model results.

For three out of the four tax incentive models, the ITC, R&D credit, and JCTC, a constant hazard provides a better fit than any other models that include duration dependence, suggesting that any duration dependence has been adequately modeled by the inclusion of the time interactions. For these three tax incentives, the results of the cloglog models are (not surprisingly) quite similar to the results of the non-proportional hazards models. The biggest difference is that for R&D credits, the unemployment rate and its time interaction are no longer significant in the cloglog model.

For apportionment changes, including a variable for logged time creates the best fitting model. The coefficient for logged time indicates that when time interaction variables are held to zero, there is still significant positive but diminishing duration dependence. In other words, apportionment changes continue to grow in popularity over time, but at a diminishing rate. Other than the inclusion of logged time, however, the results of the apportionment change cloglog model are substantively similar to the non-proportional hazards model, except that in the cloglog model, the party affiliation of the governor is no longer significant.²¹

Discussion and Conclusions

Weighing all of the evidence from the Cox models, non-proportional hazards models, and cloglog models, most of the findings about the factors that influence states' decisions to adopt tax incentives are consistent with hypothesized expectations. However, there are few similarities across the four different types of tax incentives. None of the variables were significant for all four types of tax incentives, and only ideological distance was significant for three of the four, apportionment changes, R&D credits, and JCTCs. Consistent with expectations, greater ideological distance is associated with a reduced hazard of adopting apportionment changes, and the effect diminishes over time. However, for R&D credits and JCTCs, increased ideological distance actually *increases* the hazard of adopting these incentives, all else equal, a result that is very difficult to explain. Even using different formulations of the ideological distance variable did not substantially change these conflicting results.²²

²¹ In addition, once duration dependence is modeled with logged time, if random effects are also included in the model, the random effects parameter is no longer significant. AIC and BIC scores indicate that the model with logged time (and no random effects) fits better than the model with random effects (and no duration dependence).

²² Results are robust to alternative specifications of ideological distance, including giving all prior adopters equal weight and squaring the difference instead of using the absolute value.

Unlike many other policy diffusion and tax competition studies, I do not find strong evidence that states are consistently influenced by their neighbors in deciding to enact business tax incentives. The variable measuring the percentage of border states that have already adopted an incentive significantly increases the hazard of adoption only for R&D credits. While coefficient values are greater than one for apportionment changes and JCTCs, the effects are not significant. For ITCs, results from the non-proportional hazards and cloglog models suggest that having a larger proportion of bordering state adopters may actually reduce the hazard of adoption, although the coefficient is not significant. If it were significant, this finding would be consistent with prior findings of a negative reaction function for state ITC rates (Chirinko & Wilson, 2013). Chirinko and Wilson argue that a negative reaction function does not necessarily imply that policymaking is interdependent, but only that a change in one state's tax rate can have effects on the size of the tax base in other states. Alternatively, perhaps if more neighbors have adopted, a state may believe that there is more potential benefit in distinguishing itself in other ways instead of adopting the same incentive as its neighbors.

Results for the other political and economic variables also do not suggest any strong conclusions about why states adopt tax incentives. Although not significant for all four types of incentives, coefficient estimates suggest that more manufacturing employees, higher unemployment rates, higher corporate income tax rates, and smaller populations may be associated with an increased hazard of adopting business tax incentives. However, the roles of legislative professionalism and party control of the executive and legislative branches remain unclear. Contrary to the expectation that Republicans would be more likely to support tax incentives, Democratic support sometimes appears to be a relevant factor, especially for ITCs. Professionalism is only significant for ITC adoptions, but its effect changes over time.

One clear result of this analysis of the adoption of business tax incentives is the importance of recognizing how the process of adoption may change over time. Policy diffusion is necessarily a dynamic process. The more states adopt a policy, the more their collective experience shapes and informs subsequent adoption decisions in other states. Therefore, the standard regression approach—to model time-invariant effects and view duration dependence as simply a nuisance to be controlled for—is inadequate. Event history methods provide analytical tools such as non-proportional hazards and baseline hazard plots to properly model changes over time, but they are generally under-utilized by policy diffusion scholars. As this study of tax incentives shows, the regression results are quite sensitive to the proper modeling of non-proportional hazards. When non-proportional hazards were added to the apportionment change model, for example, several variables became significant. In addition, a thorough analysis of baseline hazard plots can give an overview of how the adoption and diffusion process works at a broader level. The inverted U-shaped baseline hazards for the four tax incentives in this study suggest that perhaps states do view the use of tax incentives as a zero sum game. The hazard rises as states become aware of and race to benefit from the first mover advantage, and then as more states adopt incentives, the comparative advantage of having tax incentives shrinks, and the hazard of adoption falls.

References

- Bartik, T. J., & Erickcek, G. A. (2010). *The employment and fiscal effects of Michigan's MEGA tax credit program*. (Upjohn Institute Working Paper No. 10-164). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Basinger, S. J., Hallerberg, M. (2004). Remodeling the competition for capital: How domestic politics erases the race to the bottom. *The American Political Science Review*, 98(2), 261-276.
- Bennett, D. S. (1999). Parametric models, duration dependence, and time-varying data revisited. *American Journal of Political Science*, 43(1), 256-270.
- Bernthal, J., Gavril, D., Schumacher, K., Spencer, S., & Sydor, K. (May 2012). *Single sales-factor corporate income tax apportionment: Evaluating the impact in wisconsin*. (Prepared for the Wisconsin Department of Revenue). Madison, WI: Robert M. LaFollette School of Public Affairs, University of Wisconsin-Madison.
- Berry, W. D., Fording, R. C., Ringquist, E. J., Hanson, R. L., & Klarner, C. E. (2010). Measuring citizen and government ideology in the U.S. states: A re-appraisal. *State Politics & Policy Quarterly*, 10(2), 117-135.
- 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.
- Boehmke, F. J., & Skinner, P. (2012). State policy innovativeness revisited. *State Politics and Policy Quarterly*, 12(3), 303-329.
- Boushey, G. (2010). *Policy diffusion dynamics in America*. New York, NY: Cambridge University Press.
- Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling : A guide for social scientists*. Cambridge; New York: Cambridge University Press.
- Braun, D., Gilardi, F., Füglistner, K., & Luyet, S. (2007). Ex pluribus unum: Integrating the different strands of policy diffusion theory. In K. Holzinger, H. Jörgens & C. Knill (Eds.), *Transfer, diffusion und konvergenz von politiken* (pp. 39-55) VS Verlag für Sozialwissenschaften.
- Brennan, G., & Buchanan, J. M. (1980). *The power to tax: Analytical foundations of a fiscal constitution*. Cambridge: Cambridge University Press.
- Buckley, J., & Westerland, C. (2004). Duration dependence, functional form, and corrected standard errors: Improving EHA models of state policy diffusion. *State Politics and Policy Quarterly*, 4(1), 94-113.

- Carter, D. B., & Signorino, C. S. (2010). Back to the future: Modeling time dependence in binary data. *Political Analysis*, 18(3), 271-292.
- Chirinko, R. S., & Wilson, D. J. (2006). *State investment tax incentives: What are the facts?* (Working Paper 2006-49). San Francisco: Federal Reserve Bank of San Francisco.
- Chirinko, R. S., & Wilson, D. J. (2008). State investment tax incentives: A zero-sum game? *Journal of Public Economics*, 92(12), 2362-2384.
- Chirinko, R. S., & Wilson, D. J. (2013). *Tax competition among U.S. states: Racing to the bottom or riding on a seesaw?* (Working Paper 2008-03) San Francisco: Federal Reserve Bank of San Francisco.
- Chirinko, R. S., & Wilson, D. J. (2010). *Job creation tax credits and job growth: Whether, when, and where?* (Working Paper 2010-25) San Francisco: Federal Reserve Bank of San Francisco.
- Deskins, J., & Hill, B. (2010). Have state tax interdependencies changed over time? *Public Finance Review*, 38(2), 244-270.
- Edmiston, K. (2002). Strategic apportionment of the state corporate income tax. *National Tax Journal*, 55(2), 239.
- Efron, B. (1977). The efficiency of cox's likelihood function for censored data. *Journal of the American Statistical Association*, 72(359), 557-565.
- Elkins, Z., & Simmons, B. (2005). On waves, clusters, and diffusion: A conceptual framework. *The Annals of the American Academy of Political and Social Science*, 598(1), 33-51.
- Fox, W. F., Hill, B. C., & Murray, M. (2014). Vertical competition, horizontal competition, and mobile capital: An empirical investigation. *Public Finance Review*, advance access.
- Franzese, R. J., & Hays, J. C. (2008). Interdependence in comparative politics: Substance, theory, empirics, substance. *Comparative Political Studies*, 41(4-5), 742-780.
- Gilardi, F., & Wasserfallen, F. (forthcoming). How socialization attenuates tax competition. *British Journal of Political Science*.
- Goolsbee, A., & Maydew, E. L. (2000). Coveting thy neighbor's manufacturing: The dilemma of state income apportionment. *Journal of Public Economics*, 75(1), 125-143.
- Graham, E., Shipan, C. R., & Volden, C. (2013). The diffusion of policy diffusion research in political science. *British Journal of Political Science*, 43(3), 673-701.
- Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3), 515-526.

- Grossback, L. J., Nicholson-Crotty, S., & Peterson, D. A. M. (2004). Ideology and learning in policy diffusion. *American Politics Research*, 32(5), 521-545.
- Hayes, T. J., & Dennis, C. (2014). State adoption of tax policy: New data and new insights. *American Politics Research*, advance access.
- Hearn, J. C., Lacy, T. A., & Warshaw, J. B. (forthcoming). State research and development tax credits: The historical emergence of a distinctive economic policy instrument. *Economic Development Quarterly*, advance access.
- Karch, A. (2007). *Democratic laboratories: Policy diffusion among the American states*. Ann Arbor: University of Michigan Press.
- Klarner, C. (2013). Klarner politics. Retrieved from <http://www.indstate.edu/polisci/klarnerpolitics.htm>.
- Maggetti, M., & Gilardi, F. (2013). *How policies spread: A meta-analysis of diffusion mechanisms*. Unpublished manuscript.
- Makse, T., & Volden, C. (2011). The role of policy attributes in the diffusion of innovations. *Journal of Politics*, 73(1), 108-124.
- May, P. J. (1992). Policy learning and failure. *Journal of Public Policy*, 12(4), 331-354.
- Metcalf, L. (2012). *The innovation and diffusion of state tax expenditure budgets: Examining determinants of adoption with event history analysis*. (Unpublished PhD dissertation). The George Washington University, Washington, D.C.
- Miller, C., & Richard, B. (2010). The policy diffusion of the state R&D investment tax credit. *State and Local Government Review*, 42(1), 22-35.
- Mossberger, K. (2000). *The politics of ideas and the spread of enterprise zones*. Washington, DC: Georgetown University Press.
- Neumark, D. (2013). Spurring job creation in response to severe recessions: Reconsidering hiring credits. *Journal of Policy Analysis and Management*, 32(1), 142-171.
- Neumayer, E., & Plumper, T. (2012). Conditional spatial policy dependence: Theory and model specification. *Comparative Political Studies*, 45(7), 819-849.
- Niskanen, W. A. (1971). *Bureaucracy and representative government*. Chicago: Aldine, Atherton.
- Oates, W. E. (1972). *Fiscal federalism*. New York: Harcourt Brace Jovanovich.

- Omer, T. C., & Shelley, M. K. (2004). Competitive, political, and economic factors influencing state tax policy changes. *The Journal of the American Taxation Association*, 26(Supplement), 103-126.
- Rogers, E. M. (2003). *Diffusion of innovations*. New York: Free Press.
- Rork, J. C. (2003). Coveting thy neighbor's taxation. *National Tax Journal*, 56(4), 775-788.
- Shipan, C. R., & Volden, C. (2012). Policy diffusion: Seven lessons for scholars and practitioners. *Public Administration Review*, 72(6), 788-796.
- Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, 52(4), 840-857.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129-138.
- Squire, P. (2007). Measuring state legislative professionalism: The Squire index revisited. *State Politics and Policy Quarterly*, 7(2), 211-227.
- Story, L., Fehr, T., & Watkins, D. (2012, December 1). United States of subsidies. *New York Times*. Retrieved from <http://www.nytimes.com>.
- Tiebout, C. M. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, 64(5), 416-424.
- Wilson, D. J. (2009). Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits. *Review of Economics and Statistics*, 91(2), 431-436.
- Wilson, J. D. (1999). Theories of tax competition. *National Tax Journal*, 52(2), 269-304.
- Zodrow, G. R., & Mieszkowski, P. (1986). Pigou, Tiebout, property taxation, and the underprovision of local public goods. *Journal of Urban Economics*, 19(3), 356-370.
- Zorn, C. J. W. (2000). Modeling duration dependence. *Political Analysis*, 8(4), 367-380.

Table 1: Variables and data sources 1970-2010

Variable	Description	Data Source	Hypothesized Sign of Effect
<i>Dependent Variables</i>			
<i>ITC</i>	Dummy variable indicating whether a state adopted an Investment Tax Credit in that year. New York is excluded, and the start of risk is 1970.	Data through 2006 were provided by Dan Wilson, Federal Reserve Bank of San Francisco; State legislative records (2007-2010)	
<i>Apportionment</i>	Dummy variable indicating whether a state increased its apportionment formula sales factor in that year. Iowa, Florida, Massachusetts, New York, and Wisconsin are excluded and the start of risk is 1979.	Berenthal, et al. (2012)	
<i>R&D</i>	Dummy variable indicating whether a state adopted a Research & Development tax credit in that year. The start of risk is 1982.	Wilson (2009) (1970-2006), State legislative records (2007-2010)	
<i>Job Creation Tax Credit</i>	Dummy variable indicating whether a state adopted a Job Creation Tax Credit in that year. The start of risk is 1992.	Data through 2006 were provided by Dan Wilson, Federal Reserve Bank of San Francisco; State legislative records (2007-2010)	
<i>Independent Variables</i>			
<i>Percentage of Border States Adopting</i>	Percentage of bordering states that have already adopted the tax incentive	Author calculations using dependent variable data.	?
<i>Ideological Distance</i>	$ABS([(MostRecentAdopterIdeology + AllOtherAdoptersIdeology) / 2] - PotentialAdopterIdeology)$ <p>Underlying ideology data from Berry, et.al.'s (2010) NOMINATE measure. Scores range from 0 (most conservative) to 100 (most liberal).</p>	Author calculations using method from Grossback, Nicholson-Crotty, and Peterson (2004) and data from Berry, et al. (2010)	-
<i>Democrat Governor</i>	1 = Democratic Governor 0 = Republican Governor 0.5 = Other party	Klarner (2013)	-

<i>Democratic Control of Legislature</i>	Additive scale of Democratic power in the legislature: 1 = Democratic control of both chambers 0.75 = Democratic control of one chamber, split control of the other 0.5 = Democrats control one chamber, Republicans the other 0.25 = Republican control of one chamber, split control of the other 0 = Republican control of both chambers	Klarner (2013)	-
<i>Legislative Professionalism</i>	Index of state legislative professionalism * 100. Raw scores range from 0.0 (least professional) to 1.0 (most professional). Annual scores are interpolated or extrapolated from Squire's (2007) data.	Squire (2007)	+
<i>Manufacturing Employment (ln)</i>	Logged number of paid employees in the state's manufacturing industry as of March 12 each year	County Business Patterns (Census)	+
<i>Population (ln)</i>	Logged state population	Klarner (2013)	?
<i>Unemployment Rate</i>	Annually-averaged unemployment rate for each state	Bureau of Labor Statistics (1976-2010), Manpower Report of the President (1970-75)	+
<i>Real Per Capita Income (ln)</i>	Logged real per capita income in \$1,000s, deflated with the national CPI (1982-1984\$s)	Klarner (2013)	-
<i>Corporate Income Tax Rate</i>	Top state corporate income tax rate as of January 1 of each year	Tax Foundation, Council of State Governments	+

Table 2: Descriptive Statistics

ITC	1970-Start of Risk (45 states)		2010-Adopters (24 states)		2010 Non-Adopters (21 states)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pct. Border adopters			0.542	0.313	0.559	0.345
Ideological distance			21.230	12.059	20.968	14.402
Dem. Governor	0.378	0.490	0.625	0.495	0.476	0.512
Dem. Legislature	0.506	0.466	0.688	0.438	0.583	0.421
Professionalism	19.931	10.322	18.667	9.295	19.843	15.086
Manufacturing employees (ln)	12.211	1.372	11.827	1.047	11.809	1.175
Population (ln)	14.776	1.006	15.088	0.932	15.292	1.011
Unemployment rate	4.916	1.107	8.763	1.862	8.676	2.018
Real per capita income (ln)	2.281	0.167	2.859	0.146	2.885	0.139
CIT rate	5.971	2.924	7.085	2.104	7.426	1.635

Apportionment	1979-Start of Risk (40 states)		2010-Adopters (28 states)		2010 Non-Adopters (12 states)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pct. Border adopters			0.812	0.194	0.575	0.329
Ideological distance			22.947	19.787	17.557	10.631
Dem. Governor	0.675	0.474	0.536	0.508	0.500	0.522
Dem. Legislature	0.713	0.410	0.652	0.404	0.542	0.487
Professionalism	22.465	10.725	21.150	13.705	17.092	9.854
Manufacturing employees (ln)	12.260	1.299	12.108	0.877	10.823	1.144
Population (ln)	14.795	0.957	15.382	0.873	14.404	0.766
Unemployment rate	5.590	1.311	9.050	1.861	8.008	2.042
Real per capita income (ln)	2.470	0.146	2.862	0.148	2.850	0.115
CIT rate	6.764	2.050	7.130	1.999	7.088	1.332

R&D	1981-Start of Risk (45 states)		2010-Adopters (37 states)		2010 Non-Adopters (8 states)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pct. Border adopters			0.821	0.240	0.487	0.261
Ideological distance			21.105	15.289	18.649	12.689
Dem. Governor	0.578	0.499	0.568	0.502	0.500	0.535
Dem. Legislature	0.667	0.440	0.662	0.430	0.531	0.432
Professionalism	22.120	11.377	19.665	13.229	17.138	5.508
Manufacturing employees (ln)	12.333	1.271	11.808	1.123	11.868	1.028

Population (ln)	14.916	0.968	15.147	0.980	15.350	0.932
Unemployment rate	7.362	1.807	8.576	1.999	9.400	1.365
Real per capita income (ln)	2.457	0.145	2.878	0.142	2.838	0.144
CIT rate	7.072	2.112	7.472	1.907	6.191	1.459

JCTC	1991-Start of Risk (45 states)		2010-Adopters (24 states)		2010 Non- Adopters (21 states)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pct. Border adopters			0.512	0.271	0.484	0.309
Ideological distance			19.368	12.805	21.799	15.328
Dem. Governor	0.533	0.493	0.667	0.482	0.429	0.507
Dem. Legislature	0.778	0.325	0.698	0.383	0.571	0.475
Professionalism	21.061	13.214	20.992	9.678	17.186	14.556
Manufacturing employees (ln)	12.273	1.207	11.874	1.029	11.754	1.189
Population (ln)	14.998	0.963	15.175	0.902	15.193	1.053
Unemployment rate	6.469	1.533	8.763	1.764	8.676	2.117
Real per capita income (ln)	2.619	0.151	2.859	0.133	2.884	0.153
CIT rate	7.856	2.198	7.159	1.973	7.341	1.826

Table 3: Cox Proportional Hazards Model Results

	ITC		Apportionment		R&D		JCTC	
	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error
Pct. Border adopters	1.068	0.934	2.424	1.772	6.764 **	5.177	3.698	3.995
Ideological distance	1.021	0.018	1.001	0.017	1.059 ***	0.015	1.030 **	0.014
Dem. Governor	0.639	0.309	0.605	0.236	0.479	0.239	1.140	0.475
Dem. Legislature	5.686 **	4.135	0.893	0.481	1.577	0.860	3.669 **	2.402
Professionalism	1.023 *	0.013	0.993	0.014	0.997	0.018	1.030 *	0.016
Manufacturing employees (ln)	3.560 *	2.472	1.830	1.078	1.286	0.483	2.704 *	1.488
Population (ln)	0.157 **	0.138	0.901	0.663	0.821	0.348	0.316	0.233
Unemployment rate	1.037	0.169	1.114	0.173	1.300 *	0.183	0.870	0.179
Real per capita income (ln)	7.446	10.883	0.710	1.197	1.450	2.122	2.717	5.310
CIT rate	0.932	0.089	1.106	0.112	1.391 ***	0.122	0.999	0.105
N	1294		775		787		585	
Log pseudolikelihood	-71.298		-82.094		-97.972		-77.154	
Prob > chi2	0.0689		0.0010		0.0000		0.1071	

Standard errors are clustered by state.

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 4: Non-Proportional Hazards Model Results

	ITC		Apportionment		R&D		JCTC		
	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error	
Pct. Border adopters	0.975	0.919	2.249	1.539	5.428 **	4.335	21.837	42.729	
Ideological distance	1.019	0.016	0.921 **	0.030	1.063 ***	0.015	1.028 **	0.014	
Dem. Governor	0.667	0.320	0.441 **	0.183	0.053 ***	0.047	1.176	0.502	
Dem. Legislature	5.391 **	4.175	1.064	0.640	3.276	3.992	1.590	1.699	
Professionalism	0.943 *	0.033	0.996	0.014	0.992	0.020	1.024	0.015	
Manufacturing employees (ln)	4.634 **	3.439	3.181	4.543	1.309	0.553	2.616 *	1.419	
Population (ln)	0.117 **	0.107	0.301	0.503	0.800	0.352	0.325	0.239	
Unemployment rate	0.808	0.340	1.574 *	0.403	1.619 ***	0.263	0.883	0.174	
Real per capita income (ln)	17.423 *	26.914	0.153	0.670	1.299	1.989	1.981	3.871	
CIT rate	0.833	0.219	1.532 *	0.360	2.006 ***	0.417	1.002	0.099	
Time Interactions									
Pct. Border adopters							0.831	0.146	
Ideological distance			1.007 ***	0.002					
Dem. Governor					1.151 ***	0.055			
Dem. Legislature					0.950	0.062	1.102	0.118	
Professionalism	1.004 **	0.002							
Manufacturing employees (ln)			0.998	0.086					
Population (ln)			1.040	0.106					
Unemployment rate	1.017	0.021	0.972	0.019	0.981 *	0.010			
Real per capita income (ln)			1.123	0.256					
CIT rate	1.004	0.012	0.980	0.016	0.978 *	0.012			
N	1294		775		787		585		
Log pseudolikelihood	-69.449		-76.998		-92.599		-76.457		
Prob > chi2	0.0017		0.0001		0.0000		0.0996		

Standard errors are clustered by state.

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 5: Complementary Log-Log Model Results

	ITC		Apportionment		R&D		JCTC		
	exp(b)	Std. error	exp(b)	Std. error	exp(b)	Std. error	exp(b)	Std. error	
Pct. Border adopters	0.969	0.779	1.970	1.511	10.238 **	9.724	11.119	20.080	
Ideological distance	1.021	0.021	0.928 **	0.032	1.061 ***	0.015	1.025 *	0.014	
Dem. Governor	0.620	0.319	0.473	0.220	0.082 ***	0.073	1.113	0.450	
Dem. Legislature	5.240 *	4.794	0.897	0.517	3.147	3.307	2.071	2.131	
Professionalism	0.969	0.032	0.992	0.018	0.999	0.020	1.022	0.014	
Manufacturing employees (ln)	4.003 **	2.282	3.867	5.996	1.041	0.437	2.744 *	1.690	
Population (ln)	0.151 **	0.114	0.418	0.791	1.090	0.518	0.296	0.228	
Unemployment rate	0.888	0.175	1.550 *	0.368	1.112	0.163	0.994	0.120	
Real per capita income (ln)	3.725	5.691	1.491	6.283	1.052	1.631	2.073	4.095	
CIT rate	0.982	0.132	1.646 *	0.452	1.401 ***	0.183	0.991	0.118	
Time Interactions									
Pct. Border adopters							0.835	0.125	
Ideological distance			1.005 ***	0.002					
Dem. Governor					1.143 **	0.060			
Dem. Legislature					0.964	0.055	1.038	0.078	
Professionalism	1.003 *	0.002							
Manufacturing employees (ln)			0.979	0.075					
Population (ln)			1.016	0.101					
Unemployment rate	1.003	0.008	0.983	0.018	0.999	0.006			
Real per capita income (ln)			0.964	0.257					
CIT rate	0.997	0.008	0.973	0.023	0.999	0.008			
Logged time			152.373 ***	278.586					
Constant	18.704	102.218	1.12E-09 ***	7.41E-09	3.99E-05 **	0.000	0.479	2.057	
N	1294		775		787		585		
Log pseudolikelihood	-104.388		-100.631		-120.123		-93.646		
Prob > chi2	0.0732		0.0000		0.0000		0.0554		

Standard errors are clustered by state. *** p < 0.01, ** p < 0.05, * p < 0.1

Figure 1: A Model of Policy Adoption and Diffusion

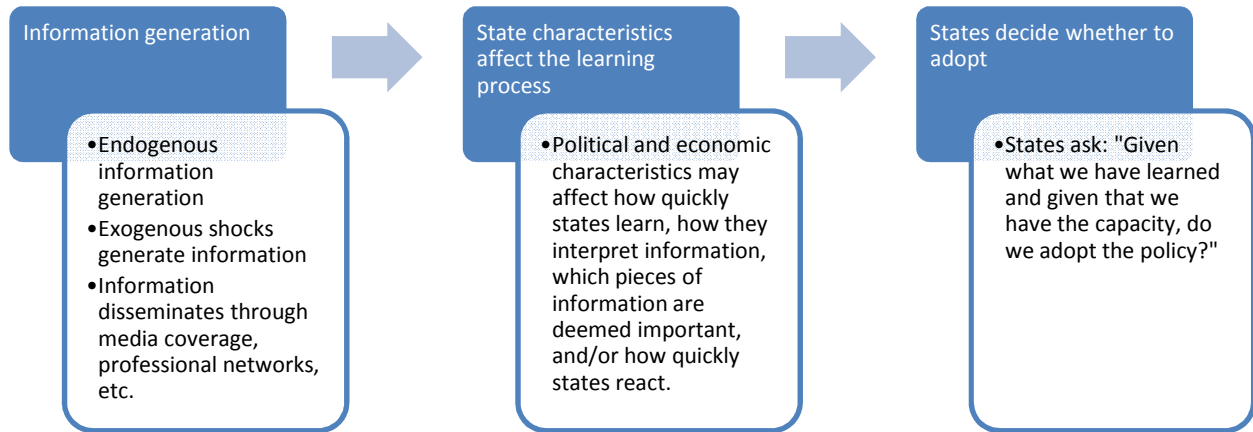


Figure 2: Kaplan-Meier Failure Plots

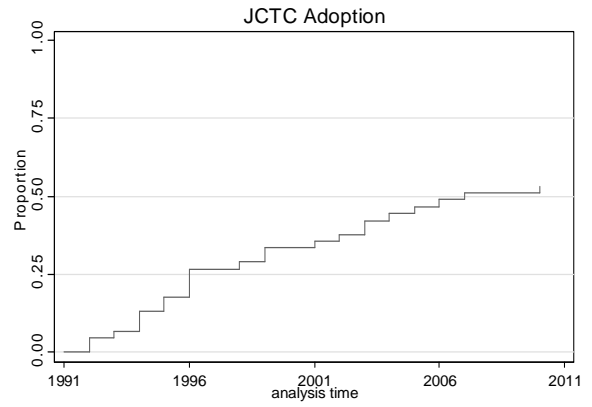
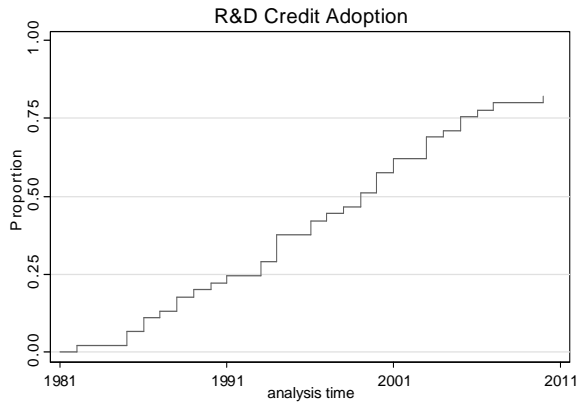
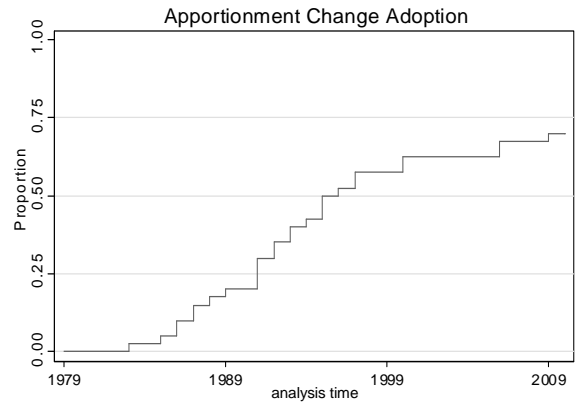
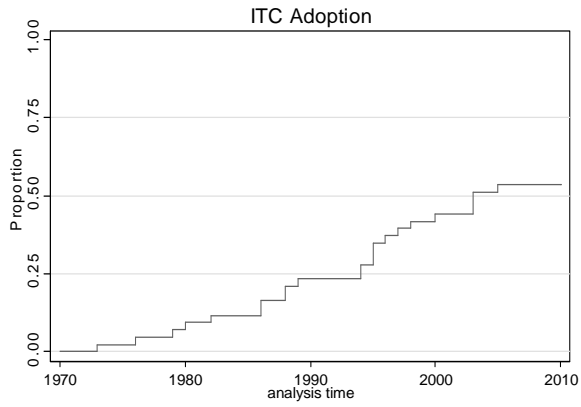
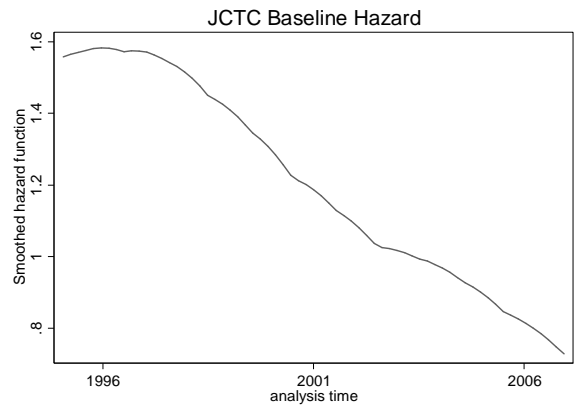
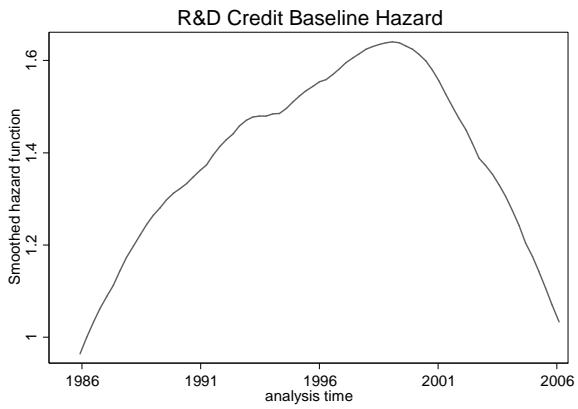
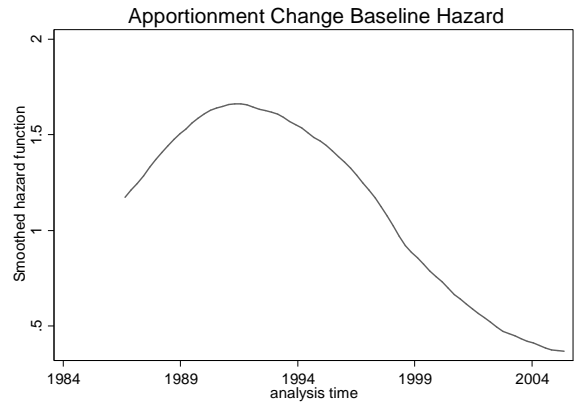
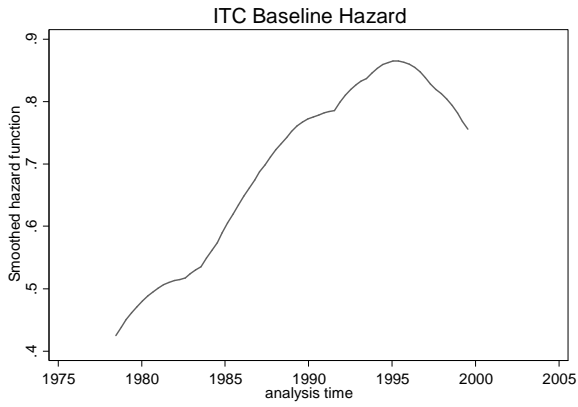


Figure 3: Baseline Hazard Plots



Essay Three: Duration Dependence in Policy Diffusion Studies

Event history analysis is concerned with the timing of events and is well-suited for addressing the types of questions policy diffusion scholars ask, such as: which factors increase the likelihood that a government will adopt a policy at any given time? The likelihood of adoption at a given time t is expressed by the hazard rate, $h(t)$. In the policy adoption context, a rising hazard rate indicates that the probability that governments will adopt a policy (given that they haven't already) is increasing over time, and a falling hazard indicates that adoption is becoming less likely over time. Duration dependence is the term used to describe the overall shape or pattern of how the hazard changes over time. For example, positive duration dependence indicates a period of rising hazards, while negative duration dependence indicates a period of falling hazards. Researchers' expectations about how the hazard changes over time guide the choice of how to model duration dependence, a key component to any event history model. However, the policy diffusion literature has yet to make full use of all of the tools event history analysis has to offer for modeling and interpreting duration dependence. In fact, some system-level diffusion processes, such as bandwagon and critical mass effects, will likely only be detected in duration dependence. This essay fills this gap in the policy diffusion literature by explaining how policy diffusion scholars can use event history techniques to properly model and interpret duration dependence in policy diffusion studies.

Beginning with Berry and Berry's (1990) introduction of event history analysis as a methodology for policy diffusion studies, it has become the technique of choice for scholars analyzing quantitative policy diffusion data. In the past 20+ years, policy diffusion scholars have deployed event history analysis in the study of how policies spread at the local, state/sub-national, federal/national, and international levels (for a review, see F. S. Berry & Berry, 2007). Unlike earlier methods used to study policy diffusion, event history methods are similar to the

conventional regression paradigm, with which social scientists are quite familiar. In the policy diffusion context, the unit of analysis is a government (e.g. states, countries, etc.), and event history methods are used to study the effects of independent variables on the timing of governments' adoption of certain policies, measured as the duration of time from a chosen start of risk to the adoption. Policy diffusion scholars have made several refinements and advances in using event history methods, including exploring ways to incorporate spatial effects (W. D. Berry & Baybeck, 2005), ways to model diffusion mechanisms (Shipan & Volden, 2008), using dyads as the unit of analysis (Volden, 2006), and using pooled analyses of multiple policies (Boehmke & Skinner, 2012).

Here, however, I argue that policy diffusion scholars have overlooked one of the central goals and strengths of event history analysis—the ability to model duration dependence in the hazard of failure. Many event history studies in other areas of research such as biostatistics, epidemiology, and demography *primarily* focus on modeling and interpreting how the hazard rate changes over time. For example, demographers differentiate a “pure” person-to-person infection process from a “contact-independent” infection process (e.g. influence by mass media) by analyzing how the hazard changes over time while controlling for a set of covariates (Braun & Engelhardt, 2002). In contrast, policy diffusion researchers have generally viewed duration dependence as something to be controlled for rather than interpreted and analyzed. They have accounted for duration dependence by following the advice of political science methodology papers like Beck, Katz, and Tucker (1998) without much consideration as to whether duration dependence has any meaning or importance in the context of policy diffusion. In international relations, another area of political science where event history methods are commonly used, some scholars have argued that duration dependence is of substantive importance and ought to be

modeled and interpreted (Bennett, 1999; Carter & Signorino, 2010; Zorn, 2000), but this line of argument has yet to be taken up in the policy diffusion literature.

In this essay, I discuss what duration dependence could mean in the context of policy diffusion theory and suggest that one source of observed duration dependence is system-level diffusion dynamics, such as bandwagon effects, critical mass, competitive races, policy “outbreaks” (Boushey, 2010; 2012), and generalized learning effects. Central to my argument is the observation that these types of effects cannot be adequately modeled with covariates and will instead manifest as duration dependence. While policy diffusion scholars have primarily modeled diffusion mechanisms using covariates, I argue that their approach can be complemented by incorporating techniques to model duration dependence for a more comprehensive examination of diffusion processes and patterns.

The next section discusses the policy diffusion literature with regard to how mechanisms of diffusion are defined and operationalized in quantitative studies. The following section provides more detail about event history analysis and discusses the meaning of duration dependence from a methodological perspective. The fourth expands on the discussion of duration dependence to explore what it may mean in the context of policy diffusion. The fifth section provides a list of recommendations for how policy diffusion scholars can incorporate the analysis of duration dependence into their studies, and the sixth section concludes.

Defining and Operationalizing Policy Diffusion Mechanisms

The policy diffusion literature studies how and why policy ideas spread across jurisdictional lines, often to become adopted by multiple governments in a population. Elkins and Simmons (2005) argue that diffusion encompasses any process of *uncoordinated*

interdependence, including learning, imitation, bandwagoning, emulation, mimicry, or competition. Boushey (2010) likens policy diffusion to the spread of diseases in the study of epidemiology: a small number of “infected” individuals transmit a “disease” to non-infected individuals, causing the total proportion of infected individuals in the population to rise.

As Elkins and Simmons (2005) argue, it is important to distinguish between defining diffusion in terms widespread adoption of a policy—*diffusion-as-outcome*—and defining diffusion as a process of interdependence—*diffusion-as-process*. The *diffusion-as-outcome* conceptualization, based on the observed prevalence of a policy in many governments, implies that prior adoptions somehow increase the probability of adoption by non-adopters. However, the *diffusion-as-process* conceptualization is more general because it does not specify how prior adoptions affect non-adopters, allowing for the possibility that prior adoptions could actually decrease the probability of adoption by non-adopters. For example, Brueckner (2003) shows that in multiple models of strategic interaction among governments, including tax competition, it is theoretically ambiguous whether governments will tend to converge on the same policies or diverge from one another. Following the advice of Elkins and Simmons (2005), this essay conceives of diffusion as a process of interdependence without making the assumption that the process leads to widespread adoption. Like Elkins and Simmons, I adopt Strang’s (1991) definition of diffusion as “any process where prior adoption of a trait or practice in a population alters the probability of adoption for remaining non-adopters.”

Mechanisms of diffusion provide the causal stories of how exactly interactions between prior and potential adopters alter the probability of adoption. Shipan and Volden (2008) describe four mechanisms:

- *Learning*: Governments consider the experiences of prior adopters in deciding whether to adopt a policy. The deemed success of policies in other jurisdictions increases the probability of adoption.
- *Economic competition*: Some policies may create positive or negative spillovers in other jurisdictions. Governments may strategically adopt or not adopt policies in order to gain an advantage or avoid being disadvantaged as a result of spillovers.
- *Imitation*: Governments often copy the actions of certain prior adopters in order to appear more like them. No learning about the policy itself is necessary.
- *Coercion*: Governments often directly or indirectly exert pressure on one another either to adopt or not adopt certain policies.

It would be very difficult, if not impossible, to directly observe any of these processes because it would require full information about the internal motivations of individual policymakers and other actors as well as the dynamics that aggregate individual actions into decisions by governments.

Instead, policy diffusion scholars have posed the following question: if a non-adopting government were learning/competing/etc., which adopting governments would influence it? For example, researchers often argue that governments are most likely to learn from their geographic neighbors. These types of studies usually include a variable that measures the number or proportion of a government's neighbors that are prior adopters and hypothesize that the more its neighbors are adopters, the more likely the government is also likely to adopt (F. S. Berry & Berry, 1990; W. D. Berry & Baybeck, 2005; F. S. Berry & Berry, 1992). For the mechanism of competition, Berry and Baybeck (2005, p. 505) argue that "states' influences on each other

should vary depending on the size and location of specific populations of individuals and firms within the states.” A current line of research in policy diffusion proposes that multiple mechanisms of diffusion may operate simultaneously (Baybeck, Berry, & Siegel, 2011; W. D. Berry & Baybeck, 2005; Shipan & Volden, 2008). For example, Shipan and Volden (2008) test for the presence of learning, competition, imitation, and coercion in the adoption of antismoking policies in US cities. They argue that the following network relations correspond to each mechanism: cities *learn* from other cities in the same state; cities *compete* with other cities within 10 miles, especially if they have large populations; cities *imitate* the nearest city with a larger population; and cities are *coerced* by their own state governments. Therefore, each mechanism of diffusion is associated with a different network of relationships among governments.

In other words, “the analyst defines social relations which link actors to each other. Such relations comprise a social structure that ‘channels’ diffusion, so actors respond to those they are socially connected to” (Strang, 1991, p. 325-6). Strang refers to this idea of diffusion channeled by network relations as “heterogeneous mixing” and contrasts it with homogeneous mixing, where all prior adopters have roughly the same ability to influence non-adopters.

To operationalize these diffusion mechanisms in quantitative studies, researchers create spatial lag or similar types of variables with the goal of capturing avenues of learning, competition, imitation, or coercion. For a given diffusion mechanism, the first step is to define the relevant network spatial feature, which can be based on measures of geography, political ideology, social characteristics, etc. The next step is to use the network to calculate a connectivity or distance matrix for all governments in the network.²³ For example, if a

²³ Technically, due to the structure of event history data, the spatial feature that defines the network should be something that does not change over time. If the spatial feature changes annually, for example, the connectivity

researcher believes competitive pressures are proportional to geographic distances between population centers, then these distances would constitute the connectivity matrix. The matrix values are then used to weight certain pieces of information about the policy choices of other governments. Usually, the information is simply a binary indicator of whether the government has already adopted the policy, but could also include other relevant information like whether policy was successful. The final step is to combine the distance/weights matrix with the information about prior adopters to create a spatial lag variable. The value of the spatial lag variable represents a kind of weighted average of the policy choices of other governments (see Maggetti, Radaelli, and Gilardi (2013) for a detailed discussion and example). Therefore, increasing values of the spatial lag variable indicate that more “closely-related” governments (based on the chosen connectivity matrix) have adopted the policy, and one may hypothesize that this leads to greater hazard of adoption. These spatial lag variables, with different weighting matrices for different diffusion mechanisms, are then included in event history regressions to estimate their respective effects on the hazard of adopting a policy.

Event History Analysis and Duration Dependence

Event history analysis has become the method of choice for large- N quantitative studies of policy diffusion (F. S. Berry & Berry, 2007). Sometimes also called duration analysis or survival analysis, event history analysis was originally developed in the context of biostatistics and used to study lifespans. More generally, it can be used in quantitative research designs for which the dependent variable is the duration of time from a specified “start of risk” until the

matrix would also need to be measured annually. However, a government that adopts early on in the study would subsequently have missing data for the spatial feature. Moreover, consistent with the logic of Grossback, Nicholson-Crotty, and Peterson’s (2004) ideological distance variable, governments should arguably be most influenced by the characteristics of adopters *at the time of adoption*, not in some subsequent time period.

occurrence of some “event,” whether the event is the death of an organism, the contraction of a virus, or a government’s adoption of a policy. At the onset of the study, individuals are assumed to be “at risk” of experiencing some event. When an individual experiences an event, the individual is considered to have “failed,” and that individual is then dropped from further observation.²⁴ Individuals that have not experienced the event by the end of the study are said to be right-censored.

Underlying all event history models are three basic functions: the probability density function, which shows the probability of failure as a function of time; the survival function, which shows the proportion of individuals that have not experienced the event at each time; and the hazard function, a ratio of the probability density function and survival function that reflects the instantaneous probability of failure, given that the individual has survived to that particular time. Most event history models in the social sciences specify the hazard as a function of covariate values. As Box-Steffensmeier and Jones (2004) explain, social scientists familiar with the regression framework are often most interested in hypotheses about the effects of certain covariates on the hazard of failure. For example, as mentioned above, policy diffusion scholars often hypothesize that governments with more neighbors that have already adopted a policy have a higher risk of also adopting the policy.

Despite the attention paid to modeling the effects of covariates, a key element of event history analysis also requires researchers to develop a strategy for specifying the shape of the hazard function. The hazard model selected represents duration dependence, the effect of time on the hazard of failure. In models that have covariates, the “baseline” hazard can be thought of as the effect of time, controlling for the effects of covariates in the model. For example, Figure 1

²⁴ There are also event history models designed for repeating events data, such as spells of unemployment. See Box-Steffensmeier and Jones (2004) for a detailed discussion.

replicates Figure 3 from the first essay, which shows the baseline hazard for U.S. states' adoptions of incentives for film and video production. This figure shows how the hazard of adoption increases over time, especially after 2003, while controlling for certain political and economic characteristics of states, as well as other covariates. This means that some factor unrelated to the covariates in the model is causing the hazard to rise over time, manifesting as duration dependence.

A variety of parametric and non-parametric methods are available for specifying the hazard function. Common parametric models use the exponential, Weibull, Gompertz, log-normal, and generalized gamma distributions, each of which require different assumptions about the shape of the hazard (see Box-Steffensmeier and Jones (2004) for a detailed discussion). These parametric models treat time as continuous, not broken into discrete periods like days or years, as it is in the data most social scientists are familiar with.

Discrete time event history models are designed for data that break time into discrete spans, similar to time series or panel data. In discrete time event history data, each row of data is indexed by an individual and a period of time (e.g. state-year or person-month), and the dependent variable becomes an indicator of whether that particular individual failed during that particular period of time. Beck, Katz, and Tucker (1998) call this type of data “time-series—cross-section data with a binary dependent variable” and explain that it is identical to duration data that have been grouped or aggregated into discrete time intervals. Because the dependent variable is binary, researchers can use one of the familiar binary data models, logit or probit, or the less familiar but arguably more appropriate complementary log-log model.²⁵

²⁵ Buckley and Westerland (2004) argue that although the complementary log-log model is less common in the social sciences, it may actually be more theoretically appropriate for event history data. Unlike the logit or probit functions, the cloglog distribution is skewed, and may better account for the fact that observations of events ($y = 1$) are relatively rare compared to observations of non-events ($y = 0$).

Unfortunately, the familiarity of the logit and probit models seems to have caused researchers to treat event history analysis just like other familiar regression methods, de-emphasizing the importance of modeling duration dependence. Unlike parametric event history models where the choice of function for duration dependence is apparent and unavoidable, choices about duration dependence in discrete time models are less straightforward. In fact, many of the early studies of policy diffusion that used event history analysis did not account for duration dependence at all, effectively constraining their models to have a constant baseline hazard (F. S. Berry & Berry, 2007). Researchers subsequently pointed out that event history data are usually temporally dependent, violating the assumption of independent observations in ordinary logit and probit models, and leading to incorrect standard errors and inefficiency in estimation (Beck et al., 1998). A variety of specifications for duration dependence have been proposed to account for this temporal dependence, including time dummies, log-transformed duration time, time polynomials, and splines (Beck et al., 1998; Box-Steffensmeier & Jones, 2004; Carter & Signorino, 2010).

Another option is for researchers to avoid making assumptions about duration dependence altogether by using the Cox proportional hazards model. In the Cox model, duration dependence is not parameterized, and the model has no intercept. In fact, with discrete time data, the Cox model is actually equivalent to the conditional logit model used for matched case-control data (Box-Steffensmeier & Jones, 2004). For a conditional logit model, event history data is grouped into “risk periods” by splitting the data at failure times. During each risk period, some individuals fail (“cases”) and some do not (“controls”). The conditional logit model estimates the likelihood that particular individuals fail, given all the individuals in the risk set during the particular risk period. In this way, the conditional logit model has fixed effects by risk period,

and the effect of time is not explicitly modeled. For the Cox/conditional logit model, a non-parametric estimate of the baseline hazard can be retrieved, but it tends to be noisy and overfitted (Box-Steffensmeier & Jones, 2004). Moreover, while non-parametric plots can be useful for visual inspection of duration dependence in the Cox model, they cannot be used for statistical inference.

As Carter and Signorino (2010, p.272) have noted, “the vast majority of researchers have treated temporal dependence in binary data models more as a statistical nuisance that needs to be ‘controlled for,’ rather than as something that is substantively interesting.” They report that of the 119 studies (as of July 2006) that use Beck, Katz, and Tucker’s (1998) time dummies or spline specification for duration dependence, only 7 actually plot and interpret the hazard. The effect of time is relegated to a methodological issue as researchers have focused on testing hypotheses related to other covariates in their models.

The lack of attention to duration dependence is regrettable because it can often have an important substantive interpretation for political processes (Bennett, 1999). It is important to remember that when covariates are included in a model, the duration dependence or baseline hazard traces out how the hazard of failure changes over time, *controlling for the effects of other covariates*.²⁶ If the hazard is increasing (decreasing), it is referred to as positive (negative) duration dependence and implies that the hazard of failure is increasing (decreasing) for some reason unrelated to the covariates in the model. Of course, it is not simply the passage of time itself that causes the hazard to rise or fall, but the passage of time might be related to other unmodeled processes or factors that have a temporal component. In other words, time is simply a

²⁶ The baseline hazard is usually plotted assuming covariates are set to zero. However, if there is no natural zero point for a covariate, it can be first mean-centered and then set to zero (Box-Steffensmeier & Jones, 2004). However, with proportional hazards models, the values of covariates don’t really matter that much because the shape of the hazard will not change.

proxy. To be more specific, the following sections describe several possible explanations for observed duration dependence.

Omitted Variables and State Dependence

One possible reason for observed duration dependence is omitted variables. The form of duration dependence has proven to be quite sensitive to the inclusion or exclusion of covariates, especially if the covariates are correlated with time (Bennett, 1999). If an important variable is omitted from a model, its effect will be absorbed into the baseline hazard, possibly resulting in duration dependence. A related idea is the concept of “state dependence,” which arises when the “state” or non-failure of an individual is somehow self-perpetuating (positive state dependence) or self-dampening (negative state dependence) over time (Bennett, 1999; Zorn, 2000). Positive state dependence is associated with a falling hazard of failure (i.e. negative duration dependence), and negative state dependence is associated with a rising hazard of failure (i.e. positive duration dependence). For example, Zorn (2000) argues that the termination of international alliances should exhibit positive state dependence (and negative duration dependence) because alliances become institutionalized and therefore more durable over time. In theory, including a variable that could measure institutionalization would account for the positive state dependence and eliminate any observed negative duration dependence.

In terms of policy diffusion, positive state dependence could imply that the longer a government goes without adopting a policy, the less likely it is to actually do so. This type of pattern may appear in any policy area where new ideas are rapidly being developed to replace old ones. Over time, governments would be less likely to adopt one particular policy—

exhibiting negative duration dependence—simply because they can now adopt a newer and better idea. In this case, the omitted variable is the presence of new policy alternatives.

Unobserved Heterogeneity

Closely related to the issue of omitted variables, unobserved heterogeneity is a source of model mis-specification that can create “spurious” duration dependence. Unobserved heterogeneity biases the observed duration dependence downward and biases parameter estimates, even if it is not correlated with the other covariates included in the model (Zorn, 2000). In the context of policy diffusion, unobserved heterogeneity represents unmeasured factors that make certain governments quicker to adopt new policies, perhaps a kind of inherent innovativeness or appetite for risk (Boehmke & Skinner, 2012). Similarly, as discussed by Ingle, Cohen-Vogel, and Hughes (2007) in their qualitative study of adoptions and non-adoptions of merit aid programs in southeastern U.S. states, sometimes certain states respond in idiosyncratic ways to diffusion-related stimuli like neighbor adoption, resulting in some states being more resistant to adoption than others. In fact, the concepts of “innovativeness” and “resistance” may be simply inverse ways of expressing the same heterogeneity.

Non-proportional Hazards

Another source of observed duration dependence stems from violations of the proportional hazards property, a key assumption in most commonly used event history models (Box-Steffensmeier & Jones, 2004). The proportional hazards property essentially requires that the effects of covariates are constant over time. In other words, changes in covariate values shift the baseline hazard up or down in a parallel fashion, but its shape does not change. However, as

Box-Steffensmeier and Zorn (2001) argue, there are reasons to believe that violations of proportional hazards are quite common in political science event history data, leading to biased estimates and flawed inference. Moreover, because diffusion is inherently a dynamic process that changes over time, factors that motivate early adopters are not necessarily the same as the factors that motivate later adopters (Rogers, 2003). Therefore, it is reasonable to expect that theoretically-derived hypotheses may imply time-varying effects for diffusion-related covariates. Shipan and Volden review several policy diffusion studies where the effects of diffusion change over time. For example, Gilardi, Füglistler, and Luyet (2009) find that in the diffusion of hospital financing reforms in OECD countries, the effect of their learning variable increases over time as the weight of accumulated evidence grows.

Estimating a proportional hazards model when hazards are actually non-proportional will result in biased coefficient estimates and incorrect standard errors (Box-Steffensmeier & Zorn, 2001; Box-Steffensmeier & Jones, 2004). In addition, variables with statistically significant non-proportional effects may actually be statistically insignificant if incorrectly modeled as proportional (Box-Steffensmeier, Reiter, & Zorn, 2003). The presence of non-proportional hazards will also affect the estimation and interpretation of the baseline hazard. As I discuss in more detail below, the proper way to model non-proportional effect is to include a term that interacts the offending covariates with a function of time (Box-Steffensmeier & Jones, 2004). Therefore, in a non-proportional hazards model with time interactions, the baseline hazard no longer has the same interpretation as the effect of time controlling for covariates. Because of the interaction term(s), the observed duration dependence will vary depending on the values of the covariates with which it is interacted. Therefore, the effect of time is inseparable from the effects of covariates with non-proportional hazards.

Exogenous Factors

Another possible explanation for changes in the observed hazard rate may be exogenous factors that affect all individuals in roughly the same way. In studies of policy diffusion in U.S. states, macroeconomic conditions or changes in federal policy may affect the hazard rate for all states. For example, an increase in federal corporate income tax rates would likely encourage businesses to pressure all states to reduce their corporate income tax rates. With exogenous pressures, time is just a proxy for unobserved causal factors that vary across time but not across individuals.

System-level Diffusion Dynamics

Similar to exogenous pressures, some types of endogenous diffusion influences vary across time but not across individuals. For example, consider the possibility that individuals are influenced by the total number of adopters, regardless of any specific network relationships. Strang (1991) calls this a homogeneous mixing process where any prior adopter has the same level of influence over any non-adopter. In other words, what matters is not *which* individuals have adopted but *how many*. Because such an effect exhibits variation across time but not across individuals in the system, it will be absorbed into the baseline hazard (i.e. the effect of time) and manifest as duration dependence. The next section addresses how homogeneous mixing and system-level diffusion dynamics relate to policy diffusion.

System-level Dynamics in Policy Diffusion

As discussed above, diffusion effects are usually modeled using spatial lag or similarly-constructed variables, and mechanisms are differentiated based on how the relevant network of governments is defined. Because each government is not equally “close” to all others in terms of physical distance, ideological distance, etc., some governments exert more influence than others on any given government, resulting in what Strang (1991) has called a heterogeneous mixing process. However, policy diffusion researchers have paid little attention to the possibility that diffusion can occur through a homogeneous mixing process in which governments have roughly equal influence over one another. The following are some examples of diffusion dynamics that would only be revealed by examining duration dependence:

Bandwagon Effects

What if governments were influenced merely by the total number of previous adopters, regardless of their identities? As more and more governments adopt a policy, remaining non-adopters will feel increased pressure to adopt, a kind of bandwagon effect. Bandwagon policy diffusion effects have been posited as a result of positive feedback cycles in the punctuated equilibrium framework and as a result of the representativeness heuristic in the bounded rationality framework (Boushey, 2010; Weyland, 2006). In the first essay of this dissertation, I find evidence of a bandwagon effect in the adoption of state film incentives. A bandwagon effect can be explained either in terms of learning or competition. The fact that several other governments have adopted a policy may be seen by non-adopters as a signal that the policy is good. In this case, non-adopters could be following a heuristic that they will adopt (or at least consider adopting) policies that are being adopted by a growing number of their peers. Alternatively, governments may interpret increased adoptions among other governments as a

competitive signal: whether the policy is good or bad, it is becoming more popular, and non-adopting governments should consider adopting the policy to keep up with their peers. As displayed in Figure 2, in a bandwagon effect, the hazard of adoption would be monotonically increasing, and likely accelerating.

Threshold Effects and Critical Mass

Similar to bandwagon effects are the concepts of threshold effects and critical mass. According to Rogers (2003, p. 355), “A *threshold* [emphasis in original] is the number of other individuals who must be engaged in an activity before a given individual will join that activity.” The logic is similar to the bandwagon concept described above except that instead of each additional adopter adding pressure to non-adopters, non-adopters feel increased pressure only after a minimum number of adoptions. Rogers explains that individual decisions—which are subject to threshold effects—are aggregated to define the critical mass at the system level. The critical mass is “the point at which enough individuals in a system have adopted an innovation so that the innovation’s further rate of adoption becomes self-sustaining” (Rogers, 2003, p. 343). In other words, at the critical mass, an additional adoption pushes the total number of adopters past another non-adopter’s threshold, triggering *its* adoption, which again increases the total number of adopters, etc. With a critical mass effect, the hazard will probably be relatively flat until the time the critical mass is reached, after which it will be monotonically increasing (Figure 2).

Competitive Races

As Franzese and Hays (2008) explain, some policies are associated with negative externalities—the adoption of a policy in one jurisdiction somehow negatively affects other

jurisdictions. The classic example is tax competition. A tax cut in one jurisdiction will presumably cause resources (capital, labor, etc.) to flow away from other jurisdictions into that jurisdiction, pressuring the other jurisdictions to also adopt tax cuts. Franzese and Hays explain that these types of situations are characterized by first mover advantages. That is, the greatest benefits of a tax cut (in terms of resource inflow) will accrue to those who cut their taxes before others, and as more and more jurisdictions cut their taxes, the benefit dissipates.

Policy diffusion scholars have argued that these types of competitive pressures will usually be felt most acutely by geographic neighbors (Baybeck et al., 2011; W. D. Berry & Baybeck, 2005). However, in some cases it may be reasonable for governments to believe that they are in relatively equal competition with *all* other governments, not just their geographic neighbors. This will be especially true when individual governments in a population are quite similar to one another, as is the case with states in the US, or other subnational governments. It will also be more common when the competition is for resources that are highly mobile, such as investment in motion pictures, which I discuss in the first essay of this dissertation.

A competitive race would be characterized by an inverted U-shaped hazard (Figure 2). At first, as governments start to become aware of a competitive policy, they would try to move quickly to dissipate the first mover advantage, and the hazard would increase. Then, as more governments adopt the policy, the potential gains would be dissipated, reducing the urgency to adopt and causing the hazard to level off or even fall.

Policy Outbreaks

Boushey (2010; 2012) describes another system-level policy diffusion dynamic driven by external events. He draws on punctuated equilibrium theory (True, Jones, & Baumgartner, 2007)

to suggest that exogenous shocks, such as a federal mandate or focusing event, can trigger policy outbreaks, sudden and rapid increases in the rate of adoption of a policy innovation. Boushey (2012) shows that state adoptions of motorcycle helmet laws between 1967 and 1985 are an example of a policy outbreak that was triggered by direct federal intervention. As Boushey explains, where the influence of exogenous events is strong, the cumulative distribution function (which plots the proportion of the population that are adopters over time) tends to be shaped like a lower-case r, rising very steeply when the exogenous event occurs and then flattening over time. The corresponding hazard function, therefore, would sharply increase when the exogenous event occurs and steadily decrease afterwards (Figure 2).

Generalized Learning

In the absence of exogenous shocks, Boushey (2012) explains that policies can spread as a result of a decentralized process of incremental learning. While policy diffusion scholars have concentrated on learning that takes place through specific network connections (e.g. learning from geographic neighbors), governments may also learn from each other in a more generalized way. With modern communication technology, it is relatively easy for policymakers to learn about what is going on in other governments. At a minimum, information about the total number of adopters of a policy should be widely available, and this information is sufficient to trigger bandwagon, critical mass, or race to the bottom effects, as described above. It is also quite likely that more detailed information about a policy—identities of adopters, reports of success/failure, mass media coverage, etc.—will be widely available to all governments in a system.

These system-wide learning opportunities do not preclude more intensive forms of learning from happening through specialized network connections, such as membership in

interstate professional organizations (Balla, 2001). However, to the extent that certain information is equally available to all individuals in a system, its effect on the rate of adoption will surface in the baseline hazard. In this way, the baseline hazard reflects the overall popularity of a policy, controlling for the effects of any covariates that reflect individual differences. An increasing hazard indicates that the policy is becoming more popular, perhaps because of positive media coverage or widely publicized reports of success. A decreasing hazard indicates that the policy is becoming less popular over time, perhaps because of implementation challenges, negative re-framing, or general waning interest.

The key point about bandwagon effects, critical mass, competitive races, policy outbreaks, and generalized learning is that they are all dynamics that occur at the system level, not at the individual level. However, quantitative studies of policy diffusion frequently utilize a regression framework that is designed to analyze variation at the individual level.²⁷ The problem with using only covariates to model diffusion is that it limits diffusion processes to being measured by individual-level (or dyad-level) characteristics. As a result, system-level dynamics that affect all individuals/governments in approximately the same way are simply absorbed into the baseline hazard and largely ignored.

Modeling Strategies and Recommendations

Given that duration dependence in event history policy diffusion studies is potentially much more than an assumptions violation or something to be controlled for, researchers should

²⁷ A recent notable example is Boushey (2010; 2012), who suggests that a “standard” policy diffusion process characterized by incremental learning should resemble a normal cumulative distribution function. He uses a Bass mixed influence model to explore how the diffusion curves for various policies diverge from normal. However, this model is not designed to examine how the characteristics of individuals affect the hazard of adoption, as is the case for most quantitative event history diffusion studies.

work to develop strategies to properly model, analyze, and interpret the baseline hazard. Below is a list of steps and recommendations policy diffusion researchers can follow when using quantitative event history methods.

1. Develop hypotheses and modeling strategies for duration dependence.

Researchers should use existing theory as well as background knowledge about specific policy contexts to develop hypotheses about temporal dynamics, including duration dependence and/or non-proportional hazards. As described above, some diffusion processes that involve the transmission of information and influence through specified network channels can be modeled with covariates. However, other processes like bandwagon effects and competitive races will appear in the duration dependence.

In some cases, it may be appropriate and feasible to actually include in a regression a variable that measures the total number of previous adopters to capture a bandwagon or similar effect. In parametric and discrete time event history models, this kind of variable can mathematically be included, but should be used with caution. System-level variables that vary across time but not across individuals are likely to be highly collinear with the specification for duration dependence, possibly resulting in inflated standard errors (Beck et al., 1998). As for the Cox model, it helps to recall that the Cox model is equivalent to a conditional logit model where the data are grouped by risk periods—that is, by time. Like fixed effects models, statistical identification for Cox model estimates comes from within-group variation, so covariates that are too closely related to time—as may be the case for the total number of previous adopters—will likely prevent the partial likelihood from converging.

In most cases, time will have to serve as a proxy for system-level diffusion dynamics. Therefore, researchers could also test hypotheses about whether system-level diffusion dynamics are conditioned by certain factors by interacting time with covariates. For example, one hypothesis could be that states with more professionalized legislatures (which have more resources, larger staffs, more days in session, etc.) are less susceptible to bandwagon effects—that is, the coefficient of an interaction term should be negative, with the primary effects of time and professionalism being positive. In this way, non-proportional hazards are re-interpreted to illustrate what Shipan and Volden (2008) call the conditional nature of diffusion.

2. Perform a preliminary analysis with a Cox model and plot the baseline hazard.

Although the Cox model does not parameterize the baseline hazard, a non-parametric baseline hazard plot can be retrieved from Cox model estimates (Box-Steffensmeier & Jones, 2004). The Cox baseline hazard plot is more flexible than other models because it does not impose any particular shape on duration dependence, and estimating a Cox model is a useful way to get an initial idea of the overall shape of the baseline hazard. Visual inspection should involve looking for major trends of positive or negative duration dependence or discontinuities that could indicate the influence of exogenous shocks. Researchers should also compare the Cox baseline hazard plots to the general shapes of system-level diffusion dynamics shown in Figure 2. This preliminary inspection of the baseline hazard will ideally provide guidance in selecting a parametric specification for duration dependence, which can then be used for statistical inference and tests of duration dependence (see Step 5 below).

3. Test for non-proportional hazards.

Non-proportional hazards indicate that the effect of a covariate changes over time and may interfere with the estimation and interpretation of duration dependence. There are several residual-based statistical tests available to assess non-proportional hazards in the Cox model, both at the covariate level and the global/model level, and many of them have been built into standard statistical packages (Box-Steffensmeier & Zorn, 2001; Box-Steffensmeier & Jones, 2004; Grambsch & Therneau, 1994). However, tests for assessing proportionality in models where time is parameterized are not as well developed and involve a degree of guesswork by the researcher (Box-Steffensmeier & Zorn, 2001).²⁸ Therefore, the availability of precise tests for non-proportionality is another argument in favor of using a Cox model as a preliminary step in the analysis.

When tests show that the proportional hazards assumption is violated for particular covariates, non-proportional hazards can be modeled by including time by covariate interaction terms in the model. The particular functional form of time is the choice of the researcher, though it is most common to use the natural log.²⁹ A likelihood ratio test comparing the original and expanded models should show that the inclusion of time interactions improves the fit of the model. Note that this does not necessarily ensure that any of the interaction term coefficients are statistically significant, which may have more to do with statistical power and sample size.

4. Test for unobserved heterogeneity.

²⁸ There are two general approaches to testing proportionality in parametric models. The first method involves dividing up time into separate periods (at the choice of the researcher), running separate regressions, and comparing coefficient values. Major differences may indicate non-proportional hazards. The second method involves stratifying the data by a covariate of interest, estimating separate regressions, and comparing baseline hazards. For more detail, see Box-Steffensmeier and Zorn (2001).

²⁹ In discussing the Cox model, Box-Steffensmeier and Zorn (2001) provide some discussion about whether to interact variables with time, time², or ln(time), which is most common, but there do not appear to be any strong arguments favoring one way or another. In their discussion of using polynomials and splines to model duration dependence in discrete time models, Carter and Signorino (2010) recommend interacting covariates with each of the spline or polynomial terms.

It is also important to try to rule out unobserved heterogeneity, which can cause “spurious” negative duration dependence even when the unobserved heterogeneity is not correlated with the observed covariates (Zorn, 2000). In event history analysis, shared frailty models account for unobserved heterogeneity by incorporating a random frailty parameter into the usual models (Box-Steffensmeier & Jones, 2004).³⁰ The use of a frailty model requires the researcher to make an assumption about the distribution of the random effect, and most commonly, a gamma distribution is assumed (Zorn, 2000). In the discrete time framework, one can simply use a random effects logit, probit, or cloglog model to account for heterogeneity. Alternatively, Darmofal (2009) proposes a Bayesian approach that allows for spatially auto-correlated random effects in frailty models. In any case, testing for unobserved heterogeneity involves testing the hypothesis that the random effects variance parameter is equal to zero. If a random effects model is used, it is important to remember that the model no longer has the proportional hazards property, and the shape of the hazard will vary from individual to individual.

5. Find the best-fitting specification for duration dependence.

Once other sources of duration dependence have been ruled out or accounted for, the final step is to create a model in which the form of duration dependence is fully specified. The general approach to specifying duration dependence has been to prioritize flexibility rather than parsimony or interpretability. For example, one of the most widely-cited sources for advice on specifying duration dependence is Beck, Katz, and Tucker (1998), who advise using either temporal dummies or natural cubic splines, which essentially impose smoothing functions on the

³⁰ As an alternative to frailty models, researchers may also use “split-population” or “immune fraction” models that assume that a certain proportion of the population will never fail (Box-Steffensmeier & Jones, 2004).

temporal dummies and therefore use up fewer degrees of freedom.³¹ However, as mentioned above, Carter and Signorino (2010) report that of 119 studies (as of July 2006) that use Beck, Katz, and Tucker's time dummies or spline specification for duration dependence, only 7 actually plot and interpret the hazard. They speculate that this may be the case because researchers do not really understand how to use and interpret splines, especially with regard to knot placement. As for time dummies, they also use simulations to show that the use of time dummies can lead to statistical separation problems. They explain that it is not uncommon in empirical studies that the time dummies (or some range of them) perfectly predict the outcome variable, especially when the data show decreasing hazards over time. In addition, if there are many time periods with many corresponding dummies, authors will be less likely to report estimates for them, and estimation efficiency will be reduced.

When we are interested in the form and meaning of duration dependence, there is also an argument to be made for parsimony. Why use a needlessly complicated specification for duration dependence when a simpler one provides an equally good or better fit? Moreover, if one has substantive hypotheses about duration dependence, time dummies or splines may not provide the most straightforward test. For example, perhaps a policy diffusion researcher hypothesizes that a particular policy will exhibit a bandwagon effect—that the hazard of adoption will rise over time simply because the total number of adopters is increasing. In other words, the researcher wishes to test for positive duration dependence. The simplest way to do this in a discrete-time model would be to use a likelihood ratio test to compare a model with no

³¹ According to Beck, Katz, and Tucker (1998, p. 1270): "Natural cubic splines fit cubic polynomials to a predetermined number of subintervals of a variable. These polynomials are joined at 'knots,' with the number and placement of the knots specified by the analyst. Smoothness is imposed by forcing the splines, and their first and second derivatives, to agree at each of the knots. Thus each knot only uses up one degree of freedom, so that we can flexibly fit a cubic spline using up only a very few degrees of freedom. The estimated spline coefficients can then be used to trace out the path of duration dependence."

specification for time (a constant hazard model) with a model that includes a function of time (e.g. linear, logged, etc.).³² If the inclusion of the time variable improves the fit of the model and has a positive coefficient, then the researcher can conclude that there is positive duration dependence.

In general, researchers can use likelihood ratio tests for nested models, as well as AIC and BIC scores for non-nested models to compare various specifications for duration dependence. Options include linear, quadratic, or higher order polynomials, logged time, and splines (Beck et al., 1998; Box-Steffensmeier & Jones, 2004; Carter & Signorino, 2010). Which time specifications to compare is at the discretion of the researcher but should be guided by theoretical expectations and visual inspection of the Cox baseline hazard (see step 2 above). For example, if a researcher hypothesizes that diffusion will be characterized by a competitive race with an inverted U-shaped hazard, then he or she should fit models with non-monotonic baseline hazards—quadratic or splines. The goal is to explore the nature of duration dependence by comparing multiple models. Just as researchers take care to properly model covariates of interest with transformations, interactions, etc., the proper modeling of time is a key component of event history models.

Conclusions

I argue that the analysis of duration dependence in policy diffusion event history studies can yield many insights in explaining the nature of policy diffusion processes. As Elkins and Simmons (2005) have stressed, policy diffusion is a process, not an outcome. A process implies change over time, so temporal dynamics should be central to the proper modeling of diffusion processes. Fortunately, event history analysis provides many tools for examining time dynamics,

³² This is analogous to comparing an exponential and Weibull model in the continuous time context.

including non-proportional hazards as well as several parametric and non-parametric techniques for modeling the baseline hazard. Policy diffusion scholars have underutilized these tools, relying instead on the traditional regression paradigm that focuses on static covariate effects and techniques to control for assumptions violations.

In particular, I argue that while policy diffusion researchers have adapted the approach of spatial models to create covariates that measure heterogeneous mixing diffusion processes, they have given much less consideration to diffusion mechanisms that involve homogeneous mixing or system-level dynamics such as bandwagon effects, critical mass effects, competitive races, policy outbreaks, and generalized learning processes. These diffusion stimuli occur at the system level, not the individual government level, so they do not exhibit variation across individuals and cannot be adequately measured with individual-level covariates. Instead, their effect will be evident in the baseline hazard or duration dependence. Modeling duration dependence can be tricky, however. Researchers also need to account for unobserved heterogeneity, omitted variables, exogenous influences, and non-proportional hazards using some of the techniques described in this essay. Ultimately, duration dependence and covariate effects are complementary tools that researchers can use to gain a more nuanced understanding of policy diffusion processes.

References

- Balla, S. J. (2001). Interstate professional organizations and the diffusion of policy innovations. *American Politics Research*, 29(3), 221-245.
- Baybeck, B., Berry, W. D., & Siegel, D. A. (2011). A strategic theory of policy diffusion via intergovernmental competition. *Journal of Politics*, 73(1), 232-247.
- Beck, N., Katz, J. N., & Tucker, R. (1998). Taking time seriously: Time-series--cross-section analysis with a binary dependent variable. *American Journal of Political Science*, 42(4), 1260-1288.
- Bennett, D. S. (1999). Parametric models, duration dependence, and time-varying data revisited. *American Journal of Political Science*, 43(1), 256-270.
- Berry, F. S., & Berry, W. D. (2007). Innovation and diffusion models in policy research. In P. Sabatier (Ed.), *Theories of the policy process* (pp. 223-260). Boulder, CO: Westview Press.
- 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.
- Berry, W. D., & Baybeck, B. (2005). Using geographic information systems to study interstate competition. *American Political Science Review*, 99(4), 505-519.
- 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.
- Boehmke, F. J., & Skinner, P. (2012). State policy innovativeness revisited. *State Politics and Policy Quarterly*, 12(3), 303-329.
- Boushey, G. (2010). *Policy diffusion dynamics in America*. New York, NY: Cambridge University Press.
- Boushey, G. (2012). Punctuated equilibrium theory and the diffusion of innovations. *Policy Studies Journal*, 40(1), 127-146.
- Box-Steffensmeier, J. M., Reiter, D., & Zorn, C. (2003). Nonproportional hazards and event history analysis in international relations. *Journal of Conflict Resolution*, 47(1), 33-53.
- Box-Steffensmeier, J. M., & Zorn, C. (2001). Duration models and proportional hazards in political science. *American Journal of Political Science*, 45(4), 972-988.
- Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling : A guide for social scientists*. Cambridge; New York: Cambridge University Press.

- Braun, N., & Engelhardt, H. (2002). *Diffusion processes and event history analysis*. (MPIDR Working Paper No. WP 2002-007). Rostock, Germany: Max Planck Institute for Demographic Research.
- Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review*, 26(2), 175-188.
- Buckley, J., & Westerland, C. (2004). Duration dependence, functional form, and corrected standard errors: Improving EHA models of state policy diffusion. *State Politics and Policy Quarterly*, 4(1), 94-113.
- Carter, D. B., & Signorino, C. S. (2010). Back to the future: Modeling time dependence in binary data. *Political Analysis*, 18(3), 271-292.
- Darmofal, D. (2009). Bayesian spatial survival models for political event processes. *American Journal of Political Science*, 53(1), 241-257.
- Elkins, Z., & Simmons, B. (2005). On waves, clusters, and diffusion: A conceptual framework. *The Annals of the American Academy of Political and Social Science*, 598(1), 33-51.
- Franzese, R. J., & Hays, J. C. (2008). Interdependence in comparative politics: Substance, theory, empirics, substance. *Comparative Political Studies*, 41(4-5), 742-780.
- Gilardi, F., Füglistler, K., & Luyet, S. (2009). Learning from others: The diffusion of hospital financing reforms in OECD countries. *Comparative Political Studies*, 42(4), 549-573.
- Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3), 515-526.
- Grossback, L. J., Nicholson-Crotty, S., & Peterson, D. A. M. (2004). Ideology and learning in policy diffusion. *American Politics Research*, 32(5), 521-545.
- Ingle, W. K., Cohen-Vogel, L., & Hughes, R. (2007). The public policy process among southeastern states: Elaborating theories of regional adoption and hold-out behavior. *The Policy Studies Journal*, 35(4), 607-628.
- Maggetti, M., Radaelli, C. M., & Gilardi, F. (2013). *Designing research in the social sciences*. Los Angeles, Calif.; London: SAGE.
- Rogers, E. M. (2003). *Diffusion of innovations*. New York: Free Press.
- Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, 52(4), 840-857.
- Strang, D. (1991). Adding social structure to diffusion models. *Sociological Methods & Research*, 19(3), 324-353.

- True, J. L., Jones, B. D., & Baumgartner, F. R. (2007). Punctuated equilibrium theory: Explaining stability and change in public policymaking. In P. A. Sabatier (Ed.), *Theories of the policy process* (Second edition ed., pp. 155-188). Boulder, CO: Westview Press.
- 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.
- Weyland, K. G. (2006). *Bounded rationality and policy diffusion : Social sector reform in latin america*. Princeton, N.J.: Princeton, N.J. : Princeton University Press.
- Zorn, C. J. W. (2000). Modeling duration dependence. *Political Analysis*, 8(4), 367-380.

Figure 1: Example of Baseline Hazard for Cox Model Estimates on State Film Incentive Adoption

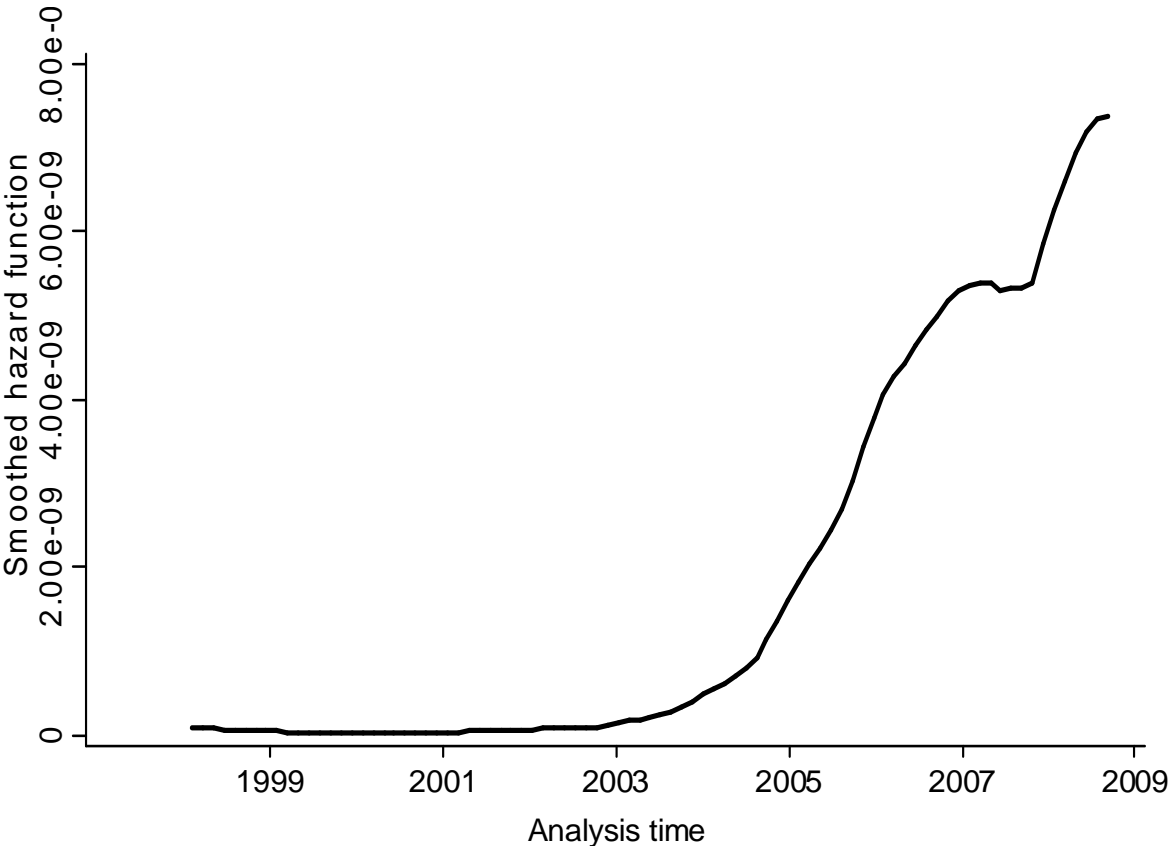
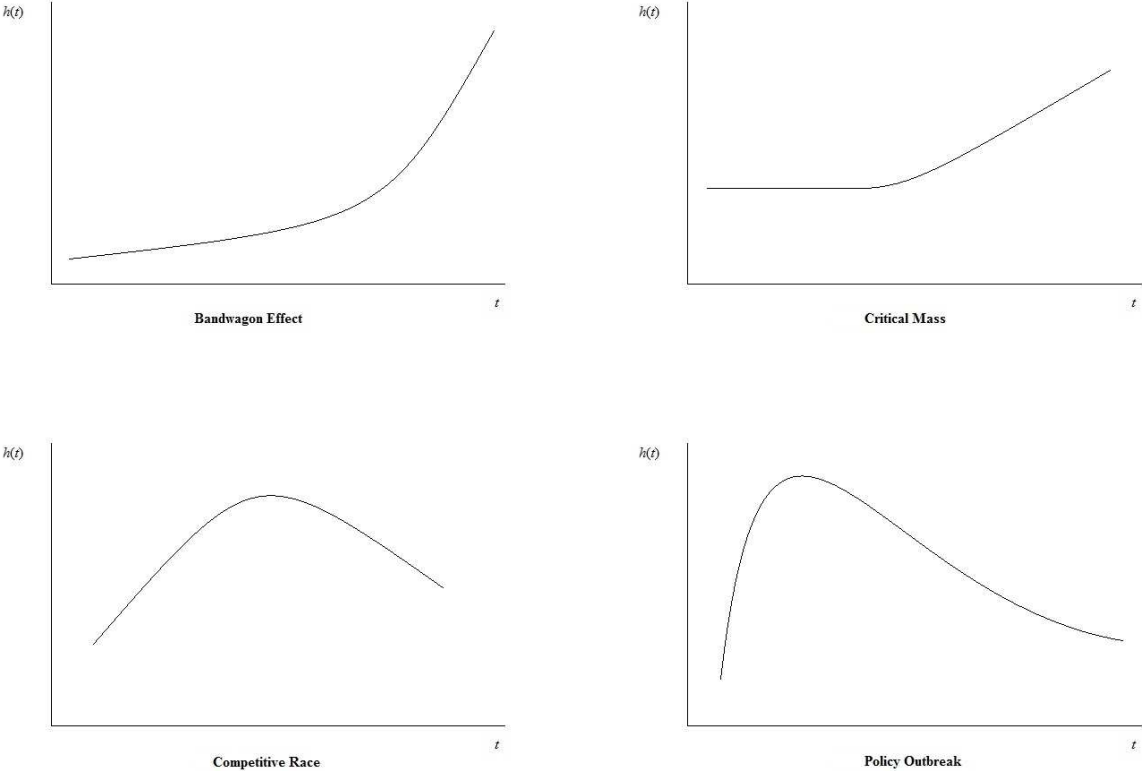


Figure 2: Baseline Hazard Plots for System-Level Diffusion Processes



Conclusions

While policy diffusion theory and methods can offer new insights into the study of state tax policy, the study of tax policy can also offer insights to the development of policy diffusion theory and methodology. In this dissertation, I have explored why states adopt tax incentives for business. I have examined the effects of various diffusion-related factors, as well as states' economic and political characteristics, on states' decisions to adopt five different types of business tax incentives: film incentives, Investment Tax Credits, apportionment changes, R&D Credits, and Job Creation Tax Credits. It turns out that what the evidence does not show is just as important as what the evidence does show. I find very few similarities in the sets of factors that influence the adoptions of these five different tax incentives.

For example, in most diffusion studies, it is hypothesized that states are more likely to adopt policies if more of their contiguous neighbors have adopted the policy, because neighbors provide either the most intense competition or most opportunities for learning. I find this to be the case only for R&D credits—neighbor effects are generally not significant for the other types of tax incentives in this dissertation. These findings are consistent with the idea that regionally-based competition is becoming an obsolete way to characterize the competitive pressures states face. As several of this study's interviewees remarked, states increasingly feel they are engaged in global competition for investment and jobs. In this case, the relevant question is not how many of a state's neighbors have already adopted tax incentives, but how many states anywhere and how many foreign countries have adopted tax incentives. While these global competitive pressures may ebb and flow over time, they exert an influence on all states.

The prevailing methodology in quantitative policy diffusion studies is ill-equipped for dealing with these types of global factors. Existing studies focus on factors that vary from state to state, such as unemployment rates, partisan control of state government, or proportion of

neighbors adopting. However, global pressures cannot be measured at the state level and are therefore likely to be overlooked by the usual methodology. As I have argued in each essay in this dissertation, the key to understanding these global diffusion factors is in examining duration dependence, or baseline hazard, which shows how the hazard of adoption changes over time controlling for other variables. In these essays, I have tried to draw on under-utilized techniques from event history analysis to show how the analysis of duration dependence can contribute to our understanding of diffusion dynamics for state tax incentives.

For example, in the first essay on state film incentives, I use a Cox model to uncover a pattern of positive duration dependence that suggests that there is a bandwagon effect in the adoption state film incentives—as more states adopt them, non-adopting states become more and more likely also to adopt. I am then able to explicitly model and test the significance of this bandwagon effect using parametric event history techniques. In the second essay, I again use Cox models to plot the baseline hazards for ITCs, apportionment changes, R&D Credits, and JCTCs. Interestingly, I find that all four baseline hazard plots follow an inverted U-shape, consistent with the idea that by adopting tax incentives, states are effectively competing in a zero-sum game for jobs and investment—a race to the bottom.

The third essay takes a more comprehensive methodological look at how to analyze and interpret duration dependence in the context of policy diffusion. It connects the specific findings about duration dependence and tax incentives in the first two essays to a larger discussion that has implications for other quantitative policy diffusion studies. In this essay, I argue that certain system-level diffusion dynamics like bandwagon effects and competitive races, will manifest as duration dependence. Therefore, an analysis of diffusion patterns is incomplete without a thorough examination of the baseline hazard. I provide a list of recommendations and strategies

researchers can use to incorporate the analysis of duration dependence into their studies, and I also explain how duration dependence is connected to other methodological issues like unobserved heterogeneity and non-proportional hazards.

Going forward, it is important to continue to combine the insights from tax competition and policy diffusion because they both address the same fundamental question of how interdependence between governments influences the policymaking process. More importantly, if we can develop a better understanding of how governments interact in making policy, we are then in a better position to ask more normative questions. Does it matter that there do not seem to be many common factors that lead states to adopt tax incentives? Should the federal government intervene to mitigate these apparent races to the bottom? Should states revise how they decide whether to adopt tax incentives? Without seriously debating and answering these questions, policymakers will likely continue to remain stuck in their “love/hate” relationship with tax incentives.