State Standards as Local Constraints: Evolution and Creationism in the High School Classroom

Michael B. Berkman
(mbb1@psu.edu)

Eric Plutzer
(EXP12@psu.edu)

The Pennsylvania State University

Abstract: Do state curricular standards the behavior of local school districts? More specifically, do they interfere with local teachers' and administrators' responsiveness to local district preferences? We explore these questions in the highly contested arena of instruction in evolutionary biology. Drawing upon an original national survey of high school biology teachers, we show that teachers are more responsive to public opinion when standards are weaker and not enforced through hi stakes testing. We further show that these constraints operate more on teachers with less knowledge about evolution, less self confident about evolution, and who have less seniority. Since teachers self select into districts with values similar to their own, our results suggest that the emerging policy focus on teacher training and certification will further constrain local behavior.

Presented at the seminar on *States as Facilitators or Obstructionists of Local Governments*, sponsored by the DeVoe Moore Center and the LeRoy Collins Institute at Florida State University. February 25-26, 2010. This research is supported by the National Science Foundation, Spencer Foundation and the John Templeton Foundation.

State Standards as Local Constraints: Evolution and Creationism in the High School Classroom

In 2005 the small central Pennsylvania community of Dover attracted national attention when its school board questioned the place of evolutionary biology in its high school classrooms. The controversy began when a senior's science project, a mural depicting human evolution, was removed from a high school science classroom and burned by a school janitor.[1] And it was kept alive by the elevation to School Board President of Alan Bonsell, an ardent creationist who believed that the earth was 10,000 years old, that humans did not evolve from other species, and that science teachers who taught evolution were lying to their students.

School board meetings in Dover were tense and conflictual. Many young evangelical families had migrated into the community, and this introduced to the debate a vocal and vigorous disapproval of the secular world and the current school science curriculum. The school board members reflected this strong anti-evolution sentiment. They debated teaching creationism outright and adopting a new text giving equal time to creationism and evolution. Compromise suggestions, such as the adoption of a comparative religions course that would teach multiple creation stories, were soundly rejected. In the end, the board decided that teachers would read a statement questioning the validity of evolutionary theory and offering students the opportunity to learn about a prominent alternative by reading *Of Pandas and People*, an intelligent design textbook, multiple copies of which had been donated to the school library by a local church.

This policy would be challenged in federal court in Harrisburg, Pennsylvania. *Kitzmiller v. Dover* attracted wide-spread attention. Prominent scholars testified at the trial that intelligent design is not really science but rather creationism in another guise.[2] Bruce Springsteen weighed in during a concert in New Jersey.[3] A Time magazine cover story featured the trial and NOVA

[1] For our discussion of events in Dover we rely on Humes (2007).

[2] The witnesses for the plaintiffs in Dover were Brown University Biologist Kenneth R. Miller, Michigan State University Philosopher Robert T. Pennock, Georgetown University Professor of Theology John (Jack) F. Haught, Berkeley's Kevin Padian, McGill University's Brian Alters and Southeastern Louisiana University Philosopher Barbara Carroll Forrest.

[3] At a concert in New Jersey shortly before the start of the Kitzmiller v. Dover case, the rock superstar introduced his song about evolution-- Part Man, Part Monkey-- by telling the crowd that "folks in Dover aren't sure about evolution. Here in New Jersey, we're counting on it."

would eventually make a documentary about it. Scientists and the ACLU clearly saw the developments in Dover as disturbing, bad science education, and a significant violation of the separation of church and state. The teachers in Dover were appalled and refused to read the statement themselves.[4] The opinion written by Judge John Jones III was scathing in its criticism of the board. He found unequivocally that the board's policy amounted to an unconstitutional establishment of religion. In addition to the resounding legal defeat, the trial subject the citizens of Dover to a media feeding frenzy and cost the district over a million dollars in court costs and legal fees. In its aftermath, five school board incumbents were defeated in the next election (though all by narrow margins).

The story of Dover has been told many times. But we suggest there is another way to look at it. And that is that the board's policy in Dover—the reading of a statement questioning evolution and advocacy of Intelligent Design as an alternative-- was quite mild <u>relative</u> to the preferences of the residents of Dover and the board. In other words, the actions of the school board officials were severely constrained in comparison with what they would have liked to have done. Indeed, in their public comments and trial testimony their distaste for and lack of understanding of evolution were on full display. Given the opportunity they probably would have discarded evolution altogether. And so it is the story of these constraints that interests us. In particular, we look systematically at the extent to which state curricular and assessment policies constrain the behavior of local school districts and teachers when it comes to instruction in this contested area of the 9th and 10th grade science curriculum. Further, we shall show that when we consider policy as it is actually implemented, we can view public school teachers as potential facilitators of local policy preferences. But teachers constitute a double edge sword because they can serve to undermine policy as well.

Constraints on School Districts

To appreciate quite how moderate the Dover school district's final decision really was it is helpful to recognize the extent to which the public dislikes instruction in evolution and prefers

[4] The Dover teachers' said in a public statement: "To refer the students to 'Of Pandas and People' as if it is a scientific resource breaches my ethical obligation to provide them with scientific knowledge that is supported by recognized scientific proof or theory."

instead the introduction into the science classroom of Christian-inspired alternatives. Public opinion polls over thirty years have been consistent in their findings about this (Plutzer and Berkman 2008). For example, a 2005 Pew poll found that 38 percent of Americans prefer the teaching of creationism instead of evolution, while a strong majority of 58 percent of Americans in a 2006 Pew poll supported teaching creationism along with evolution. Other polls show similar opposition to a biology curriculum that teaches evolution alone (Table 1).

Table 1. Public preferences for teaching evolution and creationism, 1981-2005

	Recent polls		***Older polls***	
	VCU 2005	Harris 2005	CCD 1987	NBC 1981
Teach creationism only	21%	23%	11%	10%
Teach intelligent design only	5%	4%	-	-
Teach creationism & intelligent design	4%	-	-	-
Teach a combination including evolution	43%	55%	68%	76%
Teach evolution only	15%	12%	11%	8%
Don't know or no answer	12%	6%	10%	6%
Total	100%	100%	100%	100%
(N)	(1,002)	(1,000)	(1,708)	(1,598)

The public's clear reluctance to accept evolutionary theory and openness to non-scientific alternatives runs counter to the scientific community and multiple scientific organizations that have weighed in on the matter.[5] Among scientists there is a very strong consensus that, in the iconic words of geneticist Theodosius Dobzhansky, "Nothing in biology makes sense except in the light of evolution" (1973). Indeed, the only significant portion of the population that really accepts evolutionary biology is those with advanced college degrees (Berkman and Plutzer 2010).[6] And not surprisingly given its centrality over many decades to the fundamentalist movement, opposition is strongest among evangelicals.

Absent all Supreme Court limitations, local school districts and teachers might be expected to have diverse preferences ranging from wanting to teach the Biblical story of creation in their science classes to wanting to teach evolution as rigorously as possible with every possible challenge refuted. But years of Supreme Court rulings do not allow public opinion alone to determine what students are taught. While there is no constitutional requirement that students must learn about evolution or any other topic, the purposeful introduction of creationism into the science classroom is a violation of the 1st Amendment's prohibition against establishment of religion (for a summary of these cases see Moore 2002a; for an excellent historical review of the critical court cases see Larson 1989). Judge Jones' decision in *Kitzmiller* extends this prohibition to intelligent design as well, although it technically applies only to the Middle District of Pennsylvania.

But other constraints exist as well, and these too were evident in the Dover incident. Consider carefully the statement that the school board expected to be read in Dover's science classrooms, and note the emphasis the statement puts on meeting the requirements of the Pennsylvania Academic Standards:

> The Pennsylvania Academic Standards require students to learn about Darwin's theory of evolution and eventually to take a standardized test of which evolution is a part. Because Darwin's theory is a theory, it continues to be tested as new evidence is discovered. The

---

[5] For example, the National Science Teachers Association (2003), the National Research Council (1996) and the National Academy of Sciences (2008) have all published reports about the importance and centrality of evolution to the biology curriculum and its acceptance by the scientific community.

[6] Beyond holders of advanced degrees, Jewish Americans comprise the only other identifiable group with a majority favoring the teaching of evolution (and only evolution).

> theory is not a fact. Gaps in the theory exist for which there is no evidence. A theory is defined as a well-tested explanation that unifies a broad range of observations. Intelligent design is an explanation of the origin of life that differs from Darwin's view. The reference book, Of Pandas and People, is available for students who might be interested in gaining an understanding of what intelligent design actually involves. With respect to any theory, students are encouraged to keep an open mind. The school leaves the discussion of the origins of life to individual students and their families. As a standards-driven district, class instruction focuses upon preparing students to achieve proficiency on standards-based assessments.

Pennsylvania's science standards are generally recognized as among the nation's most rigorous in the treatment of evolution (Lerner 2000; Berkman and Plutzer 2009). And these standards clearly constrained the Dover school board. The statement reflects the reality for the school board that as "standards-driven district" they could not simply organize their curriculum as they wanted. Rather, students were required to learn about Darwin's theory to "achieve proficiency on standards-based assessments."

All school districts make determinations about what to teach - determinations that consider their state content standards and their relationship to a statewide assessment exam, if there is one. Of course, these determinations are just the formal portion of local curricular policy. The actual *de facto* policy is determined by how the policy is actually implemented (Lipsky 1980, e.g., 1 and 83). Thus teachers contribute to policy making as well and this, too, was at play in Dover. Dover's teachers took the unusual step of trying to explain to a skeptical board that their current instruction in evolution was simply the minimum required by state standards (Humes 2007). But more fundamentally, as street level *bureaucrats* with broad discretion over what they teach, teachers play the crucial role in implementation and they may be influenced not only by to local school board's stated goals, but also by state content standards, assessment policies, and their own professional identity. In short, while school district policy is important, especially in cases like Dover, it is what actually happens in the classroom that determines what students will actually learn. It is the behavior of teachers, and their constraint by standards, that we examine empirically in this paper.

To do so we rely on several original data sets. First, we analyzed the content standards of each state and coded them according to the extent to which they treated evolution rigorously, or not. Second, we administered the National Survey of High School Biology Teachers – a nationally representative survey of 926 public high school science educators. Because the teacher sample is a nationally representative probability sample, the districts in which they teach is also representative of all districts (with selection proportional to district size) and the teacher reports can be considered indicators of how curricular policy is actually implemented. Or in other words, each teacher can be regarded as a randomly selected teacher from all high school biology teachers in each district.

Our analysis here focuses on teacher responses to how they taught evolution and (if applicable) creationism along with information about teacher training, familiarity with evolutionary theory, and beliefs. Third, we computed estimates of public opinion at the school district level based on the analysis of years of national level surveys that asked specific questions about whether evolution or creationism should be taught in the science classroom. Combined, these data allow us to ask whether teachers, in deciding what to teach in the 9th and 10th grade science classroom, act in response to local opinion, independently of that opinion and according to their own preferences, or are constrained in their decision making by state standards and state examinations.

State Academic Curricular Standards and Examinations

In contrast to Dover, many contemporary political battles about creationism and evolution often revolve around the content and quality of these state-level policies (e.g., conflicts in Kansas, Ohio, Texas and Florida). This is not surprising given the increasing importance of academic standards in state education politics generally. Since the 1990s state school boards and education departments have spent considerable time and effort on the writing and periodic revision of state standards. In many subject areas these standards are “contested terrain” (Placier et al. 2002, 282) and their writing often makes news because the outcomes matter to teachers, parents, interest groups and ordinary citizens.[7] By 2000, forty nine states had adopted some form of science standards; between 2000 and 2008 a quarter of the states changed their standards

[7] Binder (2002) provides detailed case studies of standards revision in California and Kansas. For a more comprehensive set of examples of these political battles, see the website for the Center for Science Education (ncesweb.org), which keeps careful track of evolution issues throughout the American states.

concerning evolution enough to be noticed by the National Center for Science Education, a pro-evolution organization.[8]

Consequently standards range dramatically across the states. A 2000 grading of state science standards conducted by physicist Lawrence Lerner evaluated the rigor of each state's expectations about instruction in evolution. Kansas, the poster child for low quality standards, received an F-minus for its "disgraceful paean to antiscience" (Lerner 2000, 16) while Ohio received one of twelve F's reserved for states "who fail so thoroughly to teach evolution as to render their standards totally useless" (Lerner 2000, 16) and "as if it were not proper conversation in polite company" (2000, 16). On the other hand, public high school students in neighboring Indiana were offered an "exemplary" treatment of evolution (Lerner 2000) while New York emphasizes evolution by including mention of it in the first sentence of its standards; indeed, in the first word: "Evolution is the change of species over time." There are several assessments of differences in standards among states (Lerner 2000; Skoog and Bilica 2002; Swanson 2005) and all find that some standards come close to the requiring students to meet the "ideal" benchmarks from the *National Science Education Standards* (NSES) while others offer a cursory coverage of the topic if it is raised at all.[9]

Content standards stand as official written policies concerning what students should learn in their public school biology classes. But research on the effects of these standards on actual school instruction is limited and inconclusive (Moore 2002, Bandoli 2008; see Berkman and Plutzer 2010 for a critique of this literature). And none of this research has looked at how policymakers seek to reform the "architecture" of these policies to better insure that they are implemented as intended (Reenock and Gerber 2008; McCubbins, Noll, and Weingast 1987).[10]

[8] Based on a state by state review of the National Center for Science Education's news archives for 2000-2008 [http://ncseweb.org/news , last accessed June 22, 2009]. Our count includes states that strengthened or weakened or their state standards through the legislature, state board of education, or state school board.

[9] The National Science Education Standards (NSES) are one of several publications from prominent national science organizations that propose sample standards rich in evolutionary theory.

[10] A test is high stakes if a student's graduation or progression to the next grade is dependent upon their test performance or if schools face sanctions if a certain number of students fall below a particular threshold. High Stakes testing in science was not a mandate of No Child Left Behind in the spring of 2007 but it was expected to become one in next academic year.

In the realm of curricular reform in the last decade, this architecture involves the introduction of high stakes student testing. State policymakers can use testing to monitor the implementation of standards, sanction teachers and school districts who fail to adhere to them, or signal the importance of different components of a state's standards. The specific choice of test questions adds "authority to selected goals and topics" (Archbald and Porter 1994, 22). Standards aligned with examinations should "drive instruction" in the direction intended by policy makers (Airasian 1988, 305) and therefore additionally constrain local school districts and teachers.

Teachers as Street Level Bureaucrats

Teachers "ultimately decide the fate of national and state science standards" (Spillane and Callahan 2000, 401-402). They determine whether standards will successfully constrain communities and their elected officials in their efforts to offer full instruction in evolution, undercut its theoretical importance, or introduce non-scientific alternatives through "the curricular and instructional decisions they enact within the specific, particular contexts of their own classrooms" (Goldstein 2008, 449). They are no different in this way from other unelected public servants who are responsible for implementing policies and programs developed in rule-making or law-making bodies. Often referred to as *street-level bureaucrats,* teachers like social service employees, police, corrections workers, and mental health counselors have wide discretionary authority in their work with client populations (Lipsky 1980; Maynard-Mooney and Musheno 2003; Smith 2003; Keiser 1999; Meier and O'Toole 2006).

Teacher implementation of standards therefore presents what is commonly referred to as a *principal-agent* problem (McCubbins, Noll, and Weingast 1987; Brehm and Gates 1997): Principals are those who write laws and rules subject to electoral constraints—for example, the state policymakers and their appointees who develop standards—while agents are those who are relied upon to carry them out. Under this "top-down" model of democratic control, bureaucrats are expected to be responsive to those who make laws and set policy (Meier and O'Toole 2006). As teachers interpret the range of "state-, district-, and school-led policies" affecting their work they, in effect, *make* education policy (Goldstein 2008, 449). Teachers are therefore critical actors in determining whether standards operate as a constraint on local preferences.

But teachers can also serve to translate local public opinion and interests into policy, therefore undermining the capacity of standards to operate as constraints. Fredrick Mosher (1982) argues that the capabilities, orientation and values of unelected bureaucrats are shaped by their background, training, and education, while Brehm and Gates find that the "very best explanation" for why agents generally do not do as their principals want is that they do not share the same values (1997, 20). Other work on street level bureaucrats also finds or suggests that they bring their own attitudes and values to bear on how they do their jobs (Meyers and Vorsanger 2003; Keiser and Soss 1998; Keiser, forthcoming; 1999). Since attitudes toward evolution are part of teachers' fundamental values they are not easily transformed (Bishop and Anderson 1990; Lawson and Weser 1990). And decades of research on street level bureaucracy have shown how local culture and politics can influence policy implementation (e.g., Wiessert 1994; Soss 2000; Wilson 1989; Whitford 2002; Percival, Johnson and Nieman 2009). Indeed, Kenneth Meier and his colleagues (e.g., Meier and O'Toole 2006) argue that popular control is most effectively achieved not from the top but from the bottom- up, when bureaucracies are responsive to the clientele and communities they serve.

This suggests two mechanisms that might enhance the representation of local sentiment. One is that teachers share values with their communities. The simplest way that teachers' values come to match those of the communities is through the processes of hiring and retention. As Percival, Johnson and Nieman note, "the staff and leadership of local agencies might … be recruited in ways that reflect the views of the local legislature [or] … School Board" (2009, 166). In this way, the policy preferences of the broader community can lead to the hiring of teachers who share the community's values and beliefs. Indeed, research on the geographic preferences of those seeking their first teaching jobs show that "In seeking their first teaching jobs, prospective teachers appear to search very close to their hometowns and in regions that are similar to those where they grew up" (Boyd, Lankford, Loeb and Wykoff 2005, 127). The authors estimate that "Sixty-one percent of teachers entering public school teaching in New York from 1999 to 2002 first taught in schools located within 15 miles of their hometown" (2005, 117). This *assortative hiring* would produce similarity between community values and those of newly hired teachers. It is in this sense that teachers can serve as a resource that can help defend local preferences from state or federal obstruction.

Local environments may also influence policy implementation via the presence and strength of local advocacy groups or by increasing the presence of individuals who would seek to directly influence instruction. Organized groups will not merely reflect a diffuse public opinion but are effective when they can apply pressure to elected officials (school boards), supervisors (superintendents and principals), and to front-line public employees such as teachers (see e.g., Scholz and Wei 1986; Sabatier, Loomis and McCarthy 1995). Or public employees may feel that it is proper for their implementation of policy to reflect local sentiment (Percival et al. 2009). In his influential *Street Level Bureaucracy*, Lipsky argued that when local communities are indifferent to how a public policy is implemented, street level bureaucrats generate their own goals and objectives. But when the community is characterized by a strongly held consensus, civil servants will respond to community sentiment (1983, 46).

Data and Measures

To assess the importance of standards as a constraint on local preferences, as well as of teachers role in translating these preferences in the classroom, we need measures of teacher behavior, teacher values and knowledge of evolution, state standards, and public opinion at the school district level.

<u>Teacher Values, Knowledge, and Evolution Policy as Enacted by Teachers</u>

Measures for teachers' values, knowledge and classroom practices were developed from the National Survey of High School Biology Teachers, which we fielded in the spring of 2007. Beginning with a random sample of two thousand teachers of biology in US public schools, we designed a brief questionnaire and fielded the survey using the *Tailored Design Method* for mail surveys (Dillman 2000), along with a web survey option and additional email follow-ups for those teachers for whom we had valid email addresses. We will utilize data from the 926 usable respondents which represent (after excluding 58 such "out of scope" respondents) a response rate of 50% (see the Appendix for details).

The questionnaire was designed in the context of previous studies and we adapted question wording from other studies whenever possible and appropriate. We then asked six high school biology teachers from our own community to look at the initial draft of the questionnaire and provide us with suggestions for improvement. The final questionnaire reflected the feedback we received and was six pages long, taking a teacher about 15 minutes to complete. The survey contained questions about the content of their most recently taught biology course, more specific questions about the teaching of evolution in particular, and a variety of background questions.

Unlike some studies which attempt to align instructor teaching to specific pedagogical benchmarks, we view the teaching of evolution and creationism as a *political outcome*. We therefore measure the extent to which classroom teachers took up the major perspectives and positions that define the political battle over evolution. From the perspective of *establishment science*, two points are crucial: that evolution is a fact, and that evolution should be the central organizing principal of any modern biology class.

On the other hand, evolution opponents cannot expect public school teachers, in the first decade of the 21st century, to devote considerable attention to scientific creationism or intelligent design creationism. Instead, their goal in recent years has been to expose students to alternatives to evolutionary biology and to suggest the possibility that evolutionary biology may be fundamentally wrong. The policy of the Dover School District illustrates this modest goal. The Board did not legislate "equal time" or specific lesson plans for intelligent design. Rather, the board appears to have had two goals. First, by emphasizing to students that "evolution is a theory" they sought to sow skepticism about the specific facts and conclusions of evolutionary biology as typically introduced in textbooks and classrooms (e.g., that anatomically modern humans evolved in Africa about 200,000 years ago). Second, the statement (intended to be read by science teachers and expressing the authority of the school board) suggested that there were valid alternatives to evolutionary biology that merit respect and consideration. Other modest goals might include the elimination or watering down of evolution instruction, even without a corresponding introduction of creationist or intelligent design perspectives.

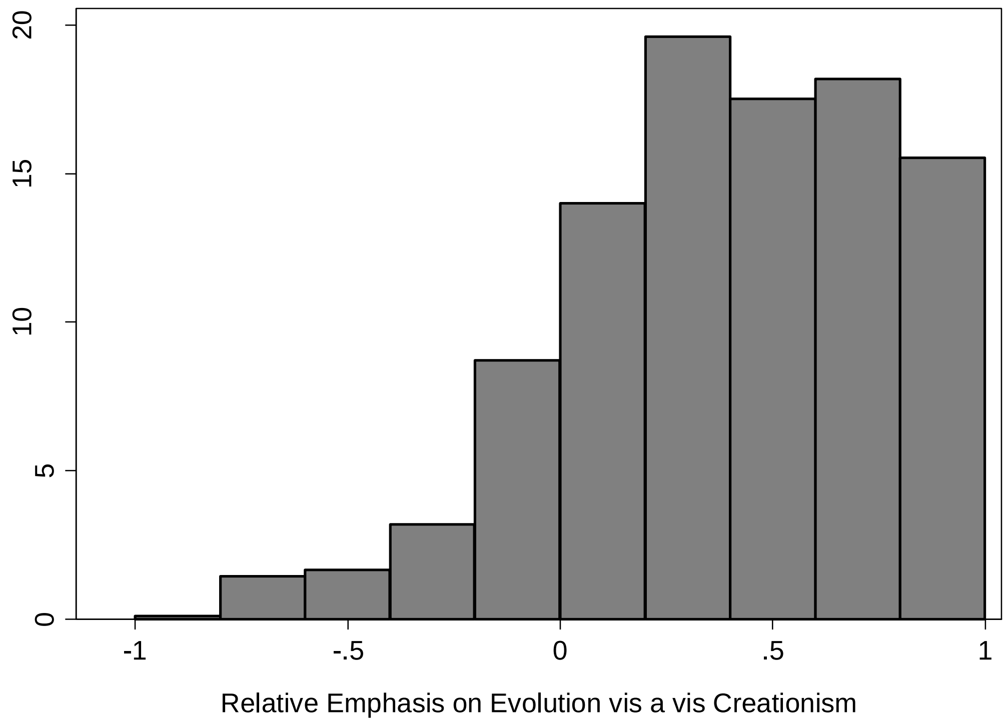
Our survey included five questions that tapped into these competing goals for science instruction. The first was, "*when I teach evolution (including answering student questions) I emphasize the broad consensus that evolution is a fact, even as scientists disagree about the specific mechanisms through which evolution occurred*." This is a key element in scientific organizations' *public science* activities and go directly to both young earth arguments and indirectly address the ways that antievolution activists exploit disagreements among scientists to undermine evolution in general.

Second, we asked "*I believe it is possible to offer an excellent general biology course for high school students that includes no mention of Darwin or evolutionary theory*." This addresses the desire to remove evolution entirely from high school biology – often urged as appropriate in order to avoid offending or infringing on the rights of some Christians. But it also measures the extent to which practicing high school biology teachers embrace or reject a key goal of the National Academy of Sciences and virtually all other scientific organizations. But the minimal inclusion of evolution is not really what establishment science recommends. They argue that a student's understanding of biology is enhanced when evolution is not just a short topic but a *unifying theme* that is woven throughout the biology curriculum. Our third question addressing scientific priorities asks teachers whether "*Evolution serves as the unifying theme for my course*."

We also had two questions concerning the endorsement of creationist ideas during formal class instruction. The first asked about whether the teacher *personally* endorses creationism: "*When I do teach about creationism of intelligent design (including answering student questions) I emphasize that this is a valid, scientific alternative to Darwinian explanations for the origin of species*." The second question was similar and contained the same introductory phrase, but focused on the scientific community, "… *I emphasize that many reputable scientists view these as valid alternatives to Darwinian theory*."

These items are highly intercorrelated and allowed us to compute a summary scale of high reliability (Cronbach's alpha = 0.79). The scale reflects the mean score of the five items after each has been standardized. Once created, the index was rescaled so that the minimum score (strong endorsement of creationism) is -1 and the maximum score (strong endorsement of evolutionary biology) is +1. The distribution of teacher reports is summarized in Figure 1, below.

Figure 1: Relative emphasis on evolution and creationism for 926 high school biology teachers.



Scores in the middle range (roughly -.4 to +.4) typically involve a mixture of positive emphases in both directions. In the tails of the distribution, teachers tend to be exclusively favoring evolutionary biology or creationism, but with differing levels of emphasis or enthusiasm. All told, 21% of the teachers endorsed at least one of the two questions concerning the teaching of creationism. In contrast, 47% of the teachers endorsed the three pro-evolution statements and the figure clearly shows that explicit endorsers of creationism are in the minority. But only 12% of the teachers strongly agreed with all three pro-evolution statements, resulting in considerable variation in teaching practices. As we note elsewhere (Berkman and Plutzer 2010), the practices of the large number of teachers in the middle of the distribution often serve to undermine science and leave students with the impression that evolution is something that can be debated by lay citizens, much like debates concerning situational ethics.

<u>Coding State Science Standards</u>

The challenge in coding standards is to distinguish between the absence of attention to evolution in some states, the marginal and coded attention it receives in others, and its prominence in others.

We rated the content standards of each state on three criteria, generating numerical scales for (1) the *prominence* of evolution in the curriculum, (2) the extent to which evolution served as the or one of several *guiding themes*, and (3) the degree to which standards were sufficiently *specific* to guide teacher behavior (see Appendix for details on our coding procedures). These three numeric ratings were then combined into a standardized scale (having a mean of zero and a standard deviation equal to one), that measures the rigor with which evolution is addressed in each state's standards. High scores indicate states whose content standards give significant attention to evolution, identify evolution as a major theme, and provide very specific guidance to teachers on the specific evolution topics that students are expected to learn. The scale has an estimated reliability of 0.88 (Cronbach's alpha).

We assessed the validity of the scale in two ways. First we calculated the correlation between our measure and the scores generated by Lerner's team of expert raters for the twenty five states *that did not revise their standards in the intervening seven years*. The correlation of 0.65 suggests that even though our content domains and grade spans were different, both scales are measuring the same general construct. Second, we compared our scale to three independent assessments of content benchmarks and showed a positive correlation in each instance. The details of how we conducted our content analysis, assigned numeric values to our observations, and how assigned we assessed validity are also detailed in the appendix.

Sanctions and Monitoring

In some states standards are coupled with assessment tests, in effect strengthening their potential to direct classroom instruction. A careful analysis by Editorial Projects in Education (2007) indicates that thirty one states had science examinations aligned with their state standards at the high school level during the 2006-2007 school year. We utilize the Editorial Projects codes.

Public Opinion Toward Evolution

We generated estimates of public opinion at the school district level in a two step process. First, we used Bayesian multilevel modeling with imputation and post-stratification (MLM-IPS; sometimes called "Mr. P" to refer to "Multilevel Regression with Poststrafication") to estimate public opinion for each of the fifty states. The general technique was developed by Little and

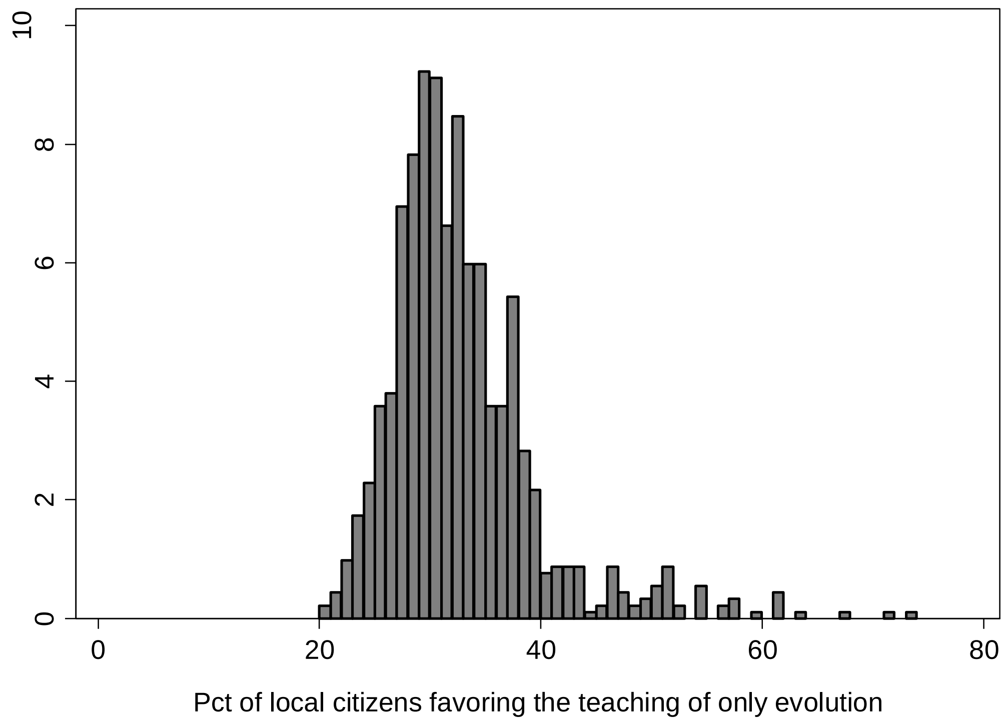
Gelman (1997) and has been applied and validated recently in a number of papers (Park, Bafumi and Gelman 2006; Lax and Phillips 2009; Pacheco 2008).

We then used Ardoin and Garand's (2003) "top down" regression method of estimating the opinion in school districts. The basic idea is that we estimated a regression of public opinion at the state level using the state's percentage of white evangelicals and the state's percentage of adults with masters or doctoral degrees. At the state level, these two variables explain 66% of the variance in the public support for teaching evolution. We then used school district level measures of these identical variables to get out of sample predictions.[11] This variable ranges from 0.20 to 0.74 and is interpreted as the percentage of the public that favors teaching evolution only (i.e., without any introduction of creationism) with a mean score of 0.33. [12] Figure 2 shows the distribution of teachers in terms of the public opinion in the communities in which they work. It shows that only the tiniest fraction (4%) of public high school high biology teachers work in communities in which evolution enjoys majority support.

[11] Educational attainment statistics were available at the level of the school district. The measure of evangelicals is for the county in which the school district is located.

[12] We note that we achieve identical results when simply using the local number of evangelicals and the local percentage of advanced degree holders. But having these measures scaled in terms of public opinion is both convenient and theoretically appropriate.

Figure 2: Local public opinion in support of teaching only evolution in public schools



Constraining Local Preferences in the Implementation of State Standards

With our focus firmly on teachers as policy implementers we can test whether standards operate as a constraint on community preferences. Our basic model posits that the emphasis a teacher places on evolutionary biology in their classroom instruction is a function of community sentiment *conditioned* by state standards. Indeed, it is the expectation (from a top-down perspective) that standards will reduce discretion at the local level, whether that discretion is exercised at the level of school district or the classroom.

$$\begin{aligned} &Evolution\ Emphasis \\ &\quad = \beta_0 + \beta_1(Public\ Opinion\ Toward\ Evolution) + \beta_2(Rigor\ of\ Standards) \\ &\quad + \beta_3(Public\ Opinion\ Toward\ Evolution * Rigor) + \ e \end{aligned}$$

However, the principal agent literature emphasizes that in the absence of both monitoring and sanctions, general policies may not be faithfully implemented by agents. High stakes assessment tests achieve both monitoring (student performance on the assessment test is typically reported at the school level, and often at the level of the classroom) and sanctions (resources may flow to high performing schools, property values may reflect school scores, and in some cases students may be penalized for poor performance). Thus we expect the basic model to hold in states with high stakes tests and for only local sentiment to matter in states without such assessments. We therefore estimated the model separately for assessment and non-assessment states. The results are reported in Table 2.

Table 2. Effect of local public opinion, standards, and testing on enacted evolution policy in the classroom

A. States without high stakes science assessments (J = 18, N=263)

	B	Robust Std Err	t	p
Local support for evolution	1.67	0.45	3.67	0.00
Rigor of state standards	-0.11	0.35	-0.30	0.77
Opinion x standards	0.44	1.03	0.43	0.68
Constant	-0.16	0.15	-1.06	0.31
R^2	0.09			

B. States with high stakes science assessments (J = 31, N=639)

	B	Robust Std Err	t	p
Local support for evolution	1.04	0.19	5.62	0.00
Rigor of state standards	0.15	0.05	2.96	0.01
Opinion x standards	-0.33	0.13	-2.65	0.01
Constant	0.08	0.07	1.07	0.30
R^2	0.06			

In the top panel we see that only local public opinion is a significant predictor of curricular standards as enacted in the classroom. Local public opinion accounts for about 9% of the overall variation and has a fairly strong impact. A one standard deviation shift in public opinion results in a shift of about a third of a standard deviation in teaching practices.

In the second model, we see a rather different picture. In a state with the mean degree of evolution rigor in standards, the impact of public opinion is about 40% lower when the state has a high stakes assessment test. And the effect of opinion diminishes further as the standards get increasingly rigorous. Thus state content standards, which have been a focal point of both pro- and anti-evolution activists and interest groups, only obstruct local policy preferences when accompanied by a high stakes examination.

Teacher Values

The models above suggest that in the absence of testing, local preferences find expression in the practices of local classroom teachers. When testing is in place, local preferences matter only when the standards lack rigor and specificity. But when rigorous standards are accompanied by a high stakes test, local preferences are thwarted.

This raises two questions: (1) When teachers have considerable autonomy, what explains their classroom choices? And (2) by what mechanism can local communities bring about desirable policy *as implemented*? The answers to both questions, we suggest, are closely related and concern the qualifications and values of the teachers themselves. If a local school district wishes its students to be exposed to the most rigorous treatment of evolutionary biology it will seek teachers with two characteristics: first, that they personally believe in organic evolution and, second, that they have outstanding scientific training in addition to their teaching certification. On the other hand, districts with strong creationist preferences will hire teachers who are themselves creationists or those with limited scientific training who will be less confident in teaching material that might be controversial in the community. In a slightly different context, we have previously referred to this as the *assortative hiring* model and we explore that next.

The assortative hiring model boils down to two propositions. The first is that individual teacher characteristics are important predictors of how they will enact policy in the classroom. And the second is that districts tend to hire teachers who share the values of their communities.

To test the first proposition, we add several teacher characteristics to our empirical models. First we add a measure of teachers' personal beliefs about human origins.

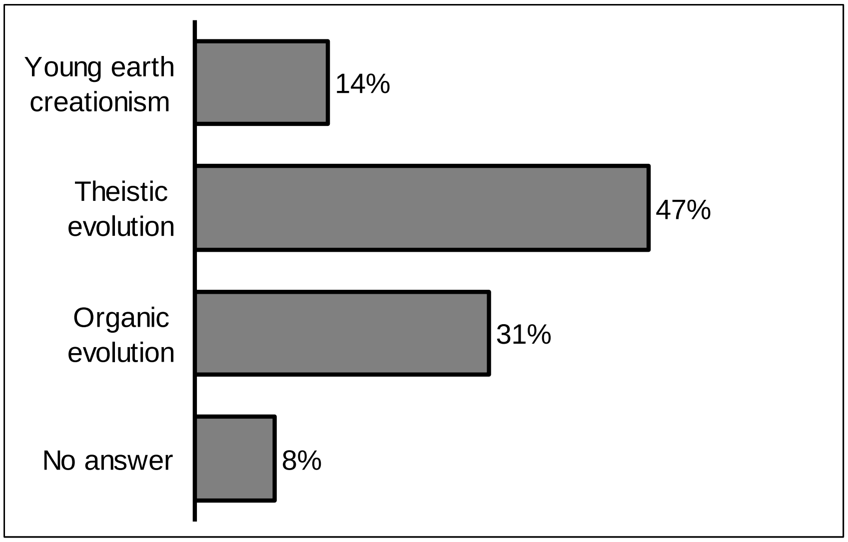
In our survey we asked each teacher the same question about beliefs about human origins that is frequently used in public opinion polls:

Now, regardless of what you do in the classroom, we would like to ask about your own personal views. Which of the following statements comes closest to your views on the origin and development of human beings?

> *Human beings developed over millions of years from less advanced forms of life, but God guided this process.*
>
> *Human beings developed over millions of years from less advanced forms of life, but God had no part in this process.*
>
> *God created human beings pretty much in their present form at one time within the last 10,000 years or so.*

The responses to this question are summarized in Figure 3 and show that only about three quarters of teachers report a personal belief in the evolution of human beings. The largest number of these teachers, 47% of the total, selected the *theistic evolution* perspective (that evolution occurred and had divine guidance) and the remainder, 31% of the sample, selected the option representing what is often called the *organic evolution* perspective. A total of 14% endorsed the *young earth* position that God created human beings recently and a sizable number of teachers declined to answer this question (in some cases telling us that their views were too nuanced to place in any of the choices). The answers to this question are dummy coded with the intermediate response, "God guided the process," serving as the contrast category.

Figure 3: Teacher beliefs about human origins (N=926)



<u>Educational Background and Certification</u>

Teachers in all states are expected to meet certain minimum requirements for certification. But this does not mean that all science teachers are equally knowledgeable about evolutionary theory or science generally. Indeed, Rutledge and Warden find among a sample of Indiana biology teachers a rather low level of understanding -- and therefore acceptance -- of evolutionary theory, well below what would "have been expected" among science teachers and more "in league with what is reported among the general public" (2000, 2). They attribute this, at least to some degree, to the teachers' college education, arguing that many teachers are not sufficiently trained to "possess a thorough knowledge of evolutionary theory and its place in the discipline of biology" (2000, 29). In general, teachers who understand a subject well will teach it differently than a subject they know less well (Carlsen 1991). Indeed, teacher knowledge and understanding may be especially germane to classroom instruction when the topic is evolution. Evolution can be a highly stressful topic many teachers, but less stressful for those teachers who are more confident and comfortable with the material (Griffith and Brem 2004). We can reasonably expect that these will be the teachers who are better trained in the subject.

Our survey allows us to look in some detail at our teachers' educational attainment so we should be able to determine the extent to which their understanding of evolution affects their instructional decisions. We know, for example, that fifty one percent of the teachers in our sample earned a bachelor's degree in biology or another scientific field; of those, about a third also received a graduate degree in a scientific field. That means that half our teachers were not science majors at all, although most had a minor concentration in biology or another scientific discipline, and some later earned a master's degree in biology or a related field. Therefore, the sample can be divided into three groups: those who do not hold any degree in science (38%), those who hold a bachelor's degree in science (37%) and those who hold a master's or doctoral degree in science (25%). Along the same lines, we classified teachers by the number of college level credit hours they completed in biology and classify teachers as having a low number (under 24 credits, about 6% of our sample), a medium number (25-40 credits, about 30%), or a high number (over 40 credits). We also asked about each teacher's certification status. Eighty percent of teachers (80%) held the normal teaching certification for their state while the remainder had temporary, provisional, or non-traditional certification.

One aspect of teachers' college level training that would likely make them more informed about evolution and more comfortable teaching it is whether or not they took a college course specifically in evolutionary biology. Randy Moore, who has studied biology teachers in a variety of states, concurs. Moore argues that evolution- related courses would "improve the status of evolution-related education" and make a difference in how well teachers understand evolution (2002, 380). This is because even though most college-level biology courses are likely to include discussion of evolution, a focused semester-long class may have a greater impact on how students think about evolution. A stand-alone course in evolution leads to both greater "acceptance" and "understanding" of evolution generally, and of human evolution in particular (Ingram and Nelson 2005; see also Bishop and Anderson 1990). Indeed, based on their study of Indiana teachers, Rutledge and Mitchell found that teachers who took a course in evolution had greater acceptance of evolution, taught more hours of it, and were able to demonstrate more sophisticated grasp of the material through concept maps (2002). Among our teachers, 43 percent took a course in evolutionary biology; not surprisingly, most of these teachers either majored in a science or took an advanced degree in one.

These measures of teacher credentials, along with their personal beliefs, are added to our previous models and the results are reported in Table 3. Looking first at the non-testing states, we see that the impact of local opinion is reduced by about 60% once we account for the characteristics of the teachers (remaining significant in the expected direction at 0.05 using a one-tailed test). Of these, two stand out as important: their personal beliefs and whether they completed an evolution class. Completing an evolution class moves teachers significantly in the pro-evolution direction, but only by about a tenth of a point on our -1 to +1 scale of teaching practices. In contrast, young earth creationists score more than a standard deviation below theistic evolution believers and nearly two standard deviations below those who believe in organic evolution. We see very similar results in the high stakes testing states (lower panel), although the cumulative impact of teachers' personal beliefs is slightly smaller (as principle-agent models would predict, given the enhanced monitoring and potential sanctions that a high stakes brings).

Table 3. Effect of local public opinion, standards, and teacher characteristics on enacted evolution policy in the classroom

A. States without high stakes science assessments (J = 18, N=244)

	B	Robust Std Err	t	p
Local support for evolution	0.71	0.39	1.82	0.09
Rigor of state standards	-0.05	0.17	-0.29	0.78
Opinion x standards	0.43	0.53	0.80	0.44
Young earth creationist	-0.43	0.08	-5.26	0.00
Organic evolutionist	0.20	0.04	5.63	0.00
Number of biology credits (0-2)	0.00	0.03	0.03	0.97
Number of science degrees (0-2)	0.01	0.02	0.69	0.50
Completed evolution course	0.09	0.03	3.06	0.01
Constant	0.07	0.14	0.50	0.62
R^2	0.36			

B. States with high stakes science assessments (J = 31, N=583)

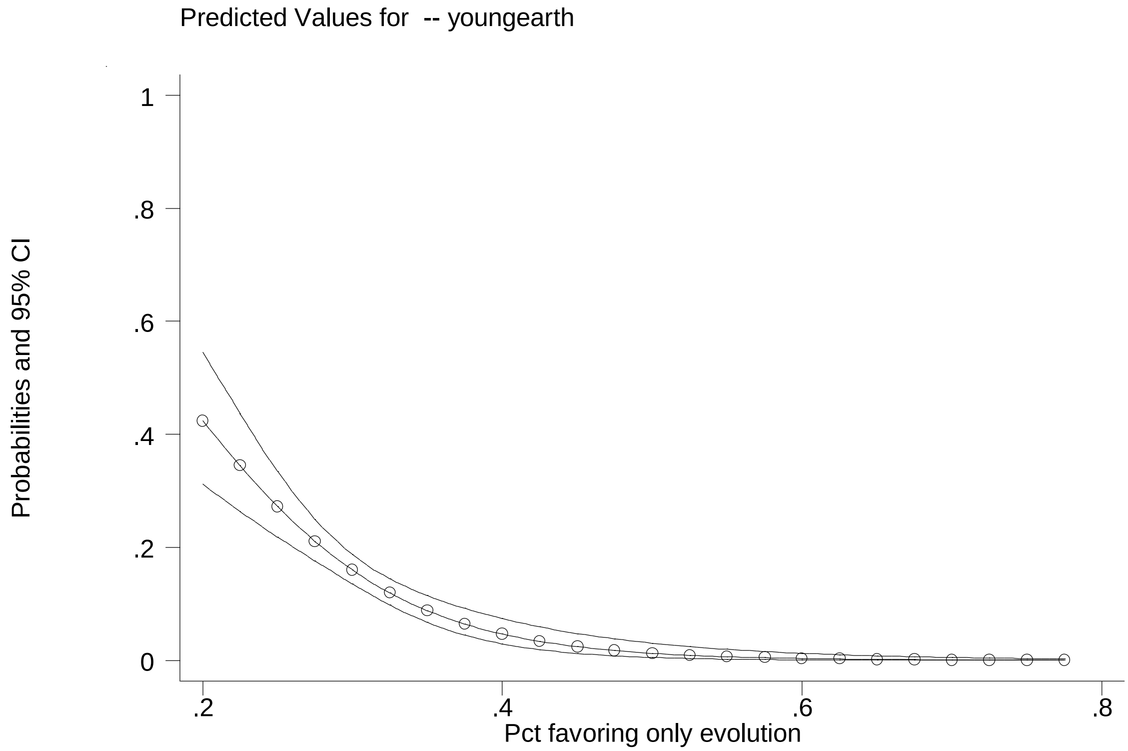
	B	Robust Std Err	t	p
Local support for evolution	0.25	0.11	2.36	0.03
Rigor of state standards	0.07	0.03	2.88	0.01
Opinion x standards	-0.14	0.07	-2.07	0.05
Young earth creationist	-0.48	0.04	-11.62	0.00
Organic evolutionist	0.16	0.02	8.74	0.00

Number of biology credits (0-2)	-0.01	0.02	-0.48	0.63
Number of science degrees (0-2)	0.01	0.01	0.42	0.67
Completed evolution course	0.09	0.02	3.56	0.00
Constant	0.33	0.06	5.57	0.00
R^2	0.40			

The second step to demonstrating the plausibility of the assortative hiring model is to see if in fact public preferences are strongly associated with the types of teachers who end up in the district. To see this, we estimated a simple logistic regression model predicting the probability that a district's teacher believes in young earth creationism with only local public opinion as a predictor. The results appear in Figure 4, below and are quite dramatic. The graph plots the logit curve only across the actual observed range of public opinion and it is useful to recall that nearly all the school districts lie between 20% and 40% support for teaching only evolution (see Figure 2, which has an equivalent horizontal axis).

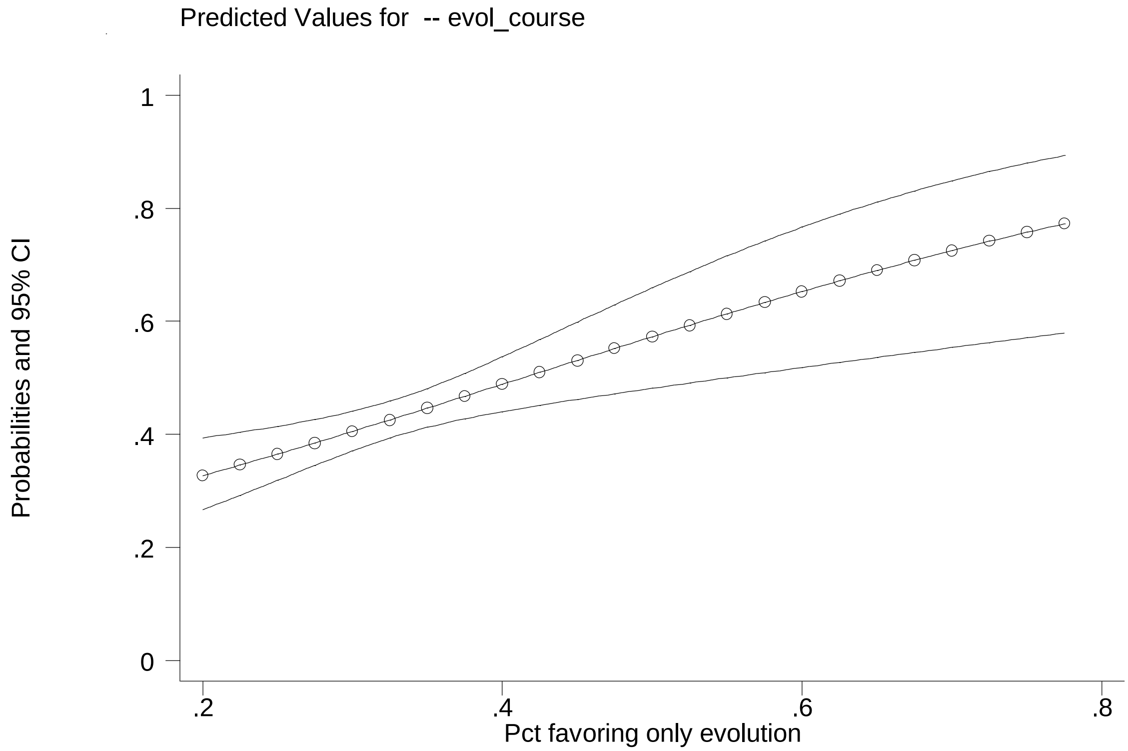
The graph shows both a strong relationship but also shows the limits of assortative hiring as a tool for local districts to fashion policy (as implemented) in accord with local preferences. Even in the most anti-evolution districts, young earth creationist teachers comprise slightly less than a majority of all high school biology teachers. Presumably, the supply of such teachers falls short of the demand and most very conservative districts must hire teachers who believe in evolution personally.

Figure 4: Predicted probability that a teacher believes in young earth creationism, by the public support for teaching only evolution in the school district.



Similarly, we examined the effect that opinion has on the qualifications of science teachers. In particular, we examined the probability of a teacher having completed an evolution course, a major predictor of their classroom conduct, as a function of local opinion. As shown in Figure 5, the likelihood goes up steadily as opinion becomes more accepting of evolution.

Figure 5: Predicted probability that a teacher completed a college-level course on evolution, by the public support for teaching only evolution in the school district.



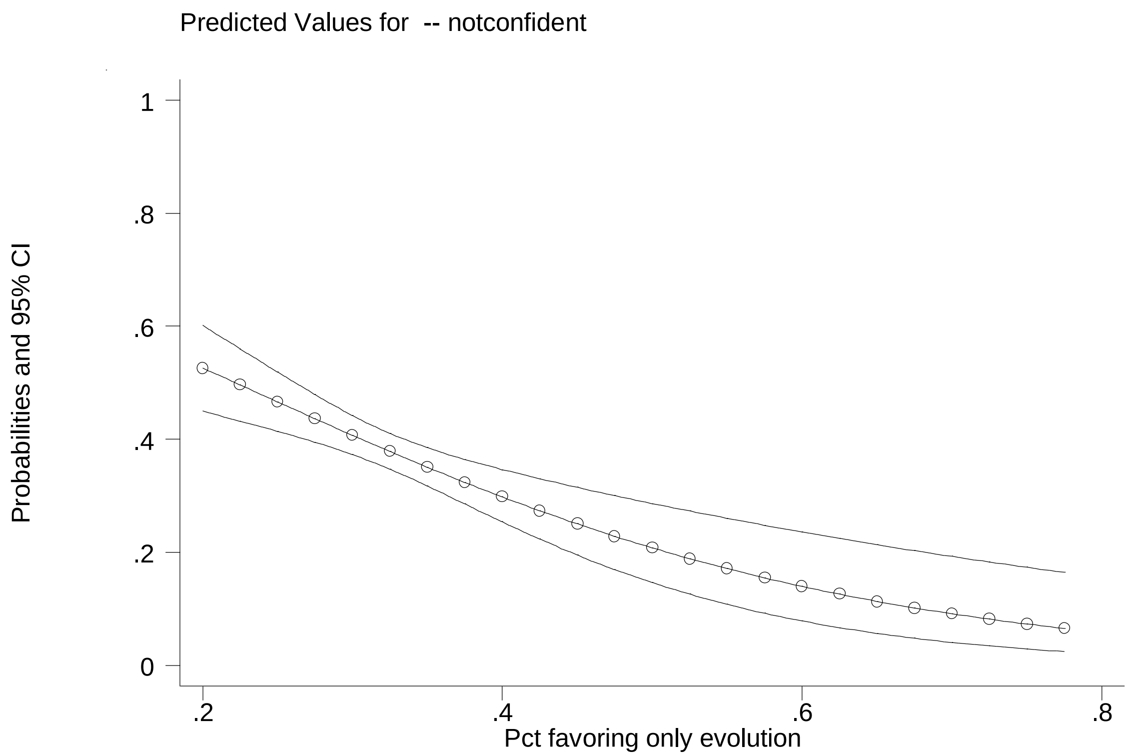
Finally, we examine one additional teacher trait that is related to local public opinion: the teacher's self-rated competence in on the topic of evolution. To assess this, we asked teachers the following self-rating question:

I would rate my knowledge of the scientific evidence bearing on the validity of evolutionary theory as:

- *Exceptional, on par with many college-level instructors (13%)*
- *Very good compared to most high school biology teachers (48%)*
- *Typical of most high school biology teachers (37%)*
- *I know less about this topic than many other high school biology teachers (2%)*

As can be seen, the question produces a "Lake Wobegon" effect as 61% rated their knowledge as "above average" or "exceptional" and only 2% rated themselves below average. For our purposes we will simply dichotomize this and examine the distribution of teachers rating themselves as only average or below average (those below the 40^{th} percentile). Figure 6 is similar to the previous two analyses except that here we statistically control for teachers' personal beliefs: the graph plots the predicted likelihood of teachers who rate themselves only average, setting their beliefs in the middle category (theistic evolution). The analysis shows that these lower confidence teachers are especially likely to end up teaching in extremely conservative districts. Since teachers lacking in confidence are prone to avoid evolution (Griffith and Brem), this pattern likely contributes to cursory teaching of evolution in conservative districts even when teachers are not themselves young-earth creationists.

Figure 6: Predicted probability that a teacher's self rating of evolution competence is below the 40th percentile (controlling for personal beliefs).



Discussion

Competition and contentious politics can be expected to surround any public policy for which responsibility is shared between different levels of government. From a top-down perspective, federalism represents a complication or a principle-agent *problem.* But from a bottom up perspective, the regulations, guidelines, incentives and sanctions developed by higher level governments are impediments to local rule. This is true of scores of policies, such as those concerning policing, environmental safety, transportation , or the conduct of elections. But there is no public policy that is more associated with local control than public education.

In this paper, we examined local policy as it is actually enacted by teachers – the street level bureaucrats entrusted with implementing educational guidelines. The teachers surveyed represent a random sample of teachers from each of the selected districts, and they function in this research design as indicators of how the policy is implemented in each locale. With most districts having only a single sampled teacher, the reliability of our measure is low. In spite of this, we found a number of crisp relationships in the data.

We have shown that local preferences concerning the desired balance between creationism and evolutionary biology are often at odds with state content standards. But in the absence of high stakes assessment tests, local public opinion is an important influence on how the policy is actually implemented. High stakes assessments give the standards a bit more bite, but in assortative hiring school districts have an important weapon in countering this state level "obstruction."

Districts tend to hire teachers who share the beliefs and values of district residents, and who therefore teach in accord with those beliefs. This can restore a degree of local control over the policy as it is actually implemented. But teachers can be a double edged sword. Their substantial autonomy and their documented tendency to teach in a manner consistent with their own *personal* values and beliefs make them difficult to monitor and sanction, something no doubt facilitated by both diffuse supervision and high job security. Nevertheless, teachers tend to wind up teaching in school districts whose public opinion is consistent with their own. Assortative hiring (like assortative mating) is an imperfect process of matching, but whether driven by the desires of school districts, of the teachers, or a combination, is a powerful engine of policy responsiveness at the local level.

Although evolution is somewhat unique as a policy area – after all, few other policy domains are alive in the popular imagination because of dramatizations like *Inherit the Wind* – it shares many characteristics of other policies transcend multiple levels of government and leave it to street level bureaucrats to implement. We think it likely that whenever local public opinion and policy preferences are thwarted by top-down restrictions, local preferences may nevertheless find expression at the implementation stage by local civil servants who are part of the community – sharing its values and subject to the more subtle influences of norms and culture.

Appendix

National Survey of High School Biology Teachers

We acquired the names and addresses of a randomly selected group of two thousand teachers of biology in US public schools. Following the *Tailored Design Method* for mail surveys (Dillman 2000), we first sent all two thousand teachers a pre-notification letter explaining that they would be receiving a survey in 7-10 days. This letter explained the purpose of the survey and encouraged them to look for our mailing and complete the questionnaire. They were then sent a survey packet containing a questionnaire booklet, along with a cover letter, a postage-paid return envelope, and a two dollar bill to get their attention and to express our thanks for completing the survey. Two weeks later, they received a reminder postcard, and (if they had not yet responded) we next sent a replacement packet with a second questionnaire and another postage paid envelope. All respondents were also given a web address that would allow them to complete the survey online. We had valid email addresses for 1500 teachers and these teachers received two email reminders as well; the emails included links to the online version of the survey. Additional technical descriptions of the sampling design, how we acquired the names and addresses of teachers, and detailed analyses showing that the survey is broadly representative of all public school teachers, are reported in the Appendix to this chapter and in an earlier scientific paper we published in the journal *PLoS Biology* (Berkman, Plutzer, and Pacheco 2008, supplementary text S1).

Coding state content standards in 2007

Two undergraduates comfortable and familiar with basic biology and evolution were hired to code each state's biology or life sciences standards that applied to grades nine and ten.[13] On their first pass through the standards the coders answered four specific questions that were intended to assess the *prominence* of evolution in the curriculum (2 questions), the extent to which evolution served as the or one of several guiding themes, and degree to which standards were sufficiently specific to guide teacher behavior.[14] On two additional passes through the standards, coders then looked specifically for presence of eleven evolution *benchmarks*. Benchmark analysis identifies "an existing statement of academic content knowledge" (Swanson 2005) and then determines how closely the standards align with that content and we will use these codes to assess the validity of our measure.

---

[13] In some states, the expectations are for a broader range of grades. In every instance, we coded the most narrow range available that encompassed the 9th or 10th grade general biology class.

[14] After the coders completed their work one of the principle investigators looked for discrepancies between the two coders. When large discrepancies were found the PI substituted his independent judgment.

Our coders established the overall prominence of evolution in the life science or biology standards by first looking at whether evolution was mentioned at all in the standards and then the level of detail in its treatment.[15] These codes were then combined into a scale that ranged from a score of 1 for standards that do not mention evolution at all, to a maximum score of 9 for standards that both devote a major, detailed section of the standards to evolution and also mention evolution in other sections of the curriculum (for example, in sections on ecology or genetics. Intermediate scores reflected coders relative rankings of the level of detail. For example, standards that included a phrase or two, typically in a bullet point, received scores of 2.5, while standards that included a short, but stand-alone section on evolution might receive a score of 5. This approach allowed us to distinguish not only between the somewhat easy to code states of New York and Iowa or Illinois, but states like Texas, which also received fairly low scores on this dimension, but not as low as Iowa or Illinois. Texas lists thirteen bullet points of "science concepts"; one of these reads that "the student knows the theory of biological evolution".

Second, the coders evaluated the overall treatment of evolution as course's unifying theme on a five point scale. The differences here are quite striking. New York, for example, states clearly that "The theory is the central unifying theme of biology." New Hampshire, introduces its science standards with this statement, different from New York's but still quite strong in its placement of evolution as one of several key topics: "Science theories are not hunches or guesses, but all have been subjected to repeated testing and verification. All scientific theories are subject to change as new evidence comes to be accepted by all scientists. Students should develop an understanding of the basic theories that are foundational in science and which guide scientific understanding. A few current scientific theories are listed": Along with the Cell, the Big Bang, Atomic Theory, and Gravity Theory we find the Theory of Evolution.

Our coders used a five point scale to complete the phrase:

Overall, evolution appears to be…

1. The unifying theme for the biology/life sciences curriculum

2. One of several unifying themes for the biology/life sciences curriculum

3. A major topic whose treatment is comparable to the cell, heredity, or ecology

4. A smaller topic, given less emphasis than others such as the cell, heredity or ecology

5. An afterthought or very minor topic

---

[15] There was high agreement among the coders over whether evolution could be found in the standards at all, and whether it was in its own or in multiple sections (r=.85); there was somewhat less agreement on the secondary question of whether it could be described as cursory, in a few sentences, or very detailed (r=.53).

Third, our coders assessed the extent to which the standards left teachers broad discretion to avoid teaching evolution or make it difficult to skip without ignoring the standards altogether. The Iowa standards, for example, certainly do not prohibit a teacher from developing detailed and in-depth lesson-plans on evolution. The current standards, in fact, are not far from Iowa's early position as the only state not to develop science standards at all but to leave discretion entirely in the hands of local districts and teachers. Iowa's emphasis is the scientific process rather than specific content is not at all inconsistent with excellent instruction in all facets of evolutionary theory; in fact, national standards recognize the role evolution can play in teaching what the scientific process is all about. But an Iowa teacher so inclined could also easily skip evolution altogether, and it is this distinction our coders attempted to quantify. For each state they used a 1 to 4 scale to complete this sentence:

The standards for biology/life sciences as a whole…

1. *"Are sufficiently specific and detailed to ensure that teachers will have to cover evolution"*
2.
3.
4. *"Are sufficiently vague that a teacher could choose not to cover evolution at all.*"

Coders could use the numbers "2" and "3" to introduce intermediate evaluations between these two assessments. Inter-coder reliability, as estimated by a simple correlation coefficients, was 0.72.

Overall, the three scores – extent of coverage, treatment as major/unifying theme, specificity in guidance to teachers – are highly correlated -indicating that they are all indicators of the same construct. We created a composite scale that is standardized (mean of zero, standard deviation of one) and measures the rigor with which evolution is addressed in the state standards. High scores indicate states whose content standards address evolution in detail, identify evolution as a major theme, and provide very specific guidance to teachers on the specific topics that students are expected to learn. The scale has an estimated reliability (Cronbach's alpha) of 0.88.

<u>Scale Validity</u>

Reliability is necessary for a measure to be a valid indicator for the concept of interest, but not all reliable measures are valid (Bohrnstedt 1983). To assess the validity of the scale—to assure that we were capturing the rigor of 9th and 10th grade biology standards – we undertook additional analyses. First, we looked to see if our rankings were positively correlated with Lerner's widely used (2000) ratings. As noted earlier, since Lerner's evaluation, twenty five states have significantly changed their standards. We defined our measure somewhat differently from Lerner and focus only on a subset of the Lerner's content domain (we focus only on grades 9 and 10, compared with the entire K-12 curriculum; and we do not consult content for the sister disciplines of earth science, physics and chemistry, focusing only on biology). Nevertheless, a valid measure should share considerable content and should correlate with Lerner's assessments for the 25 states in which we are evaluating the same standards document. This is exactly what we find: for the twenty-five states that *did not* change standards between Lerner's coding and ours the measures correlate at .65 (significant at .01), while the twenty five states that *did* change are not significantly correlated (r = .24).[16]

As a second test of our measure's validity we compare our scores with benchmarks coded by us and independently by other scholars. Benchmark analysis is different—rather than capturing the overall tone, organization, and detail of the standards it emphasizes particular content that is determined to be important. But the benchmarks are not fully independent of our other coding as it unlikely any of the standards would be considered especially strong without touching on the content in most benchmark studies.

We identified eleven benchmarks, closely following those used by Swanson (2005) and Skoog and Bilica (2001), both of whom drew heavily on the National Research Council's National Standards for Science Education (1996). Our coders went through all fifty states twice, in random order, and coded a benchmark as "achieved" when either coder determined that the basic benchmark content was present in the standards as something students should learn. The eleven benchmarks are:

The eleven benchmarks are:

1. Species evolve over time

2. Speciation

3. Diversity of organisms from evolution

4. Humans share characteristics with other concurrent living things because of shared ancestors

5. Mutations create the variation that changes an organism's offspring

6. Organisms are classified into a hierarchy of groups and subgroups based on similarities

7. Extinction of species is common

---

[16] Gross (Gross et al. 2005) used a coarse, four-category rating system of standards. In spite of the less precise measurement, the fact that many states had changed their standards, and that our analysis was limited to two grades we find the two ratings correlated at a level of r = 0.37.

8. Extinction occurs when the environment changes and the adaptive characteristics of a species are insufficient to allow its survival

9. Biological evolution accounts for the diversity of species

10. Behaviors have evolved through natural selection

11. Natural selection and its evolutionary consequences provide a scientific explanation for the fossil record of ancient life forms

Estimating local public opinion

We generated estimates of public opinion at the school district level in a two step process. First, we used Bayesian multilevel modeling with imputation and post-stratification to estimate public opinion for each of the fifty states. We then used Garand's "top down" regression method of estimating the opinion in school districts.

Step 1: Estimating state-level public opinion: We utilize a technique is known by the rather long name "Multilevel Modeling with Imputation and Post-Stratification," or the briefer "Multilevel Regression with Poststratification" or Mr.P. It was introduced by Gelman and Little (1997) and has been validated and used by many political scientists since (Park, Bafumi and Gelman 2004; Park, Gelman, and Bafumi 2006; Pacheco 2008; Lax and Phillips 2009). In an earlier book, we used a variation of this technique to validly measure public opinion for roughly 10,000 school districts (Berkman and Plutzer 2005). Others have shown that these techniques can reduce margins of error by 30-40% and reduce overall error (including bias) by as much as 50% (Park et al. 2006, Pacheco 2008) in comparison to the older aggregation method.

To estimate public opinion for each of the fifty states and the District of Columbia, we utilized data from every national poll or academic survey that met three conditions: (1) the survey contained a standard question asking specifically about teaching evolution, (2) the survey recorded the state of residence of each respondent, and (3) the original data records were available to us to analyze. In total, we were able to utilize nine different studies from 1998 through 2005 that included 9,533 respondents. The specific surveys and questions used in the estimation procedure are documented in Berkman and Plutzer (2010, Appendix section A3.8). For each survey, we recoded the question about evolution so that respondents who supported teaching only evolution (about 25%, depending on the question wording) could be compared to all others who expressed an opinion.

We then estimated a multilevel logistic regression model that allowed us to get a preliminary estimate of the average support in each state. This model also included four demographic variables – education, age, race, and sex and the effects of these variables were modeled as well. This permitted us to estimate the support for teaching only evolution for 64 different types of citizens in each state. For example, within the state of California, we not only estimate the average opinion, but also the opinion of a Californian who is female, black, a college graduate and under thirty years old. In fact, because the logistic regression slopes and state intercepts each are estimated with uncertainty (each has a margin of error), we did each estimate 1000 times to simulate random draws from sampling distributions. Thus for each of the 64 types of people in California, we have 1000 different estimates. These estimates are then combined using post-stratification weights derived from the United State Census giving us 1000 estimates of opinion in each state. We report the mean score from this simulation, a step typically called *imputation*: hence the moniker of Multilevel Models with Imputation and Post-Stratification (MLM-IPS).

Step 2: Top down estimation: Our next step was to estimate a state level prediction model for aggregate public opinion. Our state model yielded the following estimates:

$$Support\ for\ evolution = .26\ - .12(Evangelicals)\ +\ 1.07(Advanced\ Degrees)$$

Where:

> ***Support for evolution*** is our MLM-IPS estimate for the percentage of the state supporting the teaching of only evolution.
>
> ***Evangelicals*** is the number of adherents to doctrinally conservative Protestant churches (colloquially referred to as Evangelicals or Fundamentalists) expressed as a proportion of the total population.
>
> ***Advanced degrees*** is the number of adults, 25 and older, holding a masters or doctoral degree, expressed as a proportion of the over 25 adult population.

This equation was then used to make out of sample predictions at the school district level (since religious membership data is only available at the county level, we use the county in which the school district is located to make this estimate. Coding of denominations into faith traditions employs the typology developed by Steensland et al. (2000). Details on all sources of data can be found in Berkman and Plutzer (2010, appendix 3 and appendix 8).

References

Airasian, Peter. 1988. "Symbolic Validation: The Case of State-Mandated, High-Stakes Testing." *Educational Evaluation and Policy Analysis* 4 (10): 301-313.

Archbald, Douglas A. and Andrew C. Porter. 1994. "Curriculum Control and Teachers' Perceptions of Autonomy and Satisfaction." *Educational Evaluation and Policy Analysis* 16 (1): 21-39.

Ardoin, Phillip J. and James Garand. 2003. "Measuring constituency ideology in U.S. House districts: A top-down simulation approach." *Journal of Politics* 65 (November): 1165-1189.

Bandoli, James H. 2008. "Do State Science Standards Matter? Comparing Student Perceptions of the Coverage of Evolution in Indiana and Ohio Public High Schools." *American Biology Teacher* 70 (4): 212-216.

Berkman, Michael B. and Eric Plutzer. 2009. "Public Opinion and the Teaching of Evolution in the American States." *Perspectives on Politics* 7 (September): 485-500.

Berkman, Michael B. and Eric Plutzer. 2010 (forthcoming). *Evolution and Creationism: The Battle for Control of America's Classrooms.* Cambridge: Cambridge University Press

Binder, Amy. 2002. *Contentious Curricula: Afrocentrism and Creationism in American Public Schools,* Princeton University Press.

Bishop, Beth and Charles W. Anderson. 1990. "Student Conceptions of Natural Selection and its Role in Evolution." *Journal of Research in Science Teaching* *Volume* 27 (5): 415- 427.

Bohrnstedt, George. W. 1983. "Measurement." Chapter 3 in Rossi, Peter H., James D. Wright, and Andy. B. Anderson, eds., *Handbook of Survey Research*. New York: Academic Press. (69-121)

Boyd, Donald, Hamilton Lankford, Susanna Loeb and James Wykoff. 2005. "The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools." *Journal of Policy Analysis and Management* 24 (1): 113-132.

Brehm, John and Scott Gates. 1997. *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public*. Ann Arbor: University of Michigan Press.

Carlsen, William S. 1991. "Effects of New Biology Teachers Subject-Matter Knowledge on Curricular Planning." *Science Education* 75 (6): 631-647.

Dillman, Donald. 2000. *Mail and Internet Surveys: The Tailored Design Method.* New York: John Wiley.

Dobzhansky, Theodosius. 1973. "Nothing Makes Sense Except in the Light of Evolution." *American Biology Teacher* 35 (March): 125-129.

Gelman, Andrew and Thomas C. Little. 1997. "Poststratification into Many Categories Using Hierarchical Logistic Regression." *Survey Methodologist* 23 (December): 127-135.

Goldstein, Lisa. 2008. "Kindergarten Teachers Making 'Street-Level' Education Policy in the Wake of No Child Left Behind." *Early Education and Development* 19 (3): 448-478.

Griffith, Joyce A. and Sara K. Brem. 2004. "Teaching Evolutionary Biology: Pressures, Stress, and Coping." *Journal of Research in Science Teaching* 41 (8): 791–809.

Humes, Edward. 2007. *Monkey Girl: Evolution, Education, Religion, and the Battle for America's Soul*. New York: Ecco Press.

Ingram, Ella L. and Craig E. Nelson. 2005. "Relationship Between Achievement and Students' Acceptance of Evolution or Creation in an Upper-Level Evolution Course." *Journal of Research in Science Teaching.* 43 (1): 7-24.

Keiser, Lael. Forthcoming. "Understanding Street-Level Bureaucratic Decision Making: Determining Eligibility in the Social Security Disability Program." *Public Administration Review.*

Keiser, Lael R. 1999. "State Bureaucratic Discretion and the Administration of Social Welfare Programs: The Case of Social Security Disability." *Journal of Public Administration Research and Theory* 9 (1): 87-106.

Keiser, Lael R. and Joe Soss. 1998. "With Good Cause: Bureaucratic Discretion and the Politics of Child Support Enforcement." American Journal of Political Science. 42 (4): 1133-1156.

Larson, Edward J. 1989. *Trial and Error: The American Controversy over Creation and Evolution, updated edition*. Oxford: Oxford University Press

Lawson, Antono E. and John Weser. 1990. "The Rejection of Nonscientific Beliefs about Life: Effects of Instruction and Reasoning Skills." *Journal of Research in Science Teaching* 27 (6): 589 – 606.

Lax, Jeffrey R. and Justin H. Philips. 2009. "How Should We Estimate Public Opinion in the States?" *American Journal of Political Science* 53 (1): 107–121.

Lerner, Lawrence. 2000b. *Good Science, Bad Science: Teaching Evolution in the States*. Washington, DC: Thomas B. Fordham Foundation.

Lipsky, Michael. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services.* New York: Russell Sage Foundation.

Maynard-Moody, Steven and Michael Musheno. 2003. *Cops, Teachers, Counselors.* Ann Arbor: University of Michigan Press.

McCubbins, Mathew D., Roger Noll, and Barry R. Weingast. 1987. "Administrative Procedures as Instruments of Political Control." *Journal of Law, Economics, and Organization* 3 (2): 243-277.

Meier, Kenneth J. and Laurence J. O'Toole, Jr. 2006. *Bureaucracy in a Democratic State: A Governance Perspective*. Baltimore: Johns Hopkins

Meyers, Marcia. K. and Susan Vorsanger. 2003. "Street-Level Bureaucrats and the Implementation of Public Policy." Pp. 245-259 in B. Guy Peters and Jon Pierre, eds., *Handbook of Public Administration*. New York: Sage Publications.

Moore, Randy. 2002a. "Teaching Evolution: Do State Standards Matter?" *BioScience* 52 (4): 378-381.

Moore, Randy. 2002b. *Evolution in the Courtroom*. Santa Barbara: ABC-CLIO.

Mosher, Frederick C. 1982. *Democracy and the Public Service, 2nd edition.* New York: Oxford University Press.

National Research Council. 1996. *National Science Education Standards*. Washington, DC: National Academy Press. National Research Council. 1996. *National Science Education Standards*. Washington, DC: National Academy Press.

National Academy of Sciences (NAS). 2008a. *Science, Evolution and Creationism*. Washington, DC: National Academy Press.

National Science Teachers Association. 2003. "NSTA Position Statement: The Teaching of Evolution." http://www.nsta.org/about/positions/evolution.aspx accessed on November 15, 2008.

National Science Board. 1982. *Today's Problems: Tomorrow's Crises*. Washington: U.S. Government Printing Office.

Pacheco, Julianna Sandell. 2008. "Dynamic Public Opinion across the States." Paper presented at the annual State Politics and Policy Conference, May 30-31, 2008 in Philadelphia, PA.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2006. "State-Level Opinions from National Surveys: Poststratification Using Multilevel Logistic Regression." Pp. 209-228 in Jeffrey E. Cohen, ed., *Public Opinion in State Politics*. Stanford, California: Stanford University Press.

Percival, Garrick L., Martin Johnson and Max Neiman. 2009. "Representation and Local Policy: Relating County-Level Public Opinion to Policy Outputs." *Political Research Quarterly* 62 (March): 164-177.

Placier, Margaret, Michael Walker, and Bill Foster. 2002. "Writing the 'Show-Me' Standards: Teacher Professionalism and Political Control in U.S. State Curriculum Policy." *Curriculum Inquiry* 32 (3): 281-310.

Plutzer, Eric and Michael B. Berkman. 2008. "Evolution, Creationism, and the Teaching of Human Origins in Schools." *Public Opinion Quarterly* 72 (3): 540-553.

Reenock, Christopher and Brian Gerber. 2008. "Information Exchange and Interest Group Enfranchisement through Agency Design." *Journal of Public Administration Research and Theory* 18 (3): 415-440.

Rutledge, Michael L. and Melissa A. Warden. 2000. "Evolutionary Theory, the Nature of Science & High School Biology Teachers: Critical Relationships." *American Biology Teacher* 62 (1): 23-31.

Sabatier, Paul A., John Loomis and Catherine McCarthy. 1995. "Hierarchical Controls, Professional Norms, Local Constituencies, and Budget Maximization: An Analysis of U.S. Forest Service Planning Decisions." *American Journal of Political Science* 39 (Feb.): 204-242.

Scholz, John T. and FengHeng Wei. 1986. "Regulatory Enforcement in a Federalist System." *American Political Science Review* 80 (Dec): 1249-1270.

Skoog, Gerald and Kimberly Bilica. 2002. "The Emphasis Given to Evolution in State Science Standards: A Lever for Change in Evolution Education?" *Science Education*. 86 (4): 445-462.

Smith, Steven Rathgeb. 2003. "Street Level Bureaucracy and Public Policy." Pp. 354-376 in B. Guy Peters and Jon Pierre, eds., *Handbook of Public Administration*. New York: Sage Publications.

Soss, Joe. 2000. *Unwanted Claims: The Politics of Participation in the U.S. Welfare System*. Ann Arbor: University of Michigan Press.

Spillane, James P., and Karen A. Callahan. 2000. "Implementing State Standards for Science Education: What District Policy Makers Make of the Hoopla." *Journal of Research in Science Teaching* 37 (5): 401-425.

Steensland, Brian, Jerry Z. Park, Mark D. Regnerus, Lynn D. Robinson, W. Bradford Wilcox, and Robert D. Woodberry. 2000. "The Measure of American Religion: Toward Improving the State of the Art." *Social Forces* 79 (1): 291-318.

Swanson, Christopher B. 2005. "Evolution in State Science Education Standards." *Education Week Editorial Reports in Education Research Center.* November 2005. http://www.edweek.org/media/epe_evolution1105.pdf [accessed July 9, 2009]

Weatherley, Richard, and Michael Lipsky. 1978. "Street-level Bureaucrats and Institutional Innovation: Implementing Special Education Reform." *Harvard Educational Review* 47 (May): 171-197.

Weissert, Carol S. 1994. "Beyond the Organization: The Influence of Community and Personal Values on Street-Level Bureaucrats' Responsiveness." *Journal of Public Adminstration Research and Theory* 4 (2): 225-254.

Wilson, David Sloan. 2005. "Evolution for Everyone: How to Increase Acceptance of, Interest in, and Knowledge about Evolution." *PLoS Biology* 3 (12): e364: 2058-2064.

Whitford, Andrew. 2002. "Decentralization and Political Control of the Bureaucracy." *Journal of Theoretical Politics* 14 (2): 167-193.